# การประเมินทักษะการรู้สารสนเทศ สำหรับนักเรียนระดับมัธยมศึกษา

# Assessing Information Literacy Skill for High School Student. (ILS)

**นายเจนรบ โกรธา 595050089-7**

**สาขาวิชาการวัดและประเมินผลการศึกษา ชั้นปีที่ 1 ระดับปริญญาโท**

**อาจารย์ผู้สอน**

**ผู้ช่วยศาสตราจารย์ ดร. พัชรี จันทร์เพ็ง**

**สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น**


**รายวิชา 217709 ทฤษฎีการวัดและการทดสอบ**

**สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น**

# การประเมินทักษะการรู้สารสนเทศ สำหรับนักเรียนระดับมัธยมศึกษา

# Assessing Information Literacy Skill for High School Student. (ILS)

**นายเจนรบ โกรธา 595050089-7 สาขาวิชาการวัดและประเมินผลการศึกษา ระดับปริญญาโท**
**ผู้ช่วยศาสตราจารย์ ดร. พัชรี จันทร์เพ็ง สาขาวิชาการวัดและประเมินผลการศึกษา**
**คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น**
**3 ธันวาคม 2559 (1/2559) ครั้งที่ 1**
**10 ธันวาคม 2559 (1/2559) ครั้งที่ 2 (สอบปากเปล่า)**
**17 ธันวาคม 2559 (1/2559) ครั้งที่ 3 ส่งฉบับสมบูรณ์**

## บทคัดย่อ (Abstract)

ประเทศไทยต้องการเน้นให้นักเรียนเกิดการเรียนรู้ มีความรู้ และมีทักษะที่เกี่ยวกับการรู้สารสนเทศ จึงมีกระบวนการจัดการเรียนรู้ที่หลากหลาย รวมทั้งวิธีการวัดและประเมินผลเองก็หลากหลายตามมาด้วยเช่นกัน และจากการศึกษาเกี่ยวกับการทำแบบทดสอบมาตรฐานเพื่อวัดและประเมินผลทักษะการรู้สารสนเทศ พบว่ายังไม่มี การทำแบบทดสอบ หรือแบบประเมินมาตรฐานในประเทศไทย ดังนั้นผู้วิจัยจึงประสงค์จะทำการพัฒนาแบบประเมินมาตรฐาน ทักษะการรู้สารสนเทศ ภาษาไทย โดยยึดตามเกณฑ์มาตรฐานของสมาคมห้องสมุดวิทยาลัย และวิจัยแห่งสหรัฐอเมริกา (Association of College and Research Libraries- ACRL) โดยยึดหลักมาตรฐานการรู้สารสนเทศสำหรับนักเรียน (Information Literacy Standards for Schools) เนื่องจากมาตรฐานดังกล่าว มีความละเอียดครอบคลุมในหลักการรู้สารสนเทศ และเป็นมาตรฐานกำหนดใช้สำหรับระดับมัธยมศึกษาโดยเฉพาะ

โดยมีวัตถุประสงค์เพื่อพัฒนาและตรวจสอบคุณภาพกรอบการประเมินการความก้าวหน้าในการเรียนรู้ (Learning Progression) ทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษา และ เพื่อประเมินการความก้าวหน้าในการเรียนรู้ (Learning Progression) ทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษา โดยอาศัยแนวความคิดเกี่ยวกับ Four Building Blocks ได้อธิบายถึง Construct Modeling จาก The four Building blocks**.** โดยเริ่มต้นจากการอธิบายในภาพรวมและแนวคิดหลักก่อน จากนั้นอธิบายความหมายของการวัด ว่าต้องมีการกำหนดประเภทของการสังเกต คุณลักษณะของประชากรเองส่งผลต่อคุณลักษณะของการวัดด้วย การกำหนดคุณลักษณะของกลุ่มประชากรจะต้องสอดคล้องกับความจริงที่ต้องการวัด เพื่อให้ได้กระวนการวัดที่สมบูรณ์ วัถตุประสงค์หลักของการวัด ก็เพื่อ เป็นการอธิบาย เตรียมการอย่างมีเหตุผล

จากการศึกษาพบว่ากรอบการประเมินการความก้าวหน้าในการเรียนรู้ (Learning Progression) ทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษา พบว่ามีความน่าเชื่อมั่น (EAP/PV RELIABILITY: 0.864) มีความเชื่อมั่นอยู่ระดับ Coefficient Alpha**:** 0.88 แบบประเมินความก้าวหน้าในการเรียนรู้ (Learning Progression) ทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษา จำนวนทั้งหมด 9 ข้อ พบว่าข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้จำนวนทั้งหมด 9 ข้อ

คำสำคัญ: การประเมิน, ทักษะการรู้สารสนเทศ, การรู้สารสนเทศสำหรับนักเรียนมัธยมศึกษา
Keyword: Evaluation, Literacy Skill, Literacy Skill for High School Student.

**1. ที่มาและความสำคัญ (Introduction)**

ในการเตรียมนักเรียนให้พร้อมกับชีวิตในศตวรรษที่ 21 เป็นเรื่องสำคัญของกระแสการปรับเปลี่ยนทางสังคมที่เกิดขึ้นในศตวรรษที่ 21 ส่งผลต่อวิถีการดำรงชีพของสังคมอย่างทั่วถึง ครูจึงต้องมีความตื่นตัวและเตรียมพร้อมในการจัดการเรียนรู้เพื่อเตรียมความพร้อมให้นักเรียนมีทักษะสำหรับการออกไปดำรงชีวิตในโลกในศตวรรษที่ 21 ที่เปลี่ยนไปจากศตวรรษที่ 20 และ 19 โดยทักษะแห่งศตวรรษที่ 21 ที่สำคัญที่สุด คือ ทักษะการเรียนรู้ (Learning Skill) ส่งผลให้มีการเปลี่ยนแปลงการจัดการเรียนรู้เพื่อให้เด็กในศตวรรษที่ 21 นี้ มีความรู้ความสามารถ และทักษะจำเป็น ซึ่งเป็นผลจากการปฏิรูปเปลี่ยนแปลงรูปแบบการจัดการเรียนการสอนตลอดจนการเตรียมความพร้อมด้านต่างๆ

ทักษะแห่งศตวรรษที่ 21 (21st Century Skills) วิจารณ์ พานิช (2555: 16-21) ได้กล่าวถึงทักษะเพื่อการดำรงชีวิตในศตวรรษที่ 21 ดังนี้ สาระวิชาก็มีความสำคัญ แต่ไม่เพียงพอสำหรับการเรียนรู้เพื่อมีชีวิตในโลกยุคศตวรรษที่ 21 ปัจจุบันการเรียนรู้สาระวิชา (Content หรือ Subject matter) ควรเป็นการเรียนจากการค้นคว้าเองของศิษย์ โดยครูช่วยแนะนำ และช่วยออกแบบกิจกรรมที่ช่วยให้นักเรียนแต่ละคนสามารถประเมินความก้าวหน้าของการเรียนรู้ของตนเองได้

ทักษะด้านสารสนเทศ สื่อ และเทคโนโลยี เป็นอีกทักษะแห่งศตวรรษที่ 21 ที่มีความสำคัญ เนื่องด้วยในปัจจุบันมีการเผยแพร่ข้อมูลข่าวสารผ่านทางสื่อและเทคโนโลยีมากมาย ผู้เรียนจึงต้องมีความสามารถในการแสดงทักษะการคิดอย่างมีวิจารณญาณและปฏิบัติงานได้หลากหลาย โดยอาศัยความรู้ในหลายด้าน คือ ความรู้ด้านสารสนเทศ ความรู้เกี่ยวกับสื่อ ความรู้ด้านเทคโนโลยี

ในสังคมปัจจุบันซึ่งเป็นสังคมแห่งสารสนเทศ บุคคลในสังคมจำเป็นต้องรับข้อมูลข่าวสารอย่างท่วมท้น บุคคลทุกคนจำเป็นต้องมีการพัฒนาตนเอง เพื่อรับรู้ข้อมูลข่าวสารได้อย่างเหมาะสมและถูกต้อง คนในสังคมปัจจุบันจึงต้องมีการเรียนรู้อยู่ตลอดเวลา เพื่อการเท่าทันในข้อมูลข่าวสารที่หลากหลาย สังคมปัจจุบันจึงเป็นสังคมแห่งการเรียนรู้อย่างต่อเนื่อง การศึกษาต้องมุ่งเพิ่มขีดความสามารถและโอกาสในการเรียนรู้ของบุคคลและพัฒนาทักษะการรู้สารสนเทศให้แก่บุคคลในสังคมอย่างสม่ำเสมอ เพื่อสามารถนำความรู้ไปใช้ในสังคมได้อย่างยั่งยืน นักวิชาการหลายท่านได้ให้ความหมายของคำว่า การรู้สารสนเทศครอบคลุม การมีพฤติกรรมเข้าถึงสารสนเทศที่เหมาะสมอย่างมีความคิดและจริยธรรม โดยผ่านช่องทางหรือสื่อใดๆ เพื่อให้ได้มาซึ่งสารสนเทศตามความต้องการ (SUNY Council of Library Directors Information Literacy Innitiative, 2003)

การรู้สารสนเทศมีความเกี่ยวข้องกับการศึกษา ดังที่ได้กล่าวมาแล้ว โดยมี ผู้เรียนเป็นศูนย์กลางในกระบวนการเรียนรู้ ทั้งนี้เพราะการศึกษามีวัตถุประสงค์ในการพัฒนามนุษย์ ให้มี ความเท่าทันต่อการเปลี่ยนแปลง และเป็นผู้นำในสังคม ดังนั้นประเทศต่าง ๆ จึงได้มีการพัฒนามาตรฐานโดยสถาบัน องค์กรของรัฐที่เกี่ยวข้อง เช่น ประเทศสหรัฐอเมริกา สมาคมห้องสมุด (American Library Association-ALA) ได้จัดทำมาตรฐานการรู้ สารสนเทศ สำหรับนักเรียน (Information Literacy Standards for Schools) ในปี 1998 และ มาตรฐานการรู้สารสนเทศสำหรับ นักศึกษาในระดับอุดมศึกษา (Information Literacy Standards for Higher Education) ในปี 2000 เพื่อใช้เป็นแนวทางในการพัฒนาการเรียนการสอน และในประเทศอังกฤษ Society of College, National and University Libraries ได้มการกำหนดมาตรฐาน เรียกว่า Seven Pillars of Information Literacy เป็นต้น (Bruce, 2002) นอกจากนั้นเพื่อให้สามารถทราบถึงระดับทักษะการรู้สารสนเทศ จึงได้มีการพัฒนาแบบทดสอบมาตรฐานทั้งในการศึกษาระดับมัธยม และอุดมศึกษา

ในระดับโรงเรียน สมาคมห้องสมุดโรงเรียนแห่งสหรัฐอเมริกา (American Association of School Libraries : AASL, 1998) นับว่ามีบทบาทสำคัญในการจัดทำหลักสูตรการรู้สารสนเทศเพื่อใช้ เป็นมาตรฐานในการสอนทักษะการรู้สารสนเทศแก่นักเรียนระดับประถมศึกษาและมัธยมศึกษา เพื่อสอนให้นักเรียนเป็นผู้รู้สารสนเทศซึ่งเป็นพื้นฐานสำคัญของการเรียนรู้ตลอดชีวิต มาตรฐานการรู้ สารสนเทศสำหรับนักเรียน 9

มาตรฐานโดยแบ่งเป็น 3 หมวดได้แก่ หมวดความรู้ทางสารสนเทศ หมวดการเรียนแบบพึ่งพาตนเอง และหมวดการมีความรับผิดชอบต่อสังคม นักเรียนที่เป็นผู้รู้สารสนเทศใน ระดับโรงเรียนจะต้องมีคุณลักษณะตามคุณลักษณะดังต่อไปนี้ (AASL, 1998อ้างถึงในศิวราช ราชพัฒน์, 2547 : 19)

การรู้สารสนเทศมีบทบาทและความสำคัญต่อการศึกษาทุกระดับ ตั้งแต่ระดับประถมศึกษา ระดับมัธยมศึกษา และระดับอุดมศึกษา เนื่องจากสถาบันการศึกษาต่างๆ ต่างตระหนักถึงความสำคัญของการรู้สารสนเทศ ว่าเป็นพื้นฐานที่นำไปสู่การเรียนรู้ด้วยตนเองและการเรียนรู้ตลอดชีวิต ซึ่งสอดคล้องกับกรอบนโยบายเทคโนโลยีสารสนเทศ ระยะ พ.ศ.2554-2563 ของประเทศไทย" หรือ "ICT 2020" มีวิสัยทัศน์ว่า ประเทศไทยในปี 2563 ประชาชนมีความรอบรู้ เข้าถึง สามารถพัฒนาและใช้ประโยชน์ จากสารสนเทศได้อย่างรู้เท่าทันเกิดประโยชน์ต่อการเรียนรู้ การทำงาน และการดำรงชีวิตประจำวันและบุคลากร ICT มีความรู้ความสามารถและทักษะในระดับสากล และสอดคล้องกับพระราชบัญญัติการศึกษาแห่งชาติ พุทธศักราช 2542 เพราะการรู้สารสนเทศทำให้เกิดการเรียนรู้ตลอดชีวิตซึ่งเป็นกลไกนำพาให้บุคคลมีการพัฒนาคุณภาพของตนอยู่เสมอ และหากประเทศใดประชาชนมีการเรียนรู้ตลอดชีวิต ถือว่าทรัพยากรมนุษย์ของประเทศนั้นย่อมมีคุณภาพดีกว่าประเทศอื่นๆ และการรู้สารสนเทศยังเป็นวิธีแห่งการมีอำนาจของบุคคลในสังคมสารสนเทศอีกด้วย ดังนั้นประชากรที่เป็นผู้รู้สารสนเทศจึงถือได้ว่าเป็นทรัพยากรที่มีคุณค่ามากที่สุดของประเทศในยุคนี้

ทักษะการรู้สารสนเทศนั้นเป็นกระบวรการและทักษะ ที่มีความซับซ้อนในการวัดออกมา รวมถึงมีความยากที่จะวัดสิ่งที่เป็นกระบวนการออกมา นอกจากนี้เครื่องมือที่จะใช้วัดนั้นมีความสำคัญอย่างยิ่งต่อกระบวนการวัดของทักษะที่มีลักษณะที่มีโครงสร้างที่ซับซ้อนออกมา เพื่อให้สามารถวัดออกมาให้ครอบคลุมและมีประสิทธิภาพจึงจำเป็นต้องเลือกรูปแบบของการออกแบบและพัฒนาที่เหมาะสม

การศึกษาระบบการประเมินที่ได้รับการออกแบบและพัฒนาการสร้างระบบโดยใช้แบบจำลองที่เกี่ยวข้องกับการออกแบบการดำเนินงานและการประเมินผลของสี่กลุ่ม (The four building box) คือ Construct Maps, Item Design, Outcome Space และ Measurement Model.

โดยในประเด็นที่สนใจคือ Construct Maps ในส่วนที่เป็นการสร้างโครงสร้างทางความคิดสำหรับการคิดที่มีความซับซ้อน หลักการสำคัญของการสร้างแบบจำลองการสร้างคือการออกแบบการประเมิน ที่ต้องมีการพัฒนาโครงสร้างและแบบแผนเครื่องมือการประเมินที่สามารถตีความที่เกี่ยวข้องกับรูปแบบของกระบวนที่ทำให้มีความรู้ความเข้าใจ โดยรูปแบบส่วนใหญ่แล้วจะยังไม่มีความชัดเจนเท่าที่ควร ยังไม่มีแนวทางที่ชัดเจนเกี่ยวกับวิธีการสร้างแบบประเมินโครงสร้างความคิดที่ตรงตามโครงสร้างของข้อมูลเชิงลึกของผู้เชี่ยวชาญเชิงเนื้อหา แนวความคิดที่ซับซ้อนที่ว่าคือ แนวความคิดความพยายามที่จะจับการคิดด้านโดเมนของนักเรียน

ในประเด็นนี้หากเราสามารถที่จะออกแบบประเมินผลทางด้านความคิดของนักเรียน โดยใช้กระบวนการนี้ได้ โดยอาศัยความเข้าใจในการสร้างแบบจำลอง รูปแบบหรือโครงสร้างที่ซับซ้อนของการเข้าใจของนักเรียน เพื่อให้เกิดความชัดเจน มีรูปแบบการประเมินผลที่เป็นกระบวนการ วิธีการที่ยอมรับ ได้ผลเป็นข้อมูลเชิงลึก ก็จะทำให้เราได้เข้าใจถึงกระวนการคิดที่ซับซ้อน ส่งผลให้เกิดกระบวนการพัฒนาในรูปแบบต่างๆที่ตรงจุดมาขึ้น และมีความสอดคล้องมากขึ้นด้วย

ในความคิดแล้วหากเราสามารถสร้างรูปแบบหรือวิธีการ เพื่อให้ได้มาซึ่งวิธีหรือเครื่องมือซึ่งวัดประเมินความคิดที่ซับซ้อน โครงสร้างทางความคิดหรือวิธีการทำความเข้าใจของนักเรียนแล้ว ก็จะส่งผลให้เกิดรูปแบบการเรียนรู้ การจัดการเรียนการสอน รวมไปถึงวิธีการวัดประเมินผลความเข้าใจของนักเรียนได้ตรงจุดด้วย

ภาพรวมและแนวความคิดเกี่ยวกับ Four Building Blocks ได้อธิบายถึง Construct Modeling จาก The four Building blocks. โดยเริ่มต้นจากการอธิบายในภาพรวมและแนวคิดหลักก่อน จากนั้นอธิบายความหมายของการวัด ว่าต้องมีการกำหนดประเภทของการสังเกต คุณลักษณะของประชากรเองส่งผลต่อคุณลักษณะของการวัดด้วย การกำหนดคุณลักษณะของกลุ่มประชากรจะต้องสอดคล้องกับความจริงที่ต้องการ

วัด เพื่อให้ได้กระวนการวัดที่สมบูรณ์ วัถตุประสงค์หลักของการวัด ก็เพื่อ เป็นการอธิบาย เตรียมการอย่างมีเหตุผล และเป็นการสรุปผลลักษณะของคนอย่างรวดเร็วทั้งทางด้านความสำเร็จ ทัศนะคติ

สิ่งที่ได้หรือข้อสรุปจากแนวคิดจะมีความเข้าใจพื้นฐาน ลักษณะเฉพาะของเครื่องมือที่สามารถนำมาใช้ออกแบบการวัดได้หลายแบบ ทั้งแบบสำรวจ ข้อสอบ แบบสอบถาม ก็สามารถออกแบบการวัดได้หลากหลายลักษณะ หากต้องการประสบความสำเร็จในการวัดควรมีตัวอย่างที่กว้างและหลากหลายครอบคลุม

ส่วน Construct map เป็นการหาสิ่งที่ต้องการวัดและหารูปแบบหรือแนวทางที่จะนำไปสู่การออกแบบเครื่องการวัด ที่จีมีหลากหลายลักษณะ โดยจะต้องเลือกลักษณะที่เหมาะสม การให้ค่ามาตรา

ส่วน Item Design เป็นขั้นต่อมา เพื่อศึกษาทฤษฎี โครงสร้าง การพิสูจน์จากสภาพจริง เพื่อนำมาซึ่งการสร้างมาตรในการวัด การออกแบบแบบรายข้อ เพื่อใช้เป็นเครื่องมือ

ส่วน Outcome Space การตัดสินคำตอบจากเกณฑ์ ที่กำหนดขึ้น หรือเป็นการกำหนดเกณฑ์การให้คะแนนจากคำตอบ ของแต่ละประเภทที่แตกต่างกัน

ส่วนของ โมเดลการวัดผล เป็นส่วนที่จะต้องทำการตรวจสอบค่า ตรวจสอบกระบวนการวัดผลทั้งหมด รวมทั้งการหาคุณภาพของโมเดลการวัดนี้ด้วย

จะเห็นได้ว่าระบบการประเมินที่ได้รับการออกแบบและพัฒนาการสร้างระบบโดยใช้แบบจำลองที่เกี่ยวข้องกับการออกแบบการดำเนินงานและการประเมินผลของสี่กลุ่ม (The four building box) นั้นมีความเหมาะสมอย่างมากที่นำมาเป็นทฤษฎีในการออกแบบและพัฒนาการประเมินทักษะการรู้สารสนเทศ สำหรับนักเรียนระดับมัธยมศึกษา เนื่องด้วยเป็นการวัดและประเมินโครงสร้างทางความคิดสำหรับการคิดที่มีความซับซ้อน โครงสร้างของข้อมูลเชิงลึก ทั้งนี้กระบวนการพัฒนาด้วย The four building box นั้นเป็นหลักการแนวคิด และมีทฤษฎีที่มีความเหมาะสมและสอดคล้อง

จากการศึกษาพบว่าประเทศไทยต้องการเน้นให้นักเรียนเกิดการเรียนรู้ มีความรู้ และมีทักษะที่เกี่ยวกับการรู้สารสนเทศ จึงมีกระบวนการจัดการเรียนรู้ที่หลากหลาย รวมทั้งวิธีการวัดและประเมินผลเองก็หลากหลายตามมาด้วยเช่นกัน และจากการศึกษาเกี่ยวกับการทำแบบทดสอบมาตรฐานเพื่อวัดและประเมินผลทักษะการรู้สารสนเทศ พบว่ายังไม่มี การทำแบบทดสอบ หรือแบบประเมินมาตรฐานในประเทศไทย แต่จะมีพบมากในต่างประเทศ ดังที่ได้กล่าวในตอนต้น แต่เนื่องจากเป็นแบบทดสอบภาษาอังกฤษ ซึ่งอาจไม่เหมาะในการนำมาใช้ได้สอดคล้อง กับสภาพสังคมไทย และระดับทักษะการรู้สารสนเทศของนักศึกษาที่อาจแตกต่างกันในแต่ละประเทศ ตลอดจนแบบทดสอบที่ได้มีการพัฒนาเป็นแบบทดสอบที่ มีลิขสิทธิ์ และผู้ใช้ต้องเสียค่าใช้จ่ายในการนำมาใช้แต่ละครั้ง

ดังนั้นผู้วิจัยจึงประสงค์จะทำการพัฒนาแบบประเมินมาตรฐาน ทักษะการรู้สารสนเทศ ภาษาไทย โดยยึดตามเกณฑ์มาตรฐานของสมาคม ห้องสมุดวิทยาลัย และวิจัยแห่งสหรัฐอเมริกา (Association of College and Research Libraries- ACRL) โดยยึดหลักมาตรฐานการรู้สารสนเทศสำหรับนักเรียน (Information Literacy Standards for Schools) เนื่องจากมาตรฐานดังกล่าว มีความละเอียดครอบคลุมในหลักการรู้สารสนเทศ และเป็นมาตรฐานกำหนดใช้สำหรับระดับมัธยมศึกษาโดยเฉพาะ โดยใช้หลักการและแนวคิดของการออกแบบการดำเนินงานและการประเมินผลของสี่กลุ่ม (The four building box) ใช้ในการพัฒนา

**2. วัตถุประสงค์ (Objective)**

การวิจัยครั้งนี้มีวัตถุประสงค์ ดังต่อไปนี้

2.1 เพื่อพัฒนาและตรวจสอบคุณภาพกรอบการประเมินการความก้าวหน้าในการเรียนรู้ (Learning Progression) ทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษา

2.2 เพื่อประเมินการความก้าวหน้าในการเรียนรู้ (Learning Progression) ทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษา

**3. ขอบเขตการวิจัย**

**3.1 ประชากร** นักเรียนระดับมัธยมศึกษาตอนปลาย สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 25

**3.2 กลุ่มตัวอย่าง** นักเรียนระดับมัธยมศึกษาตอนปลาย โรงเรียนประชารัฐวิทยาเสริม ภาคเรียนที่ 2 ปีการศึกษา 2559 จำนวน 60 คน

**3.3 ตัวแปรที่ศึกษา** ทักษะการรู้สารสนเทศ (Information literacy Skill)

**3.5 เนื้อหาวิชา** รายวิชาเทคโนโลยีสารสนเทศ (คอมพิวเตอร์) เรื่อง ระบบเครือข่าย มาตรฐาน ง.3.1 ม.4-6/5 แก้ปัญหาด้วยกระบวนการเทคโนโลยีสารสนเทศอย่างมีประสิทธิภาพ

**3.4 พื้นที่ที่ทำการศึกษา** โรงเรียนประชารัฐวิทยาเสริม อำเภอพระยืน จังหวัดขอนแก่น

**3.5 ช่วงเวลาในการศึกษา** ปีการศึกษา 2/2559

**4. นิยามศัพท์เฉพาะ**

**การพัฒนาและตรวจสอบคุณภาพกรอบการประเมินความก้าวหน้าในการเรียนรู้ หมายถึง** กระบวนการสร้างและออกแบบการประเมินความก้าวหน้าในการเรียนรู้ การประเมินทักษะการรู้สารสนเทศสำหรับนักเรียนระดับมัธยมศึกษา

**ความก้าวหน้าในการเรียนรู้เรื่องทักษะการรู้สารสนเทศ หมายถึง** ทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษานั้นจะยึดบนหลักหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานในด้านสมรรถนะของผู้เรียน คุณลักษณะอันพึงประสงค์ รวมทั้งตัวชี้วัดซึ่งระบุสิ่งที่นักเรียนพึงรู้ที่ได้ระบุไว้ แต่ยึดหลักมาตรฐานตามสมาคมห้องสมุดโรงเรียนแห่งสหรัฐอเมริกา (American Association of School Libraries: AASL, 1998) หมวดคุณลักษณะของการรู้สารสนเทศ (Information literacy) และมาตรฐานที่ 2 นักเรียนที่เป็นผู้รู้สารสนเทศสามารถประเมินสารสนเทศได้อย่างมีหลักการ จำนวน 4 ตัวชี้วัด ได้แก่ 1) นักเรียนสามารถตัดสินได้ว่าสารสนเทศนี้นมีความถูกต้อง ตรงประเด็น และครอบคลุมใน สิ่งที่ต้องการ 2) นักเรียนรู้ถึงความแตกต่างระหว่างข้อเท็จจริงและข้อคิดเห็น 3) นักเรียนสามารถระบุถึงสารสนเทศที่คลาดเคลื่อนและไม่เป็นจริงได้ 4) นักเรียนสามารถเลือกสารสนเทศที่เหมาะสมลับปัญหาหรือข้อคำถามที่สงสัยได้

ความก้าวหน้าในการเรียนรู้เรื่องทักษะการรู้สารสนเทศ ได้แบ่งออกเป็น 5 ระดับ ได้แก่

**ระดับ 0 ไม่มีการตอบ (Non-response)** ไม่ตอบคำถาม ไม่ส่งข้อสอบ หรือไม่ได้เขียนคำตอบที่เกี่ยวกับคำถาม

**ระดับ 1 ความรู้ความจำ (Recall)** นักเรียนสามารถจำเนื้อหา หรือหลักการนั้นๆมาตอบคำถามได้ เช่นบอกความหมาย บอกขั้นตอน บอกวิธีการ แต่ไม่เข้าใจในหลักการแต่ละขั้นตอน ไม่สามารถระบุหรือทราบถึงว่า สามารถเอาสิ่งที่จำได้นั้นไปใช้ทำอะไรเพิ่มเติมได้ ใช้เพียงแค่ทำที่พบเห็นในเนื้อหา หรือเอกสารประกอบการสอนเท่านั้นมาตอบคำถามไม่สามารถบอกเหตุผล หรือวิธีการที่เลือกเอากระบวนการนั้นๆมาใช้หรือมาตอบ

**ระดับ 2 ทักษะหรือความคิดรวบยอด (Skill/Concept)** ระดับการใช้กระบวนการที่เกิดขึ้น ภายในสมอง ในระดับนี้มีความต้องการให้นักเรียนตัดสินใจ ด้วยตนเองว่าจะแก้ปัญหาอย่างไร นักเรียนสามาถเข้าใจในกระบวน หลักการ วิธีการ รวมถึงทุกขั้นตอนของวิธีที่กล่าวมา เพื่อนำมาตัดสินใจ วางแผน ในการแก้ปัญหา นักเรียนบอกการตัดสินใจในการเลือกใช้รูปแบบ ประบวนการ วิธีการนั้นๆ นักเรียนสามารถสรุปวิธีการ จากเหตุผลหลายๆปัจจัยที่มีส่วนหรือผลต่อคำตอบ นักเรียนสามารถสรุปคำตอบออกมาเป็นภาษา รูปแบบ มโนทัศน์ หรือรูปแบบอื่นๆที่เข้าใจ เป็นของตัวเอง

**ระดับ 3 ยุทธวิธีในการคิด(Strategic Thinking)** การสร้างข้อคาดเดา อธิบายข้อเท็จจริงในรูปความคิดรวบยอด ใช้ความคิดรวบยอดเพื่อแก้ปัญหา ระดับที่ต้องการให้นักเรียนให้เหตุผล การวางแผน การใช้หลักฐานเพื่อตัดสินใจ การแก้ปัญหามความซับซ้อนและเป็นนามธรรม กิจกรรมที่มีคำตอบ ที่เป็นไปได้มากกว่า 1 คำตอบ นักเรียน พิสูจน์ กิจกรรมลักษณะอื่นๆ ที่อยู่ในระดับ 3 คือ นักเรียนหาข้อสรุปจากการสังเกตหลักฐานและอ้างเหตุผล อย่างสมเหตุสมผล อธิบายข้อเท็จจริงในรูปความคิด รวบยอด และใช้ความคิดรวบยอดเพื่อแล้วปัญหา นักเรียนวิเคราะห์วิธีการ หลักการ อย่างเข้าใจและมีเหตุผล เพื่อใช้ประกอบการตัดสินใจในการเลือกวิธีที่เหมาะสม สำหรับแก้ปัญหา

**ระดับ 4 ขยายการคิด (Extended Thinking)** ใช้กระบวนการ วิธีการ หรือหลักการในการแก้ปัญหา ที่มากกว่า 1 วิธีในปัญหานั้นๆ วิเคราะห์จุดเด่น จุดดีของหลักการ วิธีการ เพื่อเลือกเอากระบวนการนั้นๆ มาปรับประยุกต์ให้เข้ากับวิธีการแก้ปัญหาของตัวเอง มีกระบวนการสร้างขั้นตอน วิธีการแก้ปัญหา และปรับปรุงเพิ่มเติมการแก้ปัญหาการวางแผนเพื่อการพัฒนา การออกแบบ การสร้างการเชื่อมโยงระหว่างความคิดรวบยอดและข้อเท็จจริง รวบรวมและ สังเคราะห์แนวคิดไปสู่ความคิดรวบยอดใหม่

**การตรวจสอบคุณภาพกรอบการประเมินความก้าวหน้าในการเรียนรู้ หมายถึง** กระบวนการการทดลอง (Try out) ข้อสอบและการประเมินโดยผู้เชี่ยวชาญ

**การประเมินความก้าวหน้าในการเรียนรู้ หมายถึง** การตรวจสอบการตอบข้อสอบของผู้เรียนที่เขียนตอบทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษา

**การประเมินทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษา** หมายถึง เป็นคำถามบนพื้นฐานของคำถามปลายเปิด (Open Ended Question) โดยมีคะแนนตั้งแต่ 0 – 4 โดยให้นักเรียนเขียนคำตอบลงในข้อคำถามข้อนั้นๆ ได้เลย

**5. ประโยชน์ที่คาดว่าจะได้รับ**

1. ได้แบบทดสอบใช้ในการทดสอบทักษะการรู้สารสนเทศ สำหรับนักเรียนระดับมัธยมศึกษาตนปลาย
2. มีเครื่องมือในการศึกษาแนวทาง หรือพัฒนาการจัดการบริการที่เหมาะสมในการพัฒนาทักษะการรู้

สารสนเทศเป็นรายบุคคล และกลุ่มเพื่อสามารถสร้างคุณลักษณะที่พึงประสงค์

3. สถาบันระดับมัธยมศึกษาที่สนใจสามารถนำไปใช้ในการทดสอบทักษะการรู้สารสนเทศได้

## 6. การศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง

### 6.1 ทักษะการรู้สารสนเทศ

#### 6.1.1 ความหมายและความสำคัญของสารสนเทศ

สารสนเทศมีความหมายที่หลากหลาย ขึ้นอยู่กับว่าจะนำไปใช้ในสถานการณ์ใด สารสนเทศในบางความหมายเหมาะกับบางสาขาวิชา แต่อาจจะไม่เหมาะที่จะนำมาเป็นความหมาย ในสาขาวิชาอื่น เพราะสารสนเทศสามารถมองได้หลายระดับ และหลายมุมมอง

สารสนเทศมีความแตกต่างกันตามบริบทที่ต่างกัน ซึ่งเกี่ยวข้องกับความสัมพันธ์ที่ แตกต่าง และองค์ประกอบตามบริบท ความหมาย

สมพร พุทธาพิทักษ์ผล (2546) ได้กล่าวถึงความหมายของสารสนเทศในเชิงของ ความสัมพันธ์ระหว่างข้อมูล สารสนเทศ และความรู้ไว้ว่า “สารสนเทศ เป็นการนำข้อมูลซึ่งเป็น ข้อเท็จจริงซึ่งอาจอยู่ในรูปของตัวเลข ตัวอักขระ ภาพ หรือเสียง มาประมวลอันเป็นกระบวนการใน การเพิ่มคุณค่าให้กับข้อมูลนั้นๆ ทำให้สามารถนำข้อมูลไปใช้ประโยชน์ได้ ในขณะที่ความรู้เป็น ความเข้าใจหรือความคิดในเรื่องใดเรื่องหนึ่ง ซึ่งบุคคลได้สังเคราะห์หรือทำความเข้าใจจาก สารสนเทศ”

ประภาวดี สืบสนธ (2543) ได้กล่าวว่าสารสนเทศในบริบทของการสื่อสารระหว่าง มนุษย์ “สารสนเทศ หมายถึง ข้อเท็จจริง เหตุการณ์ ที่ผ่านกระบวนการประมวลผล มีการถ่ายทอด และบันทึกไว้ในรูปแบบต่างๆ เช่น หนังสือ วารสาร หนังสือพิมพ์ รายงาน โสตทัศนวัสดุ เทป คอมพิวเตอร์ ตลอดจนถ่ายทอดในรูปแบบอื่นๆ เช่น คำพูด โดยมีวัตถุประสงค์เพื่อให้ผู้รับสารทราบ

มาลี ล้ำสกุล (2546) ได้สรุปความหมายของสารสนเทศไว้ว่า “สารสนเทศโดยสรุปมี ความหมายครอบคลุมข้อมูล ข้อเท็จจริง ทรัพยากรสารสนเทศ ความรู้ สื่อบันทึก ความคิด ประสบการณ์,

Keenan (2000) สรุปความหมายของคำว่า สารสนเทศ (Information) หมายถึง ความรู้ ข้อมูล ข้อเท็จจริงต่าง ๆ ที่ผ่านการประมวลหรือวิเคราะห์ผลสรุปด้วยวิธีการต่างๆ และสามารถ นำไปใช้ประโยชน์ได้

ดังนั้น สารสนเทศ จึงหมายถึง ข้อมูล ข้อเท็จจริง เหตุการณ์ ความรู้ ประสบการณ์ ความคิด ที่อยู่ในรูปของตัวเลข อักขระ ภาพ หรือเสียง ที่ผ่านการประมวลผล มีการถ่ายทอด และ บันทึกไว้ในรูปแบบต่างๆ โดยมีวัตถุประสงค์เพื่อให้ผู้รับสารทราบ และนำสารสนเทศไปใช้ ประโยชน์ต่อไป

#### 6.1.2 ความสำคัญของสารสนเทศ

ความสำคัญของสารสนเทศมีผู้กล่าวถึง ดังนี้

สารสนเทศมีความสำคัญ และเป็นเครื่องมือทางปัญญาที่ทำให้มนุษย์ดำรงชีวิตอยู่ได้ การที่มนุษย์จะดำรงชีวิตที่ดีในสังคมปัจจุบัน มนุษย์จำเป็นต้องปรับเปลี่ยนตนเองให้เข้ากับ สภาวการณ์โดยเจักแสวงหา และเลือกสรรสารสนเทศที่เหมาะสมทั้งต่อตนเอง และสังคม (มาลี ล้ำสกุล, 2546)

สารสนเทศเป็นสิ่งที่จำเป็นสำหรับการดำรงเผ่าพันธุ และการเอาชนะภัยอันตรายต่าง ๆ ทำให้สภาพความเป็นอยู่ของมนุษย์ดีขึ้น มนุษย์ได้เล็งเห็นความสำคัญของสารสนเทศ จึงได้คิดค้น ตัวอักษรที่ใช้ในการบันทึกเรื่องราวหรือเหตุการณ์เพื่อไว้ใช้ถ่ายทอดต่อไป (ยืน ภู่วรวรรณ, 2546) สารสนเทศเป็นสิ่งที่ล้ำค่า และเป็นทรัพยากรของชาติ ประภาวดี สืบสนธ (2543) ได้กล่าวถึงความสำคัญของสารสนเทศไว้ว่า “สารสนเทศเป็นทรัพยากรของชาติที่มีความสำคัญไม,ยิ่งหย่อนไปกว่าทรัพยากรประเภทอื่นๆ สารสนเทศมีประโยชน์นานัปการ นับตั้งแต่ช่วยลดความ อยากรู้ คลายความสงสัย จนถึงช่วยแค้ปัญหา ช่วยวางแผน และการตัดสินใจ

ได้อย่างถูกต้อง สารสนเทศจึงช่วยพัฒนาบุคคล ช่วยการปฏิบัติงาน ช่วยในการดำเนินชีวิต ซึ่งจะมีผลต่อการพัฒนา สังคม และประเทศ สารสนเทศจึงมีความสำคัญในหลายระดับ ทั้งต่อบุคคล ต่อองค์การ และ ต่อสังคม”

มาลี ล้ำสกุล (2546) ได้กล่าวถึงความสำคัญของสารสนเทศ ไว้ว่า สารสนเทศให้ อำนาจแก่ผู้ใช้ในการสร้างงานให้บรรลุ สารสนเทศจึงเปรียบได้กับทรัพย์สินที่มีค่า สารสนเทศ มีความสำคัญทั้งต่อบุคคล องค์การและต่อสังคมทั้งในด้านเศรษฐกิจ การเมือง การปกครอง รวมถึง การศึกษา ในแง,ของการศึกษา สารสนเทศเป็นเครื่องมือสำคัญของการจัดระบบการเรียนการสอน ในระดับต่างๆ ตั้งแต่ระดับต้นจนถึงระดับอุดมศึกษา รวมถึง การค้นคว้าในวิชาต่างๆ

ดังนั้น สารสนเทศจึงมีความสำคัญในการช่วยลดความอยากรู้ คลายความสงสัย ช่วยแก้ไขปัญหา รวมทั้ง ช่วยวางแผน การตัดสินใจ เพิ่มพูน และพัฒนาความรู้ไม่ว่าจะในด้านการ ทำงาน หรือการศึกษา สารสนเทศจึงมีความสำคัญต่อการพัฒนาตนเอง องค์การ สังคม และ ประเทศชาติ

**6.1.3 การรู้สารสนเทศ**

Paul Zurkowski (Eisenberg, 2004) เปนผู้เรมใช้ค่าว่า การรู้สารสนเทศ (Information Literacy) ในปี ค.ศ. 1974 ขึ้น เนื่องจากเห็นถึงพัฒนาทางเทคโนโลยีสารสนเทศ และการเพิ่มขึ้น ของสารสนเทศเป็นพลังขับเคลื่อนสำคัญ ที่เป็นป็จจัยที่ส่งผลให้เกิดการปรับเปลี่ยนในสังคมอย่างกว้างขวาง รวมถึงการเปลี่ยนแปลงรูปแบบการใช้สารสนเทศ การแสวงหาสารสนเทศ ทำให้ บุคคลต้องเผชิญกับทางเลือกสารสนเทศที่หลากหลาย และมากมาย จึงมีความจำเป็นที่ต้องเกี่ยวกับประเมิน เลือก และ สังเคราะห์ข้อมูลข่าวสาร รวมถึงการเปลี่ยนแปลง และเพิ่มทักษะใหม่ เช่น ทักษะในการแสวงหา การเข้าถึง สารสนเทศทักษะเทคโนโลยีเพื่อใช้เป็นเครื่องมือในการพัฒนา ความเความสามารถของมนุษย์ให้มีความสามารถในการคิดวิเคราะห์อย่างมีระบบ มีวิจารณญาณ และความสามารถในการ สร้างองค์ความใหม, อันนำไปสู่การเรียนเตลอดชีวิต รวมถึงการพัฒนา ประเทศสู่สังคมแห่งการเรียน! ในสังคมใหม่ที่มีการเปลี่ยนแปลงอย่างรวดเร็ว ด้วยเหตุดังกล่าว ป็จจุบันองค์กรสากล คือ UNESCO (2008, p.1-5) ได้กำหนดให้ทักษะการรู้สารสนเทศเป็นสิ่งที่ จำเป็นสำหรับมนุษย์ทั้งในระดับชาติ และนานาชาติ ความหมายของการรู้สารสนเทศ

Paul Zurkowski ได้อธิบายว่า การรู้สารสนเทศหมายถึง การที่บุคคลได้รับการแกให้ สามารถประยุกต์สารสนเทศมาใช้ได้อย่างมีประสิทธิภาพ ในด้านเทคนิคการใช้เครื่องมือในการ เข้าถึง และสามารถสังเคราะห์สารสนเทศมาใช้ได้ตามความต้องการ (Eisenberg, 2004, p. 3)

Chartered Institute of Library and Information Professionals (CILIP) (2008) ได้ให้ คำนิยามการรู้สารสนเทศไว้ว่า “การรู้สารสนเทศ คือ การทราบว่า เมื่อไหร่และด้วยสาเหตุอะไรที่ ต้องการสารสนเทศ และจะหาสารสนเทศที่ต้องการได้จากที่ไหน รวมทั้งวิธีการประเมิน และ สามารถใช้สารสนเทศในการสื่อสารได้อย่างมีความถูกต้อง และ มีจริยธรรม”

Association of College and Research Library-ACRL (1989) ได้ให้กำนยาม เรื่องการสารสนเทศไว้ว่า “การรู้สารสนเทศ หมายถึง ความสามารถในการเข้าถึงสารสนเทศ การประเมินการใช้สารสนเทศอย่างมีประสิทธิภาพ ประสิทธิผล จากแหล่งที่หลากหลาย และได้ อธิบายว่าการรู้สารสนเทศ หมายถึง ความ! และความสามารถของบุคคลในการระบุความต้องการ สารสนเทศของตนเอง ความสามารถในการค้นหา ประเมินคุณค่า และใช้ประโยชน์จาก

สารสนเทศได้อย่างมีประสิทธิภาพ ( American Library Association, 1989) การรู้สารสนเทศจึงครอบคลุมความ!ความสามารถของบุคคลในเรื่องต่อไปนี้

1. ตระหนักว่าสารสนเทศที่ถูกต้องสมบูรณ์สามารถนำไปใช้ประโยชน์ได้
2. รู้ว่าตนเองมีความต้องการสารสนเทศใด
3. สามารถตั้งคำถามหรือระบุความต้องการสารสนเทศของตนเองได้
4. สามารถระบุหรือชี้แหล่งสารสนเทศที่จะค้นหาได้
5. สามารถพัฒนากลวิธีการสืบค้นสารสนเทศได้
6. สามารถเข้าถึงแหล่งสารสนเทศทั้งที่จัดเก็บอยู่ในสื่อคอมพิวเตอร์และสื่อรูปแบบอื่นๆ ได้
7. สามารถประเมินคุณค่าสารสนเทศได้
8. สามารถจัดกลุ่มหรือหมวดหมู่สารสนเทศเพื่อนำไปใช้ประโยชน์ได้
9. สามารถบูรณาการสารสนเทศใหม่ ๆ เข้ากับองค์ความ!ที่มีอยู่เดิมได้
10. สามารถใช้สารสนเทศในการคิดเชิงวิเคราะห์ และใช้สารสนเทศในการแก้ปัญหาได้

**6.2 ทฤษฎีการทดสอบ**

การจัดการศึกษามีความจำเป็นที่จะต้องกำหนดจุดมุ่งหมายที่มีความชัดเจน เพื่อช่วยในการกำหนดคุณลักษณะที่ต้องการของผู้เรียน เพื่อใช้เป็นแนวทางในการวางแผนเพื่อใช้การจัดการเรียนรู้ จุดมุ่งหมายยังมีส่วนช่วยในการกำหนดทิศทางของการศึกษา การจัดประสบการณ์การเรียนรู้หรือกิจกรรมการจัดการเรียนรู้ที่เหมาะสม การวัดและการประเมินผลจึงเข้ามามีบทบาทสำคัญในการตัดสินผลการเรียนรู้ของผู้เรียนว่าสำเร็จตรงตามจุดมุ่งหมายที่กำหนดไว้หรือไม่ รวมทั้งยังช่วยให้สารสนเทศเพื่อช่วยในการปรับปรุงละพัฒนาการเรียนรู้ของผู้เรียนและการจัดกิจกรรมการเรียนรู้ของผู้สอน เมื่อการวัดและการประเมินผลเข้ามามีบทบาทสำคัญ ทำให้ทฤษฎีการทดสอบเข้ามามีบทบาทสำคัญในการวัดและประเมินผล เนื่องจากการการวัดและประเมินผลทางการศึกษานั้นเป็นสิ่งที่อยู่ภายในไม่สามารถวัดออกมาได้โดยตรง เพื่อทำความเข้าใจในสิ่งที่ต้องการวัดตามคุณลักษณะที่มุ่งวัด โครงสร้างของการวัด และการพัฒนาเครื่องมือ จึงจำเป็นต้องอาศัยทฤษฎีการทดสอบทดสอบเข้ามาช่วย โดยทฤษฎีการทดสอบนั้นหากแบ่งประเภทเป็น 2 ประเภทหลักใหญ่ๆ ได้แก่ ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) และทฤษฎีการทดสอบแนวใหม่ (Modern Test Theory)

ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) จะเน้นไปทางการตรวจสอบความสัมพันธ์ระหว่างคะแนนที่สังเกตได้กับคะแนนที่แท้จริง และวิเคราะห์คุณภาพโดยส่วนรวมของข้อสอบและแบบสอบ โดยใช้การวิเคราะห์คุณภาพของข้อสอบและแบบทดสอบ โดยใช้โมเดลคะแนนจริงแบบดั้งเดิม การใช้ทฤษฎีความเที่ยงและความคลาดเคลื่อน การใช้ความตรงเพื่อตรวจสอบความตรงต่างๆ หลักการการสร้างแบบทดสอบ และใช้หลักการวิเคราะห์ข้อสอบ ทฤษฎีการทดสอบแบบดั้งเดิมยังมีข้อด้อยของข้อตกลงเบื้องต้น มีข้อจำกัดของฐานความเชื่อเกี่ยวกับคะแนนความคลาดเคลื่อนและค่าพารามิเตอร์ของข้อสอบและแบบทดสอบ จึงส่งผลให้มีนักทฤษฎีการทดสอบหลายท่านได้ก่อการปฏิรูปแนวคิดเกิดขึ้น สู่ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theory) เพื่อคลายข้อตกลงเบื้องต้น และแก้ไขจุดอ่อนบางประการ

ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theory) จะมุ่งเน้นการขยายแนวคิดของทฤษฎีการทดสอบแบบดั้งเดิมให้มีขอบเขตที่กว้างขวาง และมีความน่าเชื่อถือมากยิ่งขึ้น โดยปรับข้อตกลงให้มีความสมเหตุสมผล ตามความจริงมากขึ้น เกี่ยวกับความคลาดเคลื่อนของการวัด ยอมให้มีความคลาดเคลื่อนจากการวัดนั้นมีความแตกต่างกันออกไปตามสถานการณ์การวัด เพื่อศึกษาความเที่ยงทั่วไป ภายใต้เงื่อนไขของการวัด

ตามแนวของ Generalizability Theory: G-Theory รวมทั้งศึกษาความคลาดเคลื่อนจากการวัดที่มีความแตกต่างกันไปตามลักษณะของข้อสอบและแบบสอบ ขึ้นกับระดับความสามารถของผู้สอบ และของข้อสอบและแบบสอบ ทั้งยังพยายามวัดคุณลักษณะภายในหรือความสามารถที่แท้จริงของบุคคลตามแนว Item Response Theory (IRT) นอกจากนี้ความคลาดเคลื่อนจากการวัดของบุคคลเองก็ส่งผลต่อการตอบข้อสอบมากกว่า 1 องค์ประกอบ ดังนั้นเพื่อวัดความสามารถของบุคคลจึงมี 2 องค์ประกอบขึ้นไป เพื่อให้ความสามารถหลายมิติของบุคคลสอดคล้องกับข้อมูลได้ดียิ่งขึ้นและส่งผลต่อความน่าจะเป็นของการตอบถูก ตามแนวของ Multidimensional Item Response Theory (MIRT). ทั้งทฤษฎีการทดสอบแบบดั้งเดิมทฤษฎีการทดสอบแนวใหม่ ต่างก็ต้องการวัดที่เน้นผลผลิต และพฤติกรรมที่เกิดจากการนำความรู้และการนำทักษะไปใช้ในบริบทหรือสถานการณ์ที่เป็นชีวิตประจำวันจริงและด้วยการใช้สถานการณ์ที่เลียนแบบชีวิตจริง

**พัฒนาการทฤษฎีการทดสอบ**

พัฒนาการของทฤษฎีการทดสอบนั้นมีความสัมพันธ์กันอย่างมากกับศาสตร์ทางด้านจิตวิทยา โดยเริ่มต้นจากนักจิตวิทยาเองที่ต้องการทดสอลและหาวิธีการแก้ปัญหาหารวัดต่างๆ เพื่อพัฒนาจนเป็นศาสตร์การวัดและตรวจสอบที่มีความน่าเชื่อถืออย่างที่เราพบเห็นและใช้กันในปัจจุบัน

ศาสตร์ทางด้านการวัดและประเมินผลนั้นมีมายาวนาน แต่ตอนปลายศตวรรษที่ 19 ที่นักวิทยาศาสตร์และนักสถิติได้ให้ความสำคัญ มีการพัฒนาทฤษฎีการทดสอบอย่างมากและมีความรวดเร็วและกว้างขวาง รวมทั้งมีความเป็นระบบมาก เนื่องมาจากการรวมตัวกันของนักวัดผลเพื่อแลกเปลี่ยนความคิดผลงานวิชาการกันจนทำให้เกิดเป็นแนวคิดใหม่ๆ ทฤษฎีการทดสอบอย่างมากมายจนเป็นที่ยอมรับและสนใจอย่างกว้างขวาง

**ทฤษฎีการทดสอบแบบดั้งเดิม (CTT)** ทฤษฎีการทดสอบแบบดั้งเดิมเป็นทฤษฎีที่ถือปฏิบัติใช้กันมานานแล้ว ส่วนใหญ่จะรู้จักกันในชื่อ *ทฤษฎีคะแนนจริงแบบดั้งเดิม* ซึ่งองค์ความรู้ทั่วไปจะเน้นเกี่ยวกับการตรวจสอบความสัมพันธ์ระหว่างคะแนนที่สังเกตได้กับคะแนนจริง รวมทั้งวิเคราะห์คุณภาพโดยส่วนรวมของข้อสอบและแบบสแบสำหรับผู้สอบเฉพาะกลุ่ม

ทฤษฎีคะแนนจริงแบบดั้งเดิมนี้จะตั้งอยู่บนพื้นฐานของ โมเดลการวัดและข้อตกลงเบื้องต้นที่สำคัญหลายประการ โดยทฤษฎีนี้หลักๆ จะเน้นเพื่อการพัฒนาหลักการสร้างข้อสอบ การวิเคราะห์คุณภาพของแบบสอบ และการวิเคราะห์คุณภาพของข้อสอบ โมเดลคะแนนจริงแบบดั้งเดิม (X) เกิดจากองค์ประกอบที่สังเกตไม่ได้ 2 ส่วน คือคะแนนจริง (T) และคะแนนความคลาดเคลื่อน (E) ซึ่งมีความสัมพันธ์ดังสมการ

$$X = T + E$$

ซึ่งโมเดลการวัดตั้งอยู่บนข้อตกลงเบื้องต้น (Model Assumption) จำนวน 6 ข้อ ดังนี้

**ข้อตกลงเบื้องต้นข้อที่ 1** คะแนนที่ได้จากการวัด มีความสัมพันธ์แบบเส้นตรงและเชิงบวกกับคะแนนจริงและคะแนนความคลาดเคลื่อน

**ข้อตกลงเบื้องต้นข้อที่ 2** คะแนนจริงมีสภาวะคงที่ ซึ่งเท่ากับค่าเฉลี่ยของคะแนนที่ได้จากการวัดซ้ำๆ หลายๆครั้ง

**ข้อตกลงเบื้องต้นข้อที่ 3** คะแนนความ คลาดเคลื่อนไม่มีความสัมพันธ์กับคะแนนจริง

**ข้อตกลงเบื้องต้นข้อที่ 4** ความคลาดเคลื่อนส่วนบุคคล ไม่มีความสัมพันธ์ซึ่งกันและกัน

**ข้อตกลงเบื้องต้นข้อที่ 5** การทดสอบ 2 ฉบับจะเป็นแบบทดสอบคู่ขนาน เมื่อคะแนนจริงของผู้สอบมีค่าเท่ากันทั้งสองฉบับ

**ข้อตกลงเบื้องต้นข้อที่ 6** แบบทดสอบ 2 ฉบับ จะถือว่าเป็นแบบทดสอบทัดเทียมกัน

**ความเที่ยง (Reliability)**

เมื่อใดก็ตามที่มีการนำแบบทดสอบไปใช้ แบบทดสอบจำเป็นต้องมีคุณลักษณะสำคัญอีกประการที่เรียกว่า ความเที่ยง ซึ่งความเที่ยงนี้มีอยู่หลายประเภท การประมาณค่าสัมประสิทธิ์ความเที่ยงของแบบทดสอบเองก็สามารถทำได้หลายวิธีเช่นเดียวกัน

ทฤษฎีความเที่ยงนั้นก็ได้ประยุกต์มาจากโมเดลของสเปียร์แมน โดยได้ตั้งข้อสังเกตว่า คะแนนจริงของผู้สอบกับคะแนนที่วัดได้มีความสัมพันธ์ใกล้ชิดกันเพียงใด เพื่อเป็นการบอกจึงมีค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจริงกับคะแนนที่วัดได้ หรือเรียกว่า ดัชนีความเที่ยง นั่นเอง

$$\sigma_X^2 = \sigma_T^2 + \sigma_E^2$$

การประมาณค่าตวามเที่ยงนั้นสามารถกระทำได้หลายวิธี

**1. ความเที่ยงแบบคงที่** มักจะใช้วิธีการสอบซ้ำ แล้วประมาณค่าโดยคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน

**2. ความเที่ยงแบบสมมูล** มักจะวัดเวลาเดียวกันโดยใช้แบบทดสอบแบบสมมูลแล้วประมาณค่าโดยคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน

**3. ความเที่ยงแบบคงที่และสมมูล** มักจะใช้วิธีการวัดเวลาเดียวกันแบบสอบซ้ำโดยใช้แบบ ทดสอบแบบสมมูลค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน

**4. ความเที่ยงแบบสอดคล้องภายใน** จะใช้กันอยู่หลายวิธีด้วยกัน ได้แก่ วิธีแบ่งครึ่งข้อสอบ วิธีคูเดอร์-ริชาร์ดสัน วิธีสัมประสิทธิ์แอลฟาของครอนบาค และวิธีวิเคราะห์ความแปรปรวนของฮอยท์ เป็นต้น

แต่ว่าความเที่ยงนั้นยังมีปัจจัยหลายประการที่ส่งผงต่อสัมประสิทธิ์ความเที่ยงนี้ ปัจจัยที่มีอิธผลและส่งผลสัมประสิทธิ์ความเที่ยง ได้แก่ ความเป็นเอกพันธ์ของกลุ่มผู้สอบ (Group Homogeneity) ความยาวของแบบทดสอบ (Test Length) ความสัมพันธ์ระหว่างข้อสอบ (Inter item Correlation) เวลาที้ในการทำข้อสอบ (Time Limit) และวิธีที่ใช้การประมาณค่าความเที่ยง (Methods of Estimating Reliabilities)

ปัจจัยเหล่านี้ควรนำมาพิจารณาเมื่อทำการศึกษาความเที่ยงของแบบสอบ โดยทั่วไปแล้วค่าสัมประสิทธิ์สหาสัมพันธ์จะสูงที่สุดเท่าที่จะเป็นไปได้ของแบบทดสอบนั้น

ความคลาดเคลื่อนมาตรฐานของการวัด (SEM) เป็นตัวบ่งชี้ถึงระดับการกระจายของคะแนนที่สังเกตได้จากคะแนนจริง การประมาณค่าความจริง สามารถกระทำได้โดยการประมาณเป็นช่วงด้วยการใช้ความคลาดเคลื่อนมาตรฐาน และประมาณเป็นจุดด้วยการประยุกต์การวิเคราะห์การถดถอย

**ความตรง (Validity)**

ความตรงเป็นคุณสมบัติที่สำคัญที่สุดของแบบสอบหรือของเครื่องมือวัดผล ซึ่งจะเกี่ยวข้องกับคุณภาพด้านความถูกต้องของผลที่ได้จากการวัด ทำให้เราสามารถนำคะแนนที่ได้ไปแปลความหมายถึงสิ่งที่มุ่งวัดได้

อย่างเหมาะสม การตรวจสอบความตรง เป็น *“กระบวนการรวบรวมและวิเคราะห์หลักฐานเพื่อการสนับสนุนความเหมาะสมและความถูกต้องของการนำคะแนนจากเครื่องมือวัดไปสรุปอ้างอิงถึงคุณลักษณะที่มุ่งวัด”* การตรวจสอบความตรงสามารถจำแนกประเภทตามเป้าหมายสำคัญได้ 3 ประเภทหลักๆ ซึ่งได้แก่

1) การตรวจสอบความตรงตามเนื้อเรื่อง (Content validation)

2) การตรวจสอบความตรงตามเกณฑ์สัมพันธ์ (Criterion-related validation) และ

3) การตรวจสอบความตรงเชิงทฤษฎี (Construct validation)

**การตรวจสอบความตรงตามเนื้อเรื่อง (Content validation)** วิธีการตรวจสอบมักจะให้ผู้เชี่ยวชาญตรวจสอบความเหมาะสมของนิยาม และขอบเขตของเนื้อเรื่องหรือประสบการณ์ที่มุ่งวัด โดยหากต้องการตรวจสอบความสอดคล้องของข้อสอบกับจุดมุ่งหมายของการวัด มักจะใช้ IOC (Item- Objective Congruence) หากเป็นต้องการตรวจ สอบความสอดคล้องของข้อสอบกับเนื้อเรื่องที่มุ่งวัดมักจะใช้ CVR (Content Validity Ratio)

**การตรวจสอบความตรงตามเกณฑ์สัมพันธ์ (Criterion-related validation)**

ความตรงตามเกณฑ์สัมพันธ์เป็นคุณสมบัติด้านความสอดคล้องสัมพันธ์กันระหว่างคะแนนจากแบบสอบกับเกณฑ์ภายนอก (Criterion) ที่สามารถใช้วัดลักษณะที่ต้องการนั้นได้ ความตรงตามเกณฑ์สัมพันธ์สามารถจำแนกได้เป็น 2 ประเภทตามลักษณะของเกณฑ์ที่นำมาใช้หาความสัมพันธ์ของข้อสอบ ดังนี้

**1. ความตรงตามสภาพหรือความตรงร่วมสมัย (Concurrent Validity)** เป็นความตรงตามเกณฑ์ซึ่งสัมพันธ์กับสมรรถนะการดพเนินงานในสภาพปัจจุบัน โดยวิธีการประมาณค่าความตรงตามสภาพนั้นเราสามารถหาได้โดยการคำนวณจากค่าสัมประสิทธิ์สหาสัมพันธ์ระหว่างคะแนนจากแบบทดสอบกับเกณฑ์

**2. ความตรงเชิงทำนาย (Predictive Validity)** เป็นความตรงตามเกณฑ์ที่สัมพันธ์กับสมรรถนะการดำเนินงานในอนาคต การประมาณค่าความตรงเชิงทำนายของแบบทดสอบเราสามารถหาได้จากการคำนวณค่าสัมประสิทธิ์สัมพันธ์ระหว่าง คะแนนจากแบบทดสอบกับคะแนนเกณฑ์

จะเห็นได้ว่าความตรงเชิงทำนายกับความตรงตามสภาพ มีความแตกต่างกันที่ช่วงเวลาของการเก็บข้อมูลคะแนนเกณฑ์ ถ้าแบบทดสอบสามารถให้คะแนนสอดคล้องกับสภาพปัจจุบัน ถือว่ามีความตรงตามสภาพ แต่ถ้าสามารถให้คะแนนสอดคล้องกับสภาพที่เกิดขึ้นในอนาคตก็จะเป็นความตรงเชิงทำนาย ดังนั้นการมีความตรงเชิงทำนายย่อมมีความตรงเชิงสภาพ แต่ถ้าแบบสอบมีความตรงตามสภาพแล้ว ไม่จำเป็นต้องมีความตรงเชิงทำนายเสมอไป การประมาณค่าความตรงตามเกณฑ์สัมพันธีนั้นยังมีปัจจัยหลายปัจจัยที่ส่งผลต่อและเป็นข้อจำกัด เช่น ปัญหาเกี่ยวกับการเลือกเกณฑ์ ความเที่ยงขแงคะแยยสอบและคะแนนเกณฑ์และช่วงจำกัดของคะแนนสอบและคะแนนเกณฑ์

**การตรวจสอบความตรงเชิงทฤษฎี (Construct validation)** การตรวจสอบด้วยวิธีนี้จะแตกต่างจาก 2 วิธีแรก คือ การตรวจสอบความตรงตามเนื้อเรื่องและการตรวจสอบความตรงตามเกณฑ์สัมพันธ์ มุ่งที่จะทรายถึงความถูกต้องและเหมาะสมในการนำคะแนนสอบไปใช้ตามจุดมุ่งหมาย แต่ในขณะเดียวกันความตรงทั้ง 2 ประเภทนี้ต่างมีข้อจำกัดที่สำคัญหลายประการดังกล่าวมาแล้ว

ในกรณีที่ต้องการแปลงคะแนนเพื่อสรุปอ้างอิงลักษณะทั่วไปของบุคคล ซึ่งอาจเป็นลักษณะทาง

จิตวิทยา คุณลักษณะเหล่านี้ถือว่าเป็นโครงสร้างทางความคิด เป็นภาวะสันนิษฐาน ซึ่งเป็นนามธรรม ไม่สามารถสังเกตได้โดยตรง แบบทดสอบเหล่านี้จึงมีความจำเป็นต้องมีความตรงเชิงทฤษฎี

ความตรงเชิงทฤษฎี จัดว่าเป็นคุณสมบัติที่สำคัญที่สุดของเครื่องมือวัดลักษณะที่เป็นนามธรรม ที่ต้องทำการวัดทางอ้อม จึงจำเป็นต้องใช้การพิจารณาลักษณะในบริบทของทฤษฎี โดยอาศัย แนวคิดเชิงทฤษฎี (Theoretical concepts) สำหรับการนิยามลักษณะที่มุ่งวัด เสนอโครงสร้างการวัด แลพกำหนดแนวทางการตั้งสมมติฐาน

การตรวจสอบความตรงเชิงทฤษฎี เป็นกระบวนการศึกษาขอบเขตความเหมาะสมของการนำคะแนนจากแบบทดสอบไปแปลผลถึงลักษณะทางจิตวิทยาที่มุ่งวัดได้ดีเพียงใด ดังนั้นการตรวจสอบความตรงเชิงทฤษฎียังคงไม่สามารถยืนยันและสรุปได้จากการศึกษาผลการทำนายเพียงครั้งเดียว แต่ต้องการผลการตรวจสอบในแง่ต่างๆ ตามโครงสร้างทฤษฎีของลักษณะที่มุ่งวัดนั้นจนได้หลักฐานเชิงประจักษ์ที่ชี้ชัดและเชื่อถือได้อย่างคงเส้นคงวา

วิธีการตรวจสอบความตรงเชิงทฤษฎี จำเป็นต้องอาศัยหลักฐานความสอดคล้องกับทฤษฎี สำหรับวิธีการหาหลักฐานเชิงประจักษ์เพื่อสนับสนุนความตรงเชิงทฤษฎีหลายวิธี วิธีที่นิยมใช้กันมาก ได้แก่ วิธีตัดสินโดยผู้เชี่ยวชาญ วิธีเปรียบเทียบคะแนนระหว่างกลุ่มที่ทราบ วิธีการเปรียบเทียบคะแนนจากการทดลอง วิธีการวิเคราะห์เมทริกซ์พหุลักษณะ-พหุวิธี วิธีวิเคราะห์ตัวประกอบ

**การประยุกต์ใช้ Confirmatory Factor Model: CFM สำหรับตรวจสอบความตรงตามเนื้อเรื่องและความเท่าเทียมกันของเนื้อเรื่อง**

เนื่องด้วยการตรวจสอบความตรงตามเนื้อเรื่องมักประสบปัญหาหลากหลายอย่าง ในการสร้างแบบสอบคู่ขนาน จำเป็นที่แบบสอบเหล่านั้นจะต้องมีความตรงจามเนื้อเรื่อง และความเท่าเทียมกันของเนื้อเรื่อง ซึ่งสามารถประยุกต์โมเดล CFM มาช่วยในการตรวจสอบได้ตามขั้นตอนดังนี้

1) ตรวจสอบความตรงตามโครงสร้างของเนื้อเรื่องที่มุ่งวัดของแบบทดสอบแต่ละฉบับ

2) ตรวจสอบความเท่าเทียมกันของเนื้อเรื่องระหว่างแบบสอบ

**ก่อนเกิดเป็นทฤษฎีการทดสอบแนวใหม่**

ทฤษฎีการทดสอบแบบดั้งเดิม Classical Test Theory (CTT) ถูกใช้เป็นแนวคิดพื้นฐานสำหรับการพัฒนาแบบทดสอบทางการศึกษาและจิตวิทยาเป็นส่วนใหญ่ ตลอดช่วงศตวรรษที่ 20 ที่ผ่านมา ในปัจจุบันก็ยังมีการใช้กันอย่างกว้างขวาง แต่ด้วยข้อจำกัดหลายประการของการวัดตามทฤษฎีการทดสอบแบบดั้งเดิม ทำให้มีนักทฤษฎีทางการทดสอบหลายท่านได้ก่อกระแสการวัดคุณลักษณะภายในแนวใหม่ เพื่อที่จะแก้ไขจุดอ่อนเกี่ยวกับข้อตกลงเบื้องต้นของทฤษฎีการทดสอบแบบดั้งเดิม เริ่มต้นด้วยเทอร์สโตน (L.L. Thurston, 1927, 1928) เป็นผู้บุกเบิกการวัดคุณลักษณะภายในบุคคล และพัฒนาเทคนิคการวิเคราะห์ตัวประกอบสำหรับศึกษาคุณลักษณะทางจิตวิทยา อันเป็นการวางรากฐานความคิดที่สำคัญเกี่ยวกับ ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theory: MTT) จากผลงานของ ครอนบาคและคณะ (Cronbach et.al., 1963, 1972) ได้เขียนบทความเรื่อง “The Dependability of Behavioral Measurement: Theory of Generalizability

of Score and Profiles” ได้เสนอแนวคิดเกี่ยวกับโมเดลความเที่ยงทั่วไปภายใต้เงื่อนไขแบบต่างๆของการทดสอบ รวมทั้งผลงานของ ลอร์ดและโนวิค (Load and Novice, 1968) เรื่อง “Statistical Theories of Mental Test Score” ซึ่งได้นำเสนอหลักการและแนวคิดสำคัญที่ช่วยผลักดันกระแสการปฏิรูประบบความคิดของการวัดสู่ทฤษฎีการทดสอบแนวใหม่ รวมทั้งการมีส่วนร่วมของนักทฤษฎีการทดสอบในยุคใหม่อีกหลายท่าน จนได้พัฒนาการทดสอบแนวใหม่ซึ่งสามารถจำแนกออกเป็น 2 แนวทางสำคัญๆ ได้แก่ ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของการวัด (Generalizability Theory) หรือ G-Theory และทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) หรือ IRT แต่ละแนวทางมีแนวความคิดและความเป็นมาที่สำคัญดังนี้

**ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory)**

ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) หรือ G-Study ได้มีการเสนอโมเดลของการศึกษาวิเคราะห์ความน่าเชื่อถือของผลการวัด หรือความเที่ยงในสถานการณ์หรือเงื่อนไขต่างๆ ของการวัด ทำให้เราทราบและสามารถควบคุมแหล่งความคลาดเคลื่อนได้ตรงประเด็น เพื่อผลการวัดมีความน่าเชื่อถือ หรือมีความเที่ยงสูงถึงระดับที่ต้องการ

จากแนวคิดทฤษฎีการทดสอบแบบดั้งเดิมมีความคิดพื้นฐานที่เชื่อว่า ความผันแปรของคะแนนที่สังเกตได้ประกอบด้วย 2 ส่วน คือ ความผันแปรของคะแนนจริง และความผันแปรของคะแนนความคลาดเคลื่อน ซึ่งไม่สามารถระบุหรือแบ่งแยกออกได้ นอกจากนี้ยังไม่ได้สนใจต่อเงื่อนไขการวัดและปฏิสัมพันธ์ระหว่างเงื่อนไขของการวัดที่สามรถส่งผลต่อความน่าเชื่อถือของการวัด ในขณะที่ G-Theory ได้เสนอแนวคิดในการแยกส่วนความคลาดเคลื่อน (Error) จากหลายแหล่ง (Multiple error sources) ประกอบด้วยความคลาดเคลื่อนอย่างเป็นระบบ (Systematic source) และความคลาดเคลื่อนแบบสุ่ม (Random source)

**แนวความคิดและข้อตกลงเบื้องต้น**

G-Theory เป็นทฤษฎีทางสถิติของการวิเคราะห์ความน่าเชื่อถือของผลการวัดในสถานการณ์ของการวัดผลลักษณะต่างๆ ที่เป็นเป้าหมายของการนำเครื่องมือไปใช้ ความถูกต้องความน่าเชื่อถือของการวัด คือ ความถูกต้องของการสรุปอ้างอิง (Generalization) จากคะแนนที่สังเกตได้ไปยังคะแนนจริงของบุคคล

การศึกษาความน่าเชื่อถือของผลการวัดตาม G-Theory ตั้งอยู่บนพื้นฐานของข้อตกลงเบื้องต้นดังนี้

1. คุณลักษณะที่มุ่งวัดของบุคคล เป็นค่าที่อยู่ในสภาวะคงที่ (Steady state)

2. ผู้สอบคนเดียวกันได้คะแนนต่างกันจากการวัดในแต่ละสถานการณ์ หรือเงื่อนไขของการวัดเนื่องมาจากความคลาดเคลื่อนที่เป็นระบบอย่างน้อย 1 แหล่ง

3. เมื่อพิจารณาผู้สอบทั้งกลุ่ม ความแปรปรวนที่สังเกตได้ ประกอบด้วย ความแปรปรวนของคะแนนจริง ซึ่งเป็นความแตกต่างที่แท้จริงระหว่างบุคคล

ผลการวิเคราะห์ด้วย G-Theory จะให้ค่าสัมประสิทธิ์สัมพันธ์เชิงสรุปที่แสดงถึงระดับความน่าเชื่อถือของคะแนนที่ได้จากการวัด (Level of dependability) เรียกว่า สัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของการวัด (Generalizability Coefficient) ซึ่งคล้ายกับสัมประสิทธิ์ความเที่ยง (Reliability Coefficient) ในทฤษฎีแบบดั้งเดิม

การตรวจสอบความน่าเชื่อถือของผลการวัดโดยใช้ G-Theory ทำให้ผู้บริหารการทดสอบสามารถทำการตัดสินความน่าเชื่อถือของผลการวัดได้ 2 ลักษณะ ได้แก่

1. ความน่าเชื่อถือของคะแนนในการนำผลไปใช้เชิงสัมพัทธ์ (Relative)

2. ความน่าเชื่อถือของคะแนนในการนำผลไปใช้สัมบูรณ์(Absolute)

**สัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient)**

สำหรับทฤษฎี G-Theory สัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient or $\rho_i^2$) เป็นสัดส่วนระหว่างความแปรปรวนของคะแนนเอกภพกับความแปรปรวนของค่าคาดหมายของคะแนนที่สังเกตได้ สัมประสิทธิ์การสรุปอ้างอิงมี 2 ประเภท คือ

1. สัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ ($\rho_{Abs}^2$)

2. สัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ ($\rho_{Rel}^2$)

**การประมาณค่าสัมประสิทธิ์สรุปอ้างอิงสำหรับการออกแบบการวัดที่มีฟาเซทเดียว (G-Coefficient for Single-Facet Design)**

ในการศึกษาคุณภาพของแบบทดสอบ นักวัดผลต้องการทราบว่าสถานการณ์ของการใช้แบบสอบ หรือเงื่อนไขทดสอบ 1 ปัจจัย เช่น ความยาวข้อสอบ (I) หรือ จำนวนครั้ง (O) หรือจำนวนผู้ตรวจ (R) จะมีผลต่อความน่าเชื่อถือของแบบสอบเพียงใด

POPULATION: Examinees or Person

UNIVERSE: All possible combinations of conditions of the facet.

FACET: Item or Occations

DESIGN: p x i design

**การประมาณค่าสัมประสิทธิ์สรุปอ้างอิงสำหรับการออกแบบการวัดที่มีสองฟาเซท (G-Coefficient for Two Facet Design)**

**1.) การออกแบบ TWO FACET CROSSED DESIGN**

ในการศึกษาคุณภาพของแบบทดสอบ นักวัดผลต้องการทราบว่าสถานการณ์ของการใช้แบบสอบ หรือเงื่อนไขทดสอบ 2 ปัจจัย ได้แก่ ความยาวข้อสอบ (I) และ จำนวนครั้ง (O) จะมีผลต่อความน่าเชื่อถือของแบบสอบเพียงใด

POPULATION: Examinees or Person

UNIVERSE: All possible combinations of conditions of the facet.

FACET 1: Item (I)

FACET 2: Occations (O)

DESIGN: p x i x o design

**2.) การออกแบบ TWO FACET NESTED DESIGN**

ในการศึกษาคุณภาพของแบบทดสอบ นักวัดผลต้องการทราบว่าสถานการณ์ของการใช้แบบสอบ หรือ

เงื่อนไขทดสอบ 2 ปัจจัย ได้แก่ ความยาวข้อสอบ (I) และ จำนวนครั้ง (O) ของการทดสอบ (เมื่อการสอบแต่ละครั้งใช้ข้อสอบต่างกัน) จะมีผลต่อความน่าเชื่อถือของแบบสอบเพียงใด

POPULATION: Examinees or Person

UNIVERSE: FACET 1: Item (I)

FACET 2: Occations (O) และ i : o

DESIGN: p x (i : o) design

**การประมาณค่าสัมประสิทธิ์สรุปอ้างอิงสำหรับการออกแบบการวัดที่มีสามฟาเซท (G-Coefficient for Three Facet Design)**

ในการศึกษาคุณภาพของแบบทดสอบ นักวัดผลต้องการทราบว่าสถานการณ์ของการใช้แบบสอบ หรือเงื่อนไขทดสอบ 1 ปัจจัย เช่น ผู้ประเมิน (R) จำนวนครั้ง (O) และแหล่งของผู้ประเมิน (R) จะมีผลต่อความน่าเชื่อถือของแบบสอบเพียงใด

**ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT)**

ทฤษฎีการตอบสนองข้อสอบ หรือ IRT สามารถจำแนกออกได้เป็น 2 ลักษณะ ได้แก่ ทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า (Dichotomous IRT) และทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous IRT)

**แนวคิดพื้นฐานของ IRT**

โมเดลการตอบสนองข้อสอบ (IRT Models) เป็นระบบความสัมพันธ์ระหว่างตัวแปรอิสระที่รวมกันสำหรับทำนายตัวแปรตาม สำหรับโมเดลการตอบสนองข้อสอบ ตัวแปรอิสระจะประกอบด้วย ตัวแปรแฝง คือ ความสามารถที่แท้จริงของผู้สอบ ($\theta$) และคุณลักษณะของผู้สอบ (B) sหรือค่าพารามิเตอร์ข้อสอบ (a, b, c) ส่วนตัวแปรอิสระเป็นตัวแปรที่สังเกตได้ คือ โอกาสการตอบข้อสอบได้ถูกต้อง

โมเดลการตอบสนองข้อสอบจะประกอบด้วยค่าพารามิเตอร์และค่าคงที่ดังนี้

1. พารามิเตอร์ของผู้สอบ

$\theta$ = ระดับความสามารถผู้สอบ โดยที่ค่า $\theta$ มีพิสัยอยู่ระหว่าง $-\infty$ ถึง $+\infty$

2. พารามิเตอร์ของข้อสอบ

$a_i$ = ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i (Discrimination parameter) ในทางทฤษฎีแล้วมีค่าระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัตินิยมใช้ $a_i$ ระหว่าง +0.50 ถึง +2.50

$b_i$ = ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i (Difficulty parameter) ในทางทฤษฎีแล้วมีค่าระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัตินิยมใช้ $b_i$ ระหว่าง -2.50 ถึง +2.50

$c_i$ = ค่าพารามิเตอร์โอกาสในการเดาข้อสอบถูก (Guessing parameter) ในทางทฤษฎีแล้วมีค่าระหว่าง 0 ถึง 1 แต่ในทางปฏิบัตินิยมใช้ $c_i$ ไม่เกิน 0.30

3. ค่าคงที่

e = 2.71828, D = 1.70

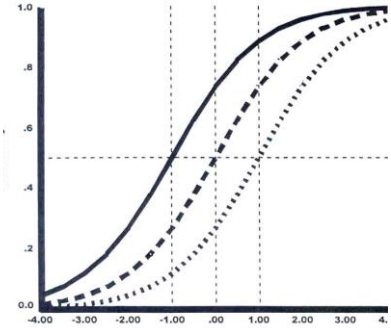
**ข้อตกลงของทฤษฎีการตอบข้อสอบ**

1. ความเป็นมิติเดียว (Unidimensionality: One trait)
2. ความเป็นอิสระ (Independence: Local Independent)
3. โค้งคุณลักษณะข้อสอบ (Item Characteristic Curve: Item Response Model)
4. ข้อสอบที่ใช้ต้องไม่เป็นข้อสอบประเภทความเร็ว

**โมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ (One parameter model or Rasch Model)**

โมเดลแรกที่มีความสำคัญที่สุดที่นำมาใช้ในโมเดลทฤษฎีการตอบข้อสอบ โค้งลักษณะข้อสอบสำหรับโมเดลโลจิสติกแบบ 1 พารามิเตอร์ แสดงในรูปฟังก์ชันทางคณิตศาสตร์ ดังนี้

$$P_i = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \text{ ; I =1,2,...,n}$$



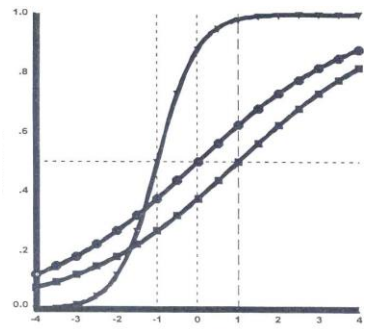

**รูปที่ 1** ลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์

โค้งลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ มีลักษณะที่สำคัญ คือ $b_i$ มีค่าแปรเปลี่ยนตามลักษณะข้อสอบแต่ละข้อ $a_i$ มีค่าคงที่ และ $c_i$ เท่ากับ 0

**โมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ (Two parameter model)**

โมเดลแรกที่มีความสำคัญที่สุดที่นำมาใช้ในโมเดลทฤษฎีการตอบข้อสอบ โค้งลักษณะข้อสอบสำหรับโมเดลโลจิสติกแบบ 2 พารามิเตอร์ แสดงในรูปฟังก์ชันทางคณิตศาสตร์ ดังนี้

$$P_i = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \text{ ; I =1,2,...,n}$$



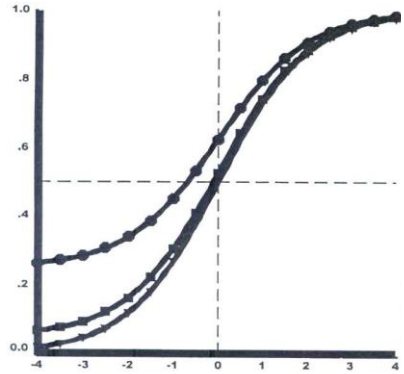

**รูปที่ 2** ลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์

โค้งลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ มีลักษณะที่สำคัญ คือ $b_i$ เป็นค่าที่แสดงตำแหน่งของ ICC ณ จุด $\theta$ ที่มีโอกาสตอบข้อสอบถูก 0.50 ส่วน $a_i$ เป็นค่าความชันของ ICC ณ จุด $b_i$ ส่วน $c_i$ เท่ากับ 0 และ e = 2.718, D = 1.7

**โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (Three parameter model)**

โมเดลแรกที่มีความสำคัญที่สุดที่นำมาใช้ในโมเดลทฤษฎีการตอบข้อสอบ โค้งลักษณะข้อสอบสำหรับโมเดลโลจิสติกแบบ 3 พารามิเตอร์ แสดงในรูปฟังก์ชันทางคณิตศาสตร์ ดังนี้

$$P_i = \boldsymbol{C_i} + (1-c_i)\frac{e^{Da_i(\theta-b_i)}}{1+e^{Da_i(\theta-b_i)}} \; ; \; I = 1,2,\ldots,n$$

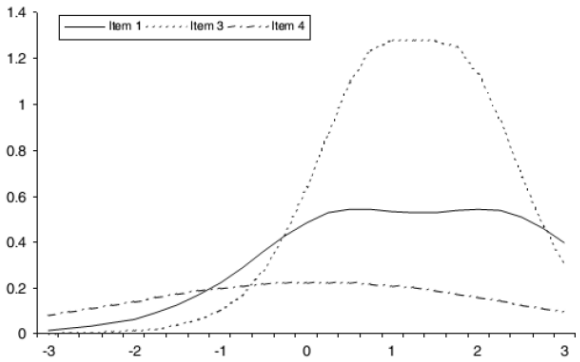



**รูปที่ 3** ลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์

โค้งลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ มีลักษณะที่สำคัญ คือ $b_i$ เป็นค่าที่แสดงตำแหน่งของ ICC ณ จุด $\theta$ ที่มีโอกาสตอบข้อสอบถูก $\frac{1+c_i}{2}$ ส่วน $a_i$ เป็นค่าความชันของ ICC ณ จุด $b_i$ ส่วน Ci คือค่าพารามิเตอร์โอกาสการเดาข้อสอบได้ถูก e = 2.718, D = 1.7

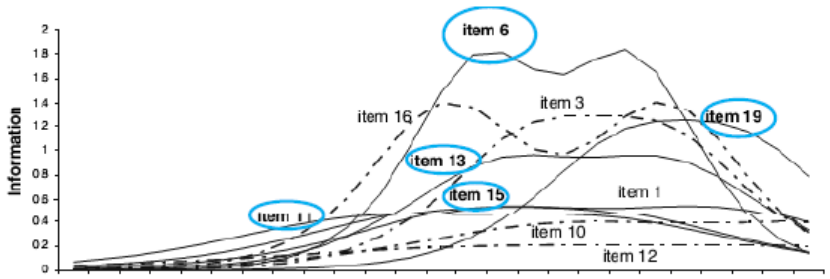
**ฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ (Item and Test Information)**

**ฟังก์ชันสารสนเทศของข้อสอบ (Item Information)** เป็นดัชนีผสมที่สร้างจากัชนีคุณลักษณะของข้อสอบหลายลักษณะ ประกอบด้วย ค่าพารามิเตอร์อำนาจจำแนก ค่าพารามิเตอร์ความยาก ค่าความแปรปรวนคะแนนรายข้อ เพื่อใช้บ่งชี้คุณภาพของข้อสอบ




**รูปที่ 4** ลักษณะกราฟฟังก์ชันสารสนเทศของข้อสอบ (Item Information)

**ฟังก์ชันสารสนเทศของแบบสอบ (Test Information)** เป็นค่าฟังก์ชันสารสนเทศของแบบสอบ อันเกิดจากผลรวมเชิงพีชคณิตของฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อเข้าด้วยกันทั้งฉบับ ณ ตำแหน่ง $\theta$ เดียวกัน




**รูปที่ 5** ลักษณะกราฟฟังก์ชันสารสนเทศของแบบสอบ (Test Information)

**ความคลาดเคลื่อนมาตรฐานของการประมาณค่า (Standard Error of Estimation)**

ความคลาดเคลื่อนมาตรฐานของการประมาณค่า [SE($\theta$)] เป็นส่วนเบี่ยงเบนมาตรฐานของการแจกแจงความน่าจะเป็น ของค่าประมาณความสามารถที่แท้จริง ($\theta$) ซึ่งเป็นค่าสัดส่วนผกผันกับความถูกต้องแม่นยำของการประมาณค่าความสามารถ หรือค่าฟังก์ชันสารสนเทศของแบบสอบ

**โมเดลการตอบสนองข้อสอบแบบหลายมิติ (Multidimensional Item Response Theory)**

แนวคิดของการพัฒนาโมเดลการตอบสนองข้อสอบแบบหลายมิติ หรือ MIRT เป็นโมเดลที่มีคุณลักษณะแฝงมากกว่า 1 องค์ประกอบหรือหลายมิติ Multidimensional Item Response Theory ได้ขยายให้รองรับ $\theta$ หลายองค์ประกอบ

โมเดลการตอบสนองข้อสอบแบบหลายมิติถือว่าคุณลักษณะแฝงของบุคคลมากกว่า 1 องค์ประกอบส่งผลต่อการตอบข้อสอบ ดังนั้นพารามิเตอร์ของผู้สอบจึงมีตั้งแต่ 2 พารามิเตอร์ขึ้นไป การพิจารณาถึงความสามารถหลายมิติของบุคคล น่าจะช่วยทำให้โมเดลมีความสอดคล้องกับข้อมูลได้ดียิ่งขึ้น

จากการศึกษาเอกสารงานวิจัยที่ได้กล่าวไว้ข้างต้นแล้ว การวัดทางการศึกษาและจิตวิทยาเป็นการวัดคุณลักษณะภายในของมนุษย์ ซึ่งไม่สามารถสังเกตได้โดยตรง แต่มีความสำคัญและจำเป็นต้องศึกษา เพราะการวัดคุณลักษณะภายในจะทำให้เข้าใจการเกิดพฤติกรรมภายนอกของมนุษย์ที่สามารถสังเกตได้โดยตรง อันจะนำไปสู่การทำนายควบคุม และพัฒนาพฤติกรรมมนุษย์ การวัดคุณลักษณะภายในของมนุษย์จำเป็นต้องอาศัย ทฤษฎีการทดสอบ เพื่อทำความเข้าใจคุณลักษณะของสิ่งที่มุ่งวัด โครงสร้างของการวัด และการพัฒนาเครื่องมือสำหรับทดสอบ

ทฤษฎีการทดสอบเป็นองค์ความรู้ที่มีนัยทั่วไปเกี่ยวกับการทดสอบ วิธีการแก้ปัญหา การทดสอบและพัฒนาเครื่องมือการทดสอบ ซึ่งจะช่วยให้นักวัดผลสามารถทำการสร้างและพัฒนาแบบสอบให้มีคุณภาพ สามารถแปลความหมายผลการวัดได้อย่างถูกต้องและสามารถนำสารสนเทศไปใช้สำหรับ การตัดสินใจได้อย่างเหมาะสม ประเภททฤษฎีการทดสอบสามารถแบ่งได้เป็น 2 ประเภทหลักๆ ได้แก่ ทฤษฎีการทดสอบแบบ

ดั้งเดิม ซึ่งจะเน้นการตรวจสอบความสัมพันธ์ระหว่างคะแนนที่สังเกตได้กับคะแนนที่แท้จริง โดยการวิเคราะห์คุณภาพรวมของข้อสอบและแบบสอบ ส่วนทฤษฎีการทดสอบแนวใหม่ มุ่งเน้นการขยายแนวคิดของทฤษฎีการทดสอบแบบดั้งเดิม เพื่อให้ได้แนวทางการวัดที่น่าเชื่อถือยิ่งขึ้นซึ่งมีพัฒนาการที่สำคัญ 2 แนวทาง ได้แก่ ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) และทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT)

ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theory) จะมุ่งเน้นการขยายแนวคิดของทฤษฎีการทดสอบแบบดั้งเดิมให้มีขอบเขตที่กว้างขวาง และมีความน่าเชื่อถือมากยิ่งขึ้น โดยปรับข้อตกลงให้มีความสมเหตุสมผล ตามความจริงมากขึ้น เกี่ยวกับความคลาดเคลื่อนของการวัด ยอมให้มีความคลาดเคลื่อนจากการวัดนั้นมีความแตกต่างกันออกไปตามสถานการณ์การวัด เพื่อศึกษาความเที่ยงทั่วไป ภายใต้เงื่อนไขของการวัดตามแนวของ Generalizability Theory: G-Theory รวมทั้งศึกษาความคลาดเคลื่อนจากการวัดที่มีความแตกต่างกันไปตามลักษณะของข้อสอบและแบบสอบ ขึ้นกับระดับความสามารถของผู้สอบ และของข้อสอบและแบบสอบ ทั้งยังพยายามวัดคุณลักษณะภายในหรือความสามารถที่แท้จริงของบุคคลตามแนว Item Response Theory (IRT)

ทฤษฎีการทดสอบแนวใหม่จึงน่าจะให้ผลการวัดที่ชัดเจนตรงประเด็นมากขึ้น อันจะเป็นประโยชน์ต่อการพัฒนาข้อสอบและแบบสอบให้สามารถนำไปใช้ในสถานการณ์ต่างๆของการทดสอบได้อย่างมีประสิทธิภาพ สามารถระบุนัยทั่วไปของคะแนนจริง คุณของแบบสอบตามเงื่อนไขของการทดสอบ รวมทั้งการประเมินความสามารถที่แท้จริงของผู้สอบ และการบรรยายพฤติกรรมการตอบสนองข้อสอบของผู้สอบได้เป็นอย่างดี

### 6.3 โมเดลเพื่อประเมินความก้าวหน้าในการเรียนรู้: โมเดลเชิงโครงสร้าง (Construct Modelling)

พัชรี จันทร์เพ็ง (2559ก, 2559ข, 2559ค, 2559ง) ได้ทำการสังเคราะห์แนวคิดเกี่ยวกับโมเดลเพื่อประเมินความก้าวหน้าในการเรียนรู้ โดยนำแนวคิดโมเดลเชิงโครงสร้างในการพัฒนาการประเมิน โดยมีรายละเอียดที่สำคัญ ดังนี้

#### 6.3.1 ความหมายและแนวคิดพื้นฐานของการประเมินความก้าวหน้าในการเรียนรู้ (Learning Progression)

การวัดและประเมินผลการเรียนรู้แนวใหม่เพื่อให้พร้อมรับสู่ศตวรรษใหม่ ไม่ได้มีจุดมุ่งเน้นเพียงเพื่อวัดให้ตรงกับสิ่งที่ต้องการวัดตามมาตรฐานหรือตัวชี้วัดตามหลักสูตรแกนกลาง พ.ศ. 2551 เท่านั้น แต่จำเป็นอย่างยิ่งที่ต้องมีการประเมินควบคู่ไปกับการจัดการเรียนรู้ในชั้นเรียนของครู เพื่อนำผลที่ได้ไปสู่การปรับปรุงพัฒนาผู้เรียนได้อย่างทันท่วงที และนำไปสู่การปรับปรุงหลักสูตรให้มีคุณภาพในอนาคต อันเป็นผลมาจากการสามารถตีความผู้เรียนได้อย่างเป็นรูปธรรมและมีความหมาย เพื่อทำให้เกิดความตรงในเชิงตีความ (Interpretation Validity) ตามมาตรฐานการประเมินทางด้านการศึกษาและจิตวิทยา (American Educational Research Association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME), 2014) ที่เป็นจุดเน้นที่สำคัญจากมาตรฐานการประเมินในอดีตที่มา

ความก้าวหน้าในการเรียนรู้ (Learning Progression; LP) หมายถึง ลักษณะของการแสดงแนวโน้มของสิ่งที่เกิดขึ้นกับผู้เรียนซึ่งเป็นผลมาจากสิ่งที่ผู้เรียนได้เรียนรู้และภายใต้มาตรฐานที่ได้มีการกำหนดไว้ หรือที่ได้มีการพัฒนาไว้ ซึ่งมีตั้งแต่ระดับความสามารถต่ำสุดไปยังระดับความสามารถที่สูงขึ้น หรือสะท้อนให้เห็นถึงสิ่งที่มีความเข้าใจซับซ้อนมากยิ่งขึ้น (Alonzo & Gearhart, 2006; Heritage, 2007; Heritage & Anderson,

2009 cited in McMillan, 2011: 150) จุดมุ่งหมายหลักคือทำอย่างไรผู้เรียนจึงจะเกิดการพัฒนาการเรียนรู้จากระดับหนึ่งไปยังอีกระดับที่สูงขึ้น ซึ่งการที่จะสามารถนำพาผู้เรียนให้ไปสู่ในระดับที่ให้เกิดการพัฒนานั้นจำเป็นต้องอาศัยข้อมูลที่ไม่เป็นทางการที่ได้จากประเมินในระหว่างที่มีการเรียนการสอน (Formative Assessment) ซึ่งข้อมูลที่ได้จะเป็นองค์ประกอบสำคัญในการพัฒนาการจัดการเรียนการสอนในอนาคต การศึกษาความก้าวหน้าของผู้เรียนจะเป็นการเพิ่มศักยภาพในการให้ข้อมูลย้อนกลับ และการปรับปรุงการเรียนการสอนให้ไปอยู่ในทิศทางที่สูงขึ้น (Heritage, 2008 cited in McMillan, 2011: 150)

การประเมินความก้าวหน้าในการเรียนรู้สามารถสรุปเป็น 2 หลักการที่สำคัญ (Popham, 2007; Wilson & Bertenthal, 2005 cited in McMillan, 2011: 150) คือ

(1) เป็นคำอธิบายที่สะท้อนถึงความสำเร็จเกี่ยวกับความซับซ้อนของกระบวนการคิดผู้เรียนซึ่งสะท้อนถึงสิ่งที่ผู้เรียนได้เกิดการเรียนรู้ และจะมีแนวทางนำไปสู่การปรับปรุงที่สูงขึ้นได้อย่างไรบ้าง

(2) เป็นสิ่งที่สะท้อนให้เห็นถึงแผนที่นำทางของผู้เรียนว่ามีความห่างจากมาตรฐานที่หลักสูตรกำหนดไว้มากน้อยเพียงใด ซึ่งในแต่ระดับจะประกอบด้วยความรู้และทักษะย่อยๆ ที่ต้องการให้เกิดกับผู้เรียน

การประเมินความก้าวหน้าในการเรียนรู้ถือเป็นแผนที่นำทางที่สำคัญ (Road Map) เพื่อให้ข้อมูลสารสนเทศที่สำคัญในการเก็บรวบรวมข้อมูลสำคัญสำหรับผู้เรียนว่าผู้เรียนมีความรู้ความเข้าใจเกี่ยวกับอะไรและมีสิ่งใดที่จำเป็นต้องรู้เกี่ยวกับผู้เรียน โดยเป็นองค์ประกอบที่เชื่อมโยงกันระหว่างความก้าวหน้าจากการจัดกิจกรรมการเรียนรู้และเกณฑ์ที่ใช้ในการประเมินระดับความสามารถ ซึ่งผู้เรียนแต่ละคนจะมีระดับความสามารถและสิ่งที่มีความจำเป็นต้องได้รับการพัฒนาที่แตกต่างกัน ซึ่งถ้าหากว่าผู้สอนสามารถนำสิ่งที่ได้เรียนรู้ในครั้งนี้นำไปสู่การจัดการเรียนการสอนให้เหมาะสมในชั้นเรียนสำหรับผู้เรียนแต่ละคน

เป้าหมายสำคัญคือการวัดการคิดและการให้เหตุผลของผู้เรียน การประเมินความก้าวหน้าในการเรียนรู้ถือเป็นวิธีการแนวใหม่ที่นำมาใช้เป็นแนวทางในการประเมินระหว่างเรียนที่ต้องทำควบคู่ไปกับการจัดการเรียนการสอน สิ่งสำคัญ คือ การประเมินความก้าวหน้าในการเรียนรู้ไม่ได้ให้รายละเอียดเพียงแค่ผู้เรียนมีความก้าวหน้าในการเรียนรู้ตามมาตรฐานที่กำหนดไว้เท่านั้นแต่ยังสามารถที่จะนำมาใช้เป็นแนวทางสำคัญในการพัฒนาหลักสูตร และการสอนได้ในคราวเดียวกัน (McMillan, 2011: 150) เช่น การประเมินการประเมินความก้าวหน้าในการเรียนรู้ รายวิชาสังคมศึกษา มีการศึกษาระดับของความก้าวหน้าในการเรียนรู้ของผู้เรียนโดยจะเพิ่มขึ้นตามความซับซ้อนของระดับความเข้าใจของผู้เรียน จากระดับหนึ่งสูงขึ้นไปยังอีกระดับหนึ่ง โดยการประเมินความก้าวหน้าในการเรียนรู้จะมีการเชื่อมให้เห็นความสำคัญกับการเรียนการสอนในชั้นเรียน โดยคำสำคัญที่ใช้จะเพิ่มจาก “การจำ (Recognizing)” ไปยัง “ความเข้าใจ (Understanding)” และไปยัง “ขั้นที่มีการคิดซับซ้อนมากขึ้น Increasing Depth” ซึ่งคำอธิบายในแต่ละระดับจะสะท้อนให้เห็นถึงแนวทางในการให้คะแนน การให้ข้อมูลย้อนกลับ และแนวทางการจัดการเรียนการสอนในแต่ละระดับ

### 6.3.2 โมเดลเชิงโครงสร้าง (Construct Modelling): หลักการพื้นฐาน 4 ประการ (4 Building Blocks)

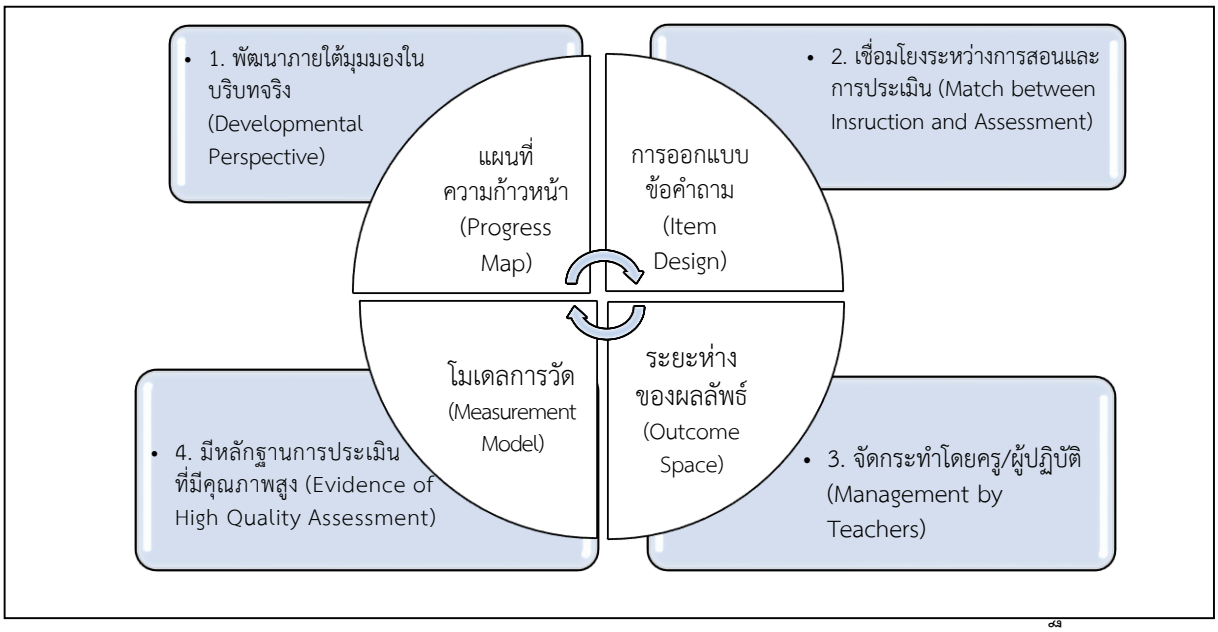
ระบบการประเมินของศูนย์การพัฒนาระบบการประเมินแห่งวิทยาลัย UC Berkeley ประเทศสหรัฐอเมริกา ซึ่งเป็นที่รู้จักกันในนาม โดย Professor Mark Wilson เป็นผู้ก่อตั้ง โดยเป็นที่รู้จักของบุคคลทั่วไปในนาม ระบบการประเมินแบบ BEAR (BEAR Assessment System) ได้มีการนำวิธีการประเมินมาเป็นส่วนที่สำคัญแทรกเข้าไปเป็นส่วนหนึ่งของกระบวนการวิจัย ซึ่งได้มีการพัฒนาอย่างต่อเนื่อง (e.g. Wilson, 1995; Wilson & Sloane, 2000; Wilson, 2009; Kennedy, & Wilson, 2006; Black, Wilson, & Yao, 2011; Wilson et al., 2012; Wilson, 2015) ได้พัฒนาหลักการที่สำคัญเกี่ยวกับบทบาทของการประเมินที่เกี่ยวข้องกับบริบทของการเรียนการสอนในชั้นเรียน (The Role of Assessment in Pedagogy) มี 4 หลักการพื้นฐานที่สำคัญ (4 Principles) ประกอบด้วย (1) มุมมองของการพัฒนา (A Developmental Perspective) ที่ต้องสอดคล้องกับการเรียนรู้ ของผู้เรียน (2) การเชื่อมโยงกันระหว่างการสอนและการประเมิน (A Match between Instructional and Assessment) (3) การบริหารจัดการโดยครูผู้สอน (Management by Teacher) และ (4) การมีหลักฐานที่มีคุณภาพ (Generating Quality Evidence) ซึ่งใช้เป็นฐานของการวิจัยมาใช้ตั้งแต่ขั้นตอนของแรกจนขั้นตอนสุดท้ายของกระบวนการวิจัย

จากหลักการพื้นฐาน 4 ประการดังกล่าวข้างต้น มีองค์ประกอบที่สำคัญในแต่ละหลักการพื้นฐาน ที่เรียกว่า 4 องค์ประกอบการประเมิน (4 Building Blocks) เป็นส่วนที่สำคัญ ประกอบด้วย (1) แผนที่เชิงโครงสร้าง/คคความก้าวหน้า (Construct/Progress Map) (2) ออกแบบข้อคำถาม/ชิ้นงาน (Item/Task Design) (3) ระยะห่างของผลลัพธ์ (Outcome Space) และ (4) โมเดลการวัด/แผนภูมิของ Wright (Measurement Model/Wright Map) เมื่อมีการบูรณาการร่วมกันระหว่างหลักการพื้นฐานและองค์ประกอบของการประเมิน จะทำให้เกิดกระบวนการพัฒนาระบบการประเมินที่เน้น การตีความอย่างมีความหมายกับการทำงานของผู้เรียนที่สอดคล้องกับการรู้คิด (Cognitive) และเป้าหมาย ของหลักสูตร (Goal of a Curriculum) ซึ่งทำให้เป็นระบบที่สอดคล้องกับหลักสูตร กิจกรรมการเรียนการ สอนและการประเมินจะฝังตัวไปด้วยกัน สามารถสรุปได้ดังแผนภาพที่ 1

| | |
|---|---|
| **หลักการที่ 1 :** | การพัฒนาแผนที่ความก้าวหน้าอยู่ภายใต้มุมมองในบริบทจริง (A Developmental Perspective) |
| **องค์ประกอบที่ 1:** | การสร้างแผนที่เชิงโครงสร้าง/ความก้าวหน้า (Construct/Progress Map) |
| **หลักการที่ 2 :** | เชื่อมโยงการสอนและการประเมิน (Match between Instruction and Assessment) |
| **องค์ประกอบที่ 2:** | การพัฒนาข้อคำถาม/ภาระงาน (Design of Items/Task) |
| **หลักการที่ 3 :** | การบริหารจัดการและมีการใช้ข้อมูลที่ได้จากการประเมินอย่างมีความหมายโดยครู หรือผู้ปฏิบัติการในชั้นเรียน (Management by Teacher) |
| **องค์ประกอบที่ 3 :** | การระบุผลการตอบของผู้เรียนออกมาเป็นระดับความก้าวหน้าเป็นเกณฑ์ การให้คะแนนจากการ Think-alounds ของผู้เรียน (Outcome Space/Progress Guides) |
| **หลักการที่ 4 :** | การประเมินในชั้นเรียนจะต้องมีมาตรฐานทั้งความตรง ความเชื่อถือได้และความยุติธรรม (High Quality Evidence) |
| **องค์ประกอบที่ 4:** | โมเดลการวัด/แผนภาพการตอบสนองของผู้เรียน (Measurement Model/Wright Map) |

**แผนภาพที่ 1** การบูรณาการร่วมกันระหว่าง 4 หลักการพื้นฐานและ 4 องค์ประกอบของการประเมิน เพื่อนำไปสู่การตีความหมายของผู้เรียนได้มีคุณภาพและน่าเชื่อถือ

จากแผนภาพที่ 1 เมื่อนำ 4 องค์ประกอบการประเมินดังกล่าวข้างต้นฝังตัวเข้าไปใน 4 หลักการพื้นฐาน เพื่อให้กระบวนการดังกล่าวมีเป้าหมายหลักเพื่อให้การสะท้อนผลมีประสิทธิภาพทั้งจากครูไปยังผู้เรียน และขณะเดียวกันจากนักเรียนไปที่ครูในระยะเวลาที่รวดเร็ว เพื่อมีการวางแผนการปรับปรุงแก้ไขทั้งระยะ สั้นและระยะ ถ้านำมาเชื่อมโยงกับแนวคิดของ Wilson (2005) เรียกว่า โมเดลเชิงโครงสร้าง (Construct Modelling) ดังปรากฏในแผนภาพที่ 2 เป็นการพัฒนาระบบการประเมินที่เน้นการตีความอย่างมีความหมายกับการทำงานของผู้เรียนที่สอดคล้องกับการคิด (Cognitive) และเป้าหมาย ของหลักสูตร ซึ่งทำให้เป็นระบบที่สอดคล้องกับหลักสูตร กิจกรรมการเรียนการ สอนและการประเมินจะฝังตัวไปด้วยกัน กับกรอบการสร้างแผนที่ความคิด ใน 4 กรอบความคิด (Wilson, 2005; Wilson, 2009; Kennedy, & Wilson, 2006; Black, Wilson, & Yao, 2011; Wilson et al., 2012) เพื่อให้ผลการประเมินที่ได้มีความตรง (Validity) ความน่าเชื่อถือ (Reliability) และความยุติธรรม (Fairness) ตามมาตรฐานการประเมินทางด้านการศึกษาและจิตวิทยา (AERA, APA, & NCME, 2014) ซึ่งได้อธิบายถึงหลักฐานสำคัญที่จะทำให้เกิดความตรงนั้นจำเป็นอย่างยิ่งที่ จะต้องมีความตรงเชื่อมโยงกัน (Coherent Validity) เพื่อนำไปสู่ความตรงในการแปลความหมายของผล การประเมินได้อย่างแท้จริง (Interpretative Validity) และสอดคล้องกับตรงตามความต้องการในการ ประเมินได้อย่างแท้จริง (Consequential Validity) ซึ่งการประเมินที่ได้เชื่อมโยงกับการสอนสุดท้ายก็จะ ฝังตัวเข้าไปในหลักสูตรในที่สุดซึ่งหาพบว่าผลการประเมินที่ได้ไม่สามารถตอบโจทย์ได้ก็สามารถที่จะปรับ เข้าสู่วงจรใหม่ (Revision Construct Map) เพื่อให้เกิดความตรงมากยิ่งขึ้นและนำไปสู่การนำไปใช้เพื่อ ปรับปรุงผู้เรียนอย่างแท้จริง




**แผนภาพที่** 2 ประเมินที่ควบคู่กับการเรียนการสอน (Wilson, 2005; Wilson & Slone, 2000) ดัดแปลงจาก Wilson, M., Beja, I., Scalise, K.,Templin, J., Wiliam, D., & Irribarra, D. T. (2012). Perspectives on Methodological Issues. In Griffin, P., McGaw, B., & Care, E. (Eds.). **Assessment and Teaching of 21st Century Skills springer New York 2012**, Springer: 129

**3.3 การประเมินความก้าวหน้าในการเรียนรู้ (Learning Progression) กับบริบทของการศึกษาในประเทศไทย**

การประเมินความก้าวหน้าในการเรียนรู้ถือเป็นกลไกที่สำคัญและจำเป็นสำหรับระบบการศึกษาที่เน้นมาตรฐานการศึกษา (Standard-based Education) โดยเฉพาะอย่างยิ่งในประเทศไทยที่ยึดมาตรฐานการเรียนรู้ตามหลักสูตรแกนกลางปี 2551 ใน 8 กลุ่มสาระการเรียนรู้ ซึ่งในแต่สาระการเรียนรู้จะมีจำนวนของมาตรฐานและตัวชี้วัดแตกต่างกันตามจุดเน้นที่แตกต่างกัน อย่างไรก็ตามแนวคิดการประเมินความก้าวหน้าในการเรียนรู้ ถือเป็นแนวคิดที่ยังไม่แพร่หลายในประเทศไทย โดยเฉพาะอย่างยิ่งการนำมาศึกษาในประเด็นของการประเมินเพื่อพัฒนาการเรียนรู้ของผู้เรียนเนื่องจากเป็นแนวคิดใหม่สำหรับบริบทการศึกษาไทย

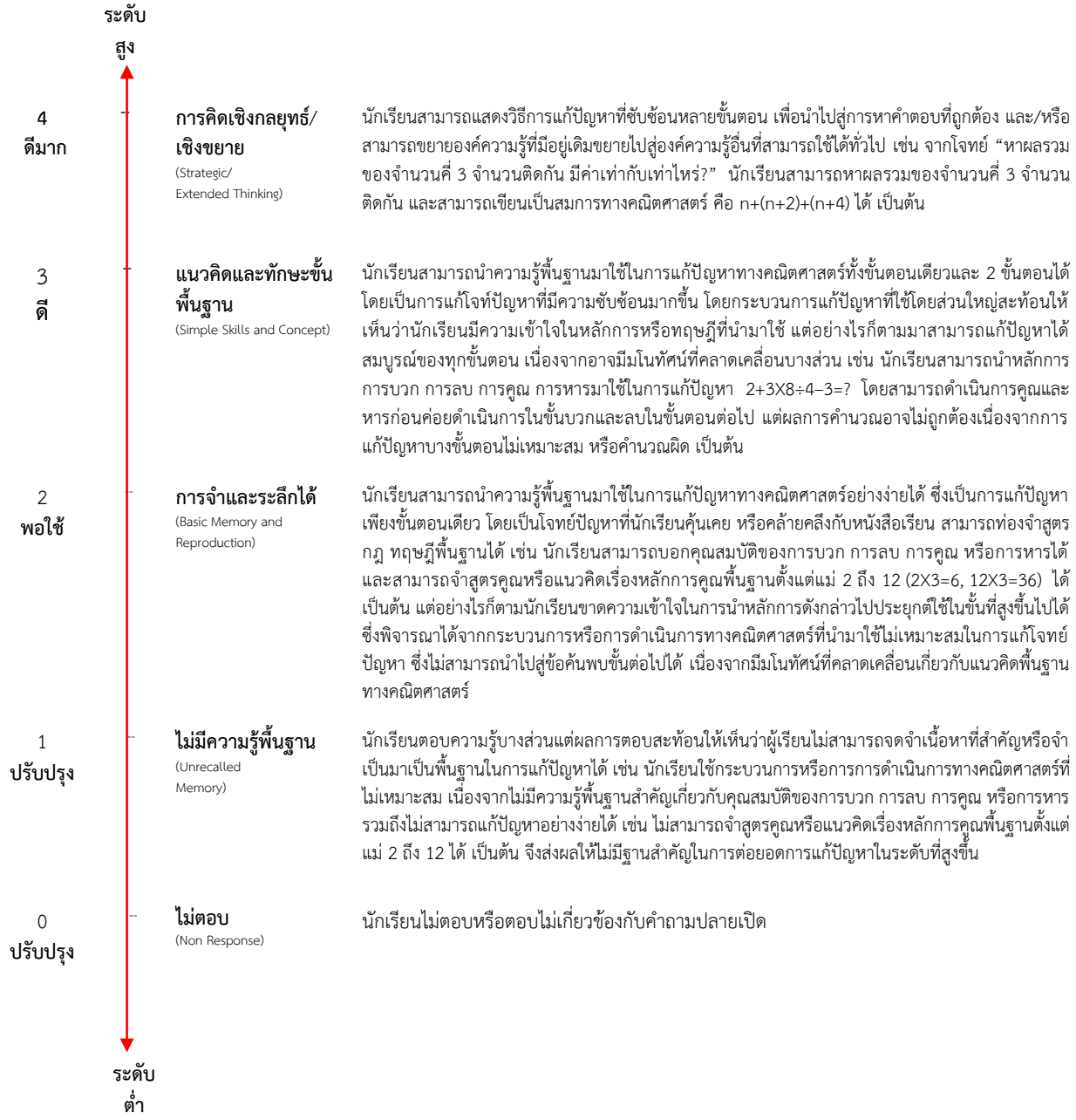
ในปี พ.ศ. 2558-2559 พัชรี จันทร์เพ็งและคณะ ซึ่งเป็นนักวิจัยจากคณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น ร่วมกับ Mark Wilson ซึ่งเป็นศาสตราจารย์ด้านการวัดและประเมินผลการศึกษา ในสาขา Quantitative Method and Evaluation Program ในภาควิชา Organization, Measurement, and Evaluation Department, คณะ Graduate School of Education มหาวิทยาลัย University of California, Berkeley (UC Berkeley) และผู้อำนวยศูนย์วิจัยทางด้านการประเมินและการประเมินผล แห่งมหาวิทยาลัย UC Berkeley โดยมีชื่ออย่างเป็นทางการว่า “The Berekeley Evaluation and Assessment Research (BEAR) Center” ซึ่งเป็นองค์กรที่สำคัญในการพัฒนาระบบการประเมินทางการศึกษาของสหรัฐอเมริกา โดยเฉพาะอย่างยิ่งในมลรัฐ California ได้นำระบบการประเมิน BEAR (BEAR Assessment System) ซึ่งมีชื่อเฉพาะในการพัฒนามาตรวัดทางจิตมิติ (Psychometrics) ว่า “โมเดลเชิงโครงสร้าง (Construct Modelling)” ที่พัฒนาโดย Wilson ในปี ค.ศ. 2005 ซึ่งปัจจุบันได้รับการยอมรับว่าสามารถนำมาใช้ในการประเมินสำหรับผู้เรียนยุคใหม่ได้อย่างเหมาะสมสอดคล้องกับบริบทจริง (National Research Council, 2014) คณะผู้วิจัยจึงได้นำมาเป็นฐานคิดที่สำคัญในการนำมาใช้ในการพัฒนากรอบการประเมินความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของนักเรียนระดับชั้นประถมศึกษาปีที่ 6 กับโรงเรียนที่นำนวัตกรรมการสอนแบบ Lesson Study และ Open Approach มาใช้กับนักเรียนในระดับประถมศึกษาปีที่ 6 เป็นระยะเวลา 1 ปี ในช่วง สิงหาคม 2558- กรกฎาคม 2559 โดยใช้ระเบียบวิธีวิจัยแบบผสมผสานแบบเปลี่ยนรูปตามกาลเวลา (Sequential Transformative Design) โดยให้ความสำคัญกับการนำผลการตรวจสอบที่มีคุณภาพสูงเชิงปริมาณมาตีความหมายของผลการประเมินประกอบกับการใช้ข้อมูลเชิงคุณภาพสนับสนุนผลการวิเคราะห์ที่ได้เพื่อร่วมกันตีความได้อย่างมีความหมาย

ประชากรและกลุ่มตัวอย่างในครั้งนี้เป็นนักเรียนที่เข้าร่วมโครงการการพัฒนาการคิดขั้นสูงทางคณิตศาสตร์ในภาคเหนือและภาคตะวันออกเฉียงเหนือของประเทศไทย ซึ่งเป็นหนึ่งในสี่ของโครงการแก้ไขปัญหาสาธารณสุขและการศึกษาในภาคตะวันออกเฉียงเหนือเพื่อลดความเหลื่อมล้ำทางสังคม โดยดำเนินการเพื่อสนองตอบเจตนารมณ์ 50 ปี แห่งการอุทิศการอุดมศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการและคณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น ภายใต้การดูแลของศูนย์วิจัยคณิตศาสตรศึกษา มหาวิทยาลัยขอนแก่น โดยมี ผู้ช่วยศาสตราจารย์ ดร. ไมตรี อินทร์ประสิทธิ์ คณบดีคณะศึกษาศาสตร์ เป็นผู้มีบทบาทสำคัญในการพัฒนาครูโดยการนำนวัตกรรมการศึกษาชั้นเรียน (Lesson Study) และวิธีการแบบเปิด (Open Approach) มาใช้เป็นครั้งแรกในประเทศไทยตั้งแต่ปี พ.ศ. 2545 ที่โรงเรียนคูคำพิทยาสรรพ์ และโรงเรียนชุมชนบ้านชนบท และได้ขยายต่อออกไปโดยครอบคลุมทั้งภาคตะวันออกเฉียงเหนือและ 4 จังหวัดในภาคเหนือ บางจังหวัดในภาคใต้ รวมทั้งโรงเรียนบางแห่งในกรุงเทพมหานคร รวมจำนวน 22

โรงเรียน ในปี 2556 (โครงการพัฒนาการคิดขั้นสูงทางคณิตศาสตร์ของนักเรียนในเขตภาคตะวันออกเฉียงเหนือ, 2559) ซึ่งปัจจุบันมีการนำแนวคิดดังกล่าวมาใช้ในการพัฒนาการเรียนรู้และการจัดการเรียนการสอนในชั้นเรียนได้เกิดผลเป็นรูปธรรม

โดยหัวใจสำคัญของการนำรูปแบบการสอนดังกล่าวมาใช้คือไม่ได้เน้นเฉพาะผลลัพธ์การเรียนรู้ของผู้เรียนเท่านั้น แต่มุ่งเน้นไปที่กระบวนการคิดของผู้เรียนด้วยเช่นกัน แต่อย่างไรก็ตามมักมีคำถามเกิดขึ้นเสมอว่าจะมีวิธีการออกแบบการประเมินอย่างไร ที่สามารถประเมินได้สอดคล้องกับรูปแบบการสอนดังกล่าว คณะผู้วิจัยจึงได้พัฒนากรอบการประเมินความก้าวหน้าในการเรียนรู้ความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของผู้เรียนที่เน้นทั้งกระบวนการและผลลัพธ์ของการแก้ปัญหาทางคณิตศาสตร์ของผู้เรียนเพื่อนำข้อมูลที่ได้มาใช้ในการประเมินการเรียนรู้ผู้เรียนควบคู่กันกับการจัดการเรียนรู้ในชั้นเรียน และประเมินผลลัพธ์การเรียนรู้ผู้เรียนได้อย่างมีความถูกต้องและน่าเชื่อถือ

คณะผู้วิจัยจึงทำการศึกษาความเป็นไปได้และหาแนวทางในการพัฒนากรอบการประเมินความสามารถทางคณิตศาสตร์ และตอบโจทย์วิจัยดังกล่าวในระยะแรก โดยคณะผู้วิจัยร่วมกับศูนย์การพัฒนาระบบการประเมินแห่งวิทยาลัย UC Berkeley ประเทศสหรัฐอเมริกา ทำการศึกษากับครู และนักเรียนระดับชั้นประถมศึกษาปีที่ 6 กับโรงเรียนที่นำนวัตกรรมการศึกษาชั้นเรียน (Lesson Study) และวิธีการแบบเปิด (Open Approach) มาใช้กับนักเรียนในระดับประถมศึกษาปีที่ 6 เป็นระยะเวลา 1 ปี ในช่วง สิงหาคม 2558- กรกฎาคม 2559 โดยนำระบบการประเมินตามโมเดลเชิงโครงสร้าง (Construct Modelling) (Wilson, 2005; Wilson, 2009; Kennedy, & Wilson, 2006; Black, Wilson, & Yao, 2011; Wilson et al., 2012) เป็นฐานในการพัฒนากรอบการประเมินดังกล่าว ทำให้ พัชรี จันทร์เพ็งและคณะสามารถพัฒนากรอบการประเมิน 3 ด้าน คือ กระบวนการทางคณิตศาสตร์ (Mathematical Procedures) กลยุทธ์ทางคณิตศาสตร์ (Mathematical Strategies) และการให้เหตุผลทางคณิตศาสตร์ (Mathematical Reasoning) ซึ่งมุ่งเน้นกระบวนการทางความคิด (Cognitive Process) ปรากฏเป็นแผนภาพความก้าวหน้า (Progress/Construct Learning) (Junpeng, Wilson, Inprasite, & Marweng, unpublished data; Junpeng, & Wilson, in press) ได้ทำการพัฒนากรอบการประเมินที่เหมาะสมสำหรับการประเมินความก้าวหน้าในการเรียนรู้ แบ่งออกเป็น 3 ด้าน ในที่นี้จะนำเสนอตัวอย่างความก้าวหน้าในการเรียนรู้ของการแก้ปัญหาทางคณิตศาสตร์ด้านกระบวนการทางคณิตศาสตร์ ซึ่งมีรายละเอียดของระดับ ชื่อ และคำอธิบายระดับความสามารถในการแก้โจทย์ปัญหาดังแผนภาพที่ 3




**แผนภาพที่** 3 แผนที่ความก้าวหน้าในการเรียนรู้ของการแก้ปัญหาทางคณิตศาสตร์ด้านกระบวนการทางคณิตศาสตร์ (Mathematical Procedures) (Junpeng, & Wilson, 2016)

จากแผนภาพที่ 3 เป็นการพัฒนาแผนที่ความก้าวหน้าด้านกระบวนการแก้ปัญหาทางคณิตศาสตร์ ซึ่งประกอบไปด้วย 5 ระดับ โดยมีรายละเอียดทั้ง 5 ระดับ โดยมีการให้คะแนนตั้งแต่ 0 – 4 คะแนน ตั้งแต่ระดับปรับปรุงจนถึงระดับดีมาก โดยมีรายละเอียดระดับ คำอธิบายในแต่ละระดับ แนวทางการให้คะแนน และแนวทางการปรับปรุงแก้ไขผู้เรียนเพื่อพัฒนาให้ก้าวหน้าขึ้น ดังต่อไปนี้

**ระดับที่ 1** มีการตรวจให้คะแนนเท่ากับ 0 คะแนน ซึ่งอยู่ในระดับที่ต้องปรับปรุงแก้ไขอย่างเร่งด่วน สำหรับครูใรการปรับปรุงพัฒนาผู้เรียน เรียกระดับนี้ว่า “ไม่มีความรู้พื้นฐาน” (Unrecalled

Memory) เนื่องจากนักเรียนไม่ตอบหรือตอบไม่เกี่ยวข้องกับโจทย์ปัญหา

**ระดับที่ 2** มีการตรวจให้คะแนนเท่ากับ 1 คะแนน อยู่ในระดับปรับปรุง โดยเรียกระดับนี้ว่า เรียกว่าไม่มีความรู้พื้นฐาน (Unrecalled Memory) เป็นระดับที่นักเรียนตอบความรู้บางส่วนแต่ผลการตอบสะท้อนให้เห็นว่าผู้เรียนไม่สามารถจดจำเนื้อหาที่สำคัญหรือจำเป็นมาเป็นพื้นฐานในการแก้ปัญหาได้ เช่น นักเรียนใช้กระบวนการหรือการการดำเนินการทางคณิตศาสตร์ที่ไม่เหมาะสม เนื่องจากไม่มีความรู้พื้นฐานสำคัญเกี่ยวกับคุณสมบัติของการบวก การลบ การคูณ หรือการหาร รวมถึงไม่สามารถแก้ปัญหาอย่างง่ายได้ เช่น ไม่สามารถจำสูตรคูณหรือแนวคิดเรื่องหลักการคูณพื้นฐานตั้งแต่แม่ 2 ถึง 12 ได้ เป็นต้น จึงส่งผลให้ไม่มีฐานสำคัญในการต่อยอดการแก้ปัญหาในระดับที่สูงขึ้น จำเป็นอย่างยิ่งที่ครูต้องพัฒนาผู้เรียนให้มีความรู้พื้นฐานที่สำคัญเพื่อเป็นฐานในการแก้ปัญหาในขั้นตอนเดียวหรือหลายขั้นตอนได้

**ระดับที่ 3** มีการตรวจให้คะแนนเท่ากับ 2 คะแนน อยู่ในระดับพอใช้ เรียกระดับนี้ว่า “การจำและระลึกได้” (Basic Memory and Reproduction) เป็นระดับที่นักเรียนสามารถนำความรู้พื้นฐานมาใช้ในการแก้ปัญหาทางคณิตศาสตร์อย่างง่ายได้ ซึ่งเป็นการแก้ปัญหาเพียงขั้นตอนเดียว โดยเป็นโจทย์ปัญหาที่นักเรียนคุ้นเคย หรือคล้ายคลึงกับหนังสือเรียน สามารถท่องจำสูตร กฎ ทฤษฎีพื้นฐานได้ เช่น นักเรียนสามารถบอกคุณสมบัติของการบวก การลบ การคูณ หรือการหารได้ และสามารถจำสูตรคูณหรือแนวคิดเรื่องหลักการคูณพื้นฐานตั้งแต่แม่ 2 ถึง 12 (2X3=6, 12X3=36) ได้ เป็นต้น แต่อย่างไรก็ตามนักเรียนขาดความเข้าใจในการนำหลักการดังกล่าวไปประยุกต์ใช้ในขั้นที่สูงขึ้นไปได้ ซึ่งพิจารณาได้จากกระบวนการหรือการดำเนินการทางคณิตศาสตร์ที่นำมาใช้ไม่เหมาะสมในการแก้โจทย์ปัญหา ซึ่งไม่สามารถนำไปสู่ข้อค้นพบขั้นต่อไปได้ เนื่องจากมีมโนทัศน์ที่คลาดเคลื่อนเกี่ยวกับแนวคิดพื้นฐานทางคณิตศาสตร์ ครูควรพัฒนาผู้เรียนให้สามารถแก้ปัญหาที่ซับซ้อนมากกว่า 1 ขั้นตอน และเน้นให้นักเรียนสามารถนำความรู้ไปสู่การนำไปใช้แก้ปัญหาในชีวิตประจำเป็นวันมากขึ้น

**ระดับที่ 4** มีการตรวจให้คะแนนเท่ากับ 3 คะแนน อยู่ในระดับดี เรียกระดับนี้ว่า “แนวคิดและทักษะขั้นพื้นฐาน” (Simple Skills and Concept) เป็นระดับที่นักเรียนสามารถนำความรู้พื้นฐานมาใช้ในการแก้ปัญหาทางคณิตศาสตร์ทั้งขั้นตอนเดียวและ 2 ขั้นตอนได้ โดยเป็นการแก้โจทย์ปัญหาที่มีความซับซ้อนมากขึ้น โดยกระบวนการแก้ปัญหาที่ใช้โดยส่วนใหญ่สะท้อนให้เห็นว่านักเรียนมีความเข้าใจในหลักการหรือทฤษฎีที่นำมาใช้ แต่อย่างไรก็ตามมาสามารถแก้ปัญหาได้สมบูรณ์ของทุกขั้นตอน เนื่องจากอาจมีมโนทัศน์ที่คลาดเคลื่อนบางส่วน เช่น นักเรียนสามารถนำหลักการ การบวก การลบ การคูณ การหารมาใช้ในการแก้ปัญหา 2+3X8÷4–3=? โดยสามารถดำเนินการคูณและหารก่อนค่อยดำเนินการในขั้นบวกและลบในขั้นตอนต่อไป แต่ผลการคำนวณอาจไม่ถูกต้องเนื่องจากการแก้ปัญหาบางขั้นตอนไม่เหมาะสม หรือคำนวณผิด เป็นต้น ครูควรมุ่งเน้นให้ผู้เรียนสามารถแก้ปัญหา โดยอาศัยความรู้และความเข้าใจขั้นพื้นฐานนำไปสู่การแก้โจทย์ที่สอดคล้องกับสถานการณ์ในชีวิตประจำวันได้ที่มีซับซ้อนหลายขั้นตอนในโจทย์ที่ไม่คุ้นเคยมาก่อนได้ เพื่อให้สามารถบูรณาการ ต่อยอดหรือขยายไปยังองค์ความรู้อื่นหรือหรือหาอื่นมากขึ้น

**ระดับที่ 5** มีการตรวจให้คะแนนเท่ากับ 4 คะแนน อยู่ในระดับดีมาก เรียกว่า “การคิดเชิงกลยุทธ์/เชิงขยาย” (Strategic/Extended Thinking) เป็นระดับที่นักเรียนสามารถแสดงวิธีการแก้ปัญหาได้อย่างหลากหลายวิธีในโจทย์ที่ซับซ้อนหรือมีการแก้ปัญหาหลายขั้นตอน เพื่อนำไปสู่การหาคำตอบที่ถูกต้อง หรือสามารถขยายองค์ความรู้ที่มีอยู่เดิมขยายไปสู่องค์ความรู้อื่นที่สามารถใช้ได้ทั่วไป เช่น จากโจทย์ “หาผลรวมของจำนวนคี่ 3 จำนวนติดกัน มีค่าเท่ากับเท่าไหร่?” นักเรียนสามารถหาผลรวมของจำนวนคี่ 3 จำนวนติดกันและสามารถเขียนเป็นสมการทางคณิตศาสตร์ คือ n+(n+2)+(n+4) ได้ เป็นต้น

ผลการวิจัยในครั้งนี้เป็นฐานสำคัญในการพัฒนากรอบการแก้ปัญหาทางคณิตศาสตร์ที่สามารถนำมาใช้ในบริบทของประเทศไทยได้ อย่างไรก็ตามจากผลการวิจัย พบว่ายังมีจุดอ่อนหลายประการ โดยเฉพาะการกำหนดมิติในการประเมิน โมเดลที่ใช้ประเมิน เครื่องมือในการประเมิน และวิธีการให้คะแนน ที่จำเป็นต้องมีการแก้ไขปรับปรุงอีกครั้ง (Re-design) ข้อค้นพบที่ได้จากการศึกษาในครั้งนี้ได้จุดประกายสำคัญให้ผู้วิจัยนำวิธีการดังกล่าวไปต่อยอดสู่วิธีการกำหนดคะแนนมาตรฐาน และคะแนนจุดตัดที่เป็นแกนกลางได้ และ นำไปสู่การกำหนดเป็นคะแนนจุดตัดในระดับประเทศ ที่น่าเชื่อถือได้เนื่องจากมีการใช้โมเดลการวัด ตรวจสอบความตรง ความเที่ยงและความยุติธรรมของการประเมิน ที่นำไปสู่การแปลความหมายของผล การประเมินได้ และสามารถนำไปสู่การออกแบบการประเมินในชั้นเรียนที่สอดรับกับการประเมิน ระดับชาติ ในขณะเดียวกันแนวคิดโมเดลเชิงโครงสร้างจะเป็นฐานสำคัญในการยอดไปถึงการรายงานผล ของผู้เรียนที่สะท้อนให้เห็นถึงจุดเด่นและจุดที่ควรพัฒนา ทั้งรายบุคคลและภาพรวมได้ เช่นเดียวกัน เนื่องจากแนวคิดดังกล่าวจะนำไปสู่ข้อค้นพบและองค์ความรู้ใหม่ที่จะเป็นประโยชน์ต่อการพัฒนาประเทศในอนาคต

**3.4 โจทย์วิจัยที่เหมาะสมต่อการนำโมเดลประเมินความก้าวหน้าในการเรียนรู้ไปใช้ในบริบทของภาคตะวันออกเฉียงเหนือ และแนวโน้มการขยายผลการนำไปใช้ในอนาคต**

ในช่วงปี พ.ศ. 2558-2559 ภายหลังจากที่พัชรี จันทร์เพ็งและคณะ ได้ทำการพัฒนากรอบการประเมินความก้าวหน้าในการเรียนรู้โดยเริ่มแรกมุ่งศึกษากับกลุ่มสาระการเรียนรู้คณิตศาสตร์ ในระดับประถมศึกษาปีที่ 6 จำนวน 1 ระดับชั้น โดยมีการปรับเปลี่ยนให้เหมาะสมกับบริบทจริงในเขตภาคตะวันออกเฉียงเหนือที่มีความหลากหลายของวัฒนธรรมในแต่ละสถานศึกษา ซึ่งปัจจัยสำคัญที่จะสามารถขับเคลื่อนแนวคิดการประเมินความก้าวหน้าในการเรียนรู้ให้เกิดผลในทางปฏิบัติได้นั้น จำเป็นต้องมีกำลังในการขับเคลื่อน โดยเฉพาะการพัฒนาวิชาชีพครู ดังนั้น พัชรี จันทร์เพ็ง ซึ่งเป็นอาจารย์ประจำอยู่สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น ได้นำแนวคิดนี้ให้เกิดผลเป็นรูปธรรม โดยนำไปเป็นส่วนหนึ่งในการพัฒนาวิชาชีพ โดยสอดแทรกเข้าไปในหลักสูตรทั้งในระดับปริญญาตรี ปริญญาโท และปริญญาเอก ใน 5 รายวิชา คือ (1) วิชาการวัดและประเมินการศึกษาในระดับปริญญาตรี สาขาวิทยาศาสตรศึกษา และสาขาสังคมศึกษา (2) วิชาทฤษฎีการวัดและการทดสอบ ระดับปริญญาโท สาขาการวัดและประเมินผลการศึกษา (3) วิชาระเบียบวิธีวิจัยการศึกษา ในระดับปริญญาโท สาขาวิทยาศาสตรศึกษา (4) วิชาระเบียบวิธีวิจัยขั้นสูงทางด้านการวัดและประเมินผลการศึกษา สาขาการวัดและประเมินผลการศึกษา และ (5) วิชากระบวนทัศน์และระเบียบวิธีวิจัยขั้นสูงทางวิทยาศาสตรศึกษา ในระดับปริญญาเอก สาขาวิทยาศาสตรศึกษา โดยรวมทั้งหมด 150 คน ซึ่งอยู่ในช่วงของการดำเนินการและพัฒนากรอบการประเมิน ภายใต้โจทย์ปัญหาที่สำคัญคือ

(1) อะไรคือเป้าหมายสำคัญของการประเมินผู้เรียนยุคใหม่ในบริบทที่แตกต่างกัน?

(2) ทำอย่างไรจึงจะเกิดการประเมินเพื่อพัฒนาความสามารถผู้เรียนอย่างแท้จริง?

(3) ทำอย่างไรครูและบุคลากรทางการศึกษาสามารถวินิจฉัยผู้เรียนและสามารถประเมิน ความก้าวหน้าของผู้เรียน ว่าขณะนี้ผู้เรียนอยู่ ณ ตำแหน่งของความสามารถ มีลักษณะใดบ้างที่ผู้เรียน สามารถทำได้ และมีลักษณะใดบ้างที่ต้องการพัฒนาเพื่อนำไปการปรับปรุงแก้ไขและไปถึงมาตรฐานที่ตั้งไว้ โดยคำนึงถึงความแตกต่างระหว่างบุคคล?

(4) เมื่อบริบทในชั้นเรียนที่แตกต่างกัน วิธีการเรียนการสอนที่แตกต่างกัน เครื่องมือการประเมินที่ต่างกัน วัฒนธรรมองค์กรที่ต่างกัน อะไรจะเป็นหลักยึดให้อยู่ในเป้าหมายเดียวกัน และสามารถนำไปสู่การปฏิบัติได้จริง?

(5) อะไรเป็นแนวคิดสำคัญสำคัญในการพัฒนาการกำหนดมาตรฐานการประเมิน การออกแบบการประเมิน และการรายงานผล ที่เรียบง่าย เอาไปใช้ได้จริง แต่มีความน่าเชื่อถือและถูกต้อง?

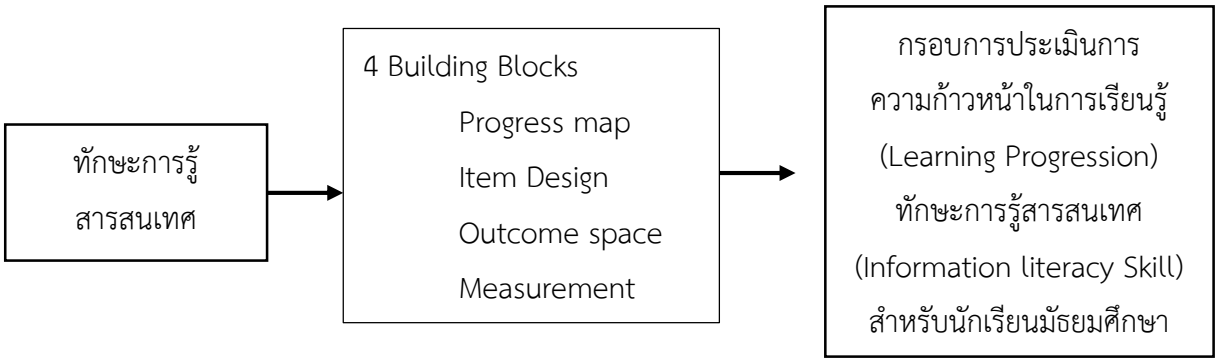
(6) ทำอย่างไรจึงจะมีการบูรณาการการประเมินร่วมกันระหว่างระดับชาติและชั้นเรียน?

(7) การประเมินกระบวนการและผลลัพธ์ของผู้เรียน จะมีวิธีการออกแบบการประเมินอย่างไร โดยมุ่งเน้นไปที่กระบวนการทางความคิด ไม่เน้นเพียงคำตอบถูกหรือผิดเพียงอย่างเดียวเท่านั้น และอะไรจะเป็นเครื่องยืนยันว่าการออกแบบนั้นมีความตรง ความเที่ยง และยุติธรรม?

(8) ควรมีการรายงานการประเมินอย่างไร ที่ง่าย และสามารถนำไปใช้ได้จริง โดยสะท้อนให้เห็นถึงจุดเด่นและจุดที่ควรพัฒนาของผู้เรียนทั้งรายบุคคลและภาพรวม?

จากที่กล่าวมาจะเห็นได้ว่าแนวคิดการประเมินความก้าวหน้าในการเรียนรู้ได้แผ่ขยายไปสู่กลุ่มสาระอื่นนอกเหนือจากคณิตศาสตร์ รวมถึงสาขาวิชาอื่นที่เกี่ยวข้องเพิ่มเติม ครอบคลุมทั้งนักศึกษาที่กำลังปฏิบัติการสอนจริงในสถานศึกษา ครูในระดับสถานศึกษาทั้งประถมศึกษาและมัธยมศึกษา และครอบคลุมไปยังอาจารย์มหาวิทยาลัย ผู้บริหารศึกษา และศึกษานิเทศน์ ที่ได้ทดลองนำระบบดังกล่าวไปใช้ในการพัฒนาในวิชาชีพของตนเอง นอกจากนี้ในระดับอุดมศึกษาได้มีการดำเนินการนำวิธีการดังกล่าวไปปรับใช้ในการพัฒนากรอบการประเมินพนักงานและบุคลากรในระดับมหาวิทยาลัย สำหรับการนำระบบดังกล่าวไปดำเนินการใช้จริงในสถานศึกษา เมื่อพิจารณาถึงการนำมาแก้ปัญหาจริงในชั้นเรียนโดยเน้นไปที่การพัฒนาวิธีการประเมินของครูในชั้นเรียนและการพัฒนาผู้เรียนไปพร้อมกัน พัชรี จันทร์เพ็งและคณะได้เริ่มดำเนินการทำงานร่วมกับผู้บริหารและครูจำนวน 2 กลุ่ม สาระการเรียนรู้ คือ คณิตศาสตร์และภาษาต่างประเทศ ที่โรงเรียนสาธิตศึกษาศาสตร์ มหาวิทาลัยขอนแก่น เพื่อให้เกิดผลเป็นรูปธรรมและเกิดผลต่อผู้เรียนอย่างแท้จริง โดยแนวโน้มในอนาคต จะมีการขยายไปยังนักศึกษาสาขาวิชาอื่น ของคณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น และสาขาวิชาชีพอื่นในอนาคต ซึ่งอยู่ในช่วงของการวางแผนดำเนินการขยายผล เพื่อสามารถนำระบบดังกล่าวไปสู่การประยุกต์ใช้ให้เหมาะสมกับสภาพปัญหาของในแต่บริบท

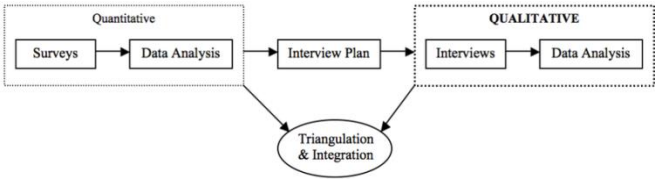
**6. กรอบแนวคิดในการวิจัย**




**แผนภาพที่** 4 กรอบแนวคิดในการวิจัย

**4. ระเบียบวิธีวิจัย (Research Methodology)**

สำหรับการวิจัยครั้งนี้คณะวิจัยได้มีการใช้รูปแบบการวิจัยแบบผสมผสาน (Mixed Method Research) โดยมีการเก็บข้อมูลทั้งเชิงปริมาณและเชิงคุณภาพ เพื่อเป็นระเบียบวิธีวิจัยที่สำคัญในการการพัฒนากรอบการประเมินความก้าวหน้าในการเรียนรู้ของนักเรียน โดยมีลักษณะที่เรียกว่า “แบบเปลี่ยนรูปตามกาลเวลา” (Sequential Transformative Design) ตามแนวคิดของ Cresswell (2011)




โดยในระยะแรกเริ่มต้นด้วยวิธีการเชิงปริมาณและตามด้วยวิธีการเชิงคุณภาพ จากนั้นมีการบูรณาการวิธีการเชิงปริมาณและคุณภาพเข้าด้วยกันในระยะตีความหมาย โดยอาศัยกรอบโมเดลเชิงโครงสร้าง (Construct Modelling) (Wilson, 2005)มาเป็นแนวทางในการพัฒนากรอบการประเมินใน 4 ขั้นตอน (Four Building Blocks) คือ (1) แผนที่ความก้าวหน้า (Progress Map) (2) การออกแบบข้อสอบ/ภาระงาน (Item Design) (3) การให้คะแนนของผลลัพธ์การเรียนรู้ (Outcome Space) และ (4) โมเดลการวัด (Measurement Model/Wright Map รายละเอียดในแต่ละขั้นตอนดังนี้

**4.1 แผนที่ความก้าวหน้า (Progress Map)**

การประเมินทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษา (ILS) มีการพัฒนาและได้มีการกล่าวถึง มาตรฐานการรู้สารสนเทศระดับมัธยมศึกษา มาตรฐานการรู้สารสนเทศ 9 มาตรฐาน สำหรับนักเรียนในระดับโรงเรียน สมาคมห้องสมุดโรงเรียนแห่งสหรัฐอเมริกา (American Association of School Libraries: AASL, 1998), (ศิวราช ราชพัฒน์, 2547: 19), (California Media and Library Educator Association, 1994, pp.6-7) โดยในมาตรฐานการรู้สารสนเทศระดับมัธยมศึกษานั้นได้กำหนดไว้หลายมาตรฐานและหลายตัวชี้วัด ผู้ศึกษาจึงได้เลือก หมวดคุณลักษณะของการรู้สารสนเทศ (Information literacy) และมาตรฐานที่ 2 นักเรียนที่เป็นผู้รู้สารสนเทศสามารถประเมินสารสนเทศได้อย่างมีหลักการ

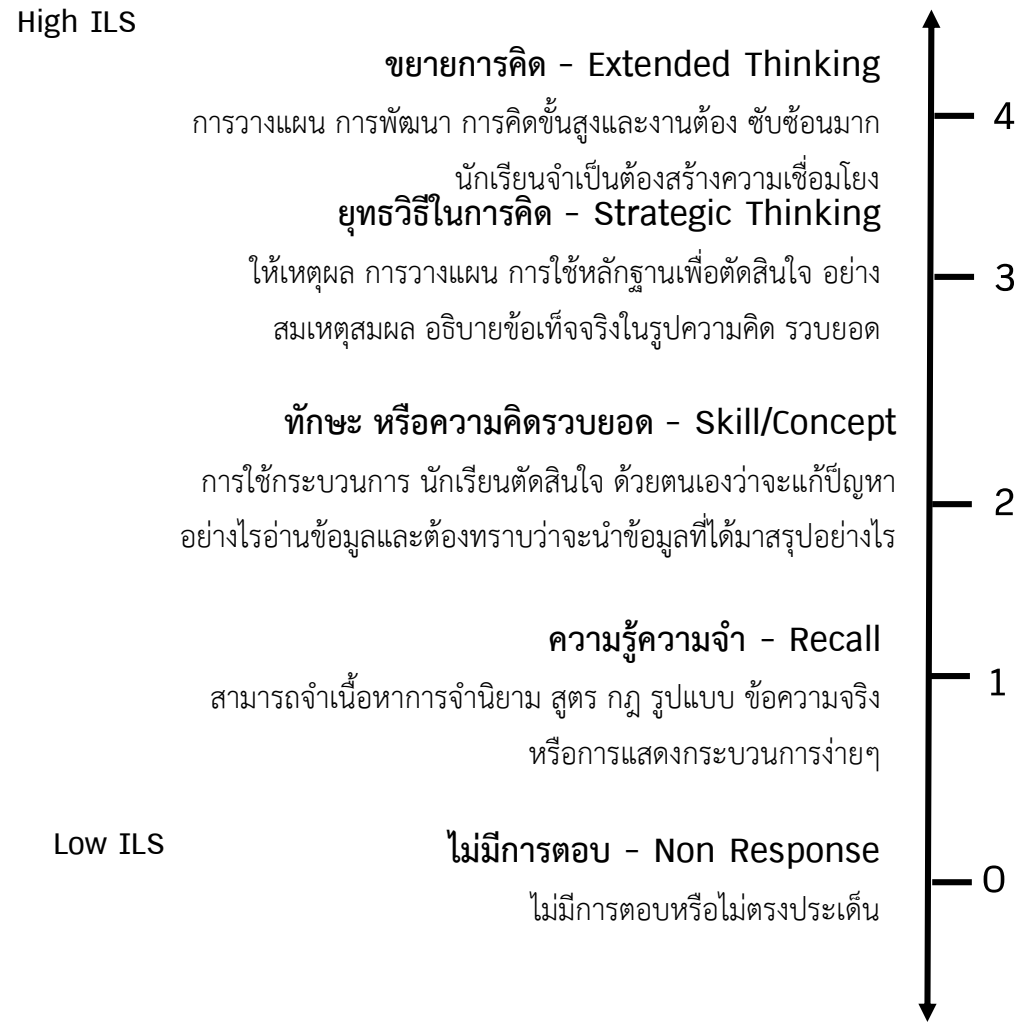
ในการศึกษาครั้งนี้ได้พัฒนาคำถามบนพื้นฐานของคำถามปลายเปิด (Open Ended Question) เป็นคำถามที่เปิดโอกาสให้ผู้ตอบสามารถตอบได้อย่างเต็มที่ ซึ่งคาดว่าน่าจะได้คำตอบที่แน่นอน สมบูรณ์ ตรงกับสภาพความเป็นจริง และเนื้อหาได้ใช้รายวิชาเทคโนโลยีสารสนเทศ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานในด้านสมรรถนะของผู้เรียน คุณลักษณะอันพึงประสงค์ รวมทั้งตัวชี้วัดซึ่งระบุสิ่งที่นักเรียนพึงรู้และปฏิบัติได้ เพื่อศึกษารูปแบบ กระบวนการได้มาซึ่งคำตอบ และแนวความคิดเกี่ยวกับประเด็นที่ตอบ

การออกบบระดับความรู้ความสามารถด้านพุทธิพิสัย ได้ยึดตามองค์การเครือข่ายผลิตเครื่องมือ (Web Alignment Tool, 1999: Online) และ เวปป์ (Webb Norman L, 2002: 4 - 5) เพื่อใช้ในการออกแบบและสร้าง Construct map โดยผู้ศึกษาได้แบ่งระดับออกเป็น 5 ระดับที่แตกต่างกัน ดังนี้ Non Response, ระลึกได้ (Recall), ทักษะ หรือความคิดรวบยอด (Skill/Concept), ยุทธวิธีในการคิด (Strategic Thinking) และขยายการคิด (Extended Thinking)

ทักษะการรู้สารสนเทศ (Information literacy Skill) นั้นมีการแบ่งถึง 3 หมวด 9 มาตรฐาน และมี 29 ตัวชี้วัด เพื่อให้ครอบคลุมและวัดได้ตรง ควรมีการออกแบบเครื่องมือให้ครบทุกด้านให้มากที่สุด

การประเมินทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษา เป็นคำถามบนพื้นฐานของคำถามปลายเปิด (Open Ended Question) โดยมีคะแนนตั้งแต่ 0 – 4 โดยให้นักเรียนเขียนคำตอบลงในข้อคำถามข้อนั้นๆ ได้เลย

โดยที่การพัฒนาการประเมินทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษานั้นจะยึดบนหลักหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานในด้านสมรรถนะของผู้เรียน คุณลักษณะอันพึงประสงค์ รวมทั้งตัวชี้วัดซึ่งระบุสิ่งที่นักเรียนพึงรู้ที่ได้ระบุไว้ แต่ยึดหลักมาตรฐานตามสมาคมห้องสมุดโรงเรียนแห่งสหรัฐอเมริกา (American Association of School Libraries: AASL, 1998) หมวดคุณลักษณะของการรู้สารสนเทศ (Information literacy) และมาตรฐานที่ 2 นักเรียนที่เป็นผู้รู้สารสนเทศสามารถประเมินสารสนเทศได้อย่างมีหลักการ จำนวน 4 ตัวชี้วัด ได้แก่ 1) นักเรียนสามารถตัดสินได้ว่าสารสนเทศนี้นมีความถูกต้อง ตรงประเด็น และครอบคลุมใน สิ่งที่ต้องการ 2) นักเรียนรู้ถึงความแตกต่างระหว่างข้อเท็จจริงและข้อคิดเห็น 3) นักเรียนสามารถระบุถึงสารสนเทศที่คลาดเคลื่อนและไม่เป็นจริงได้ 4) นักเรียนสามารถเลือกสารสนเทศที่เหมาะสมลับปัญหาหรือข้อคำถามที่สงสัยได้



**4.2 การออกแบบข้อสอบ/ภาระงาน (Item Design)**

ผู้วิจัยเลือกพัฒนาเครื่องมือ การประเมินทักษะการรู้สารสนเทศ สำหรับนักเรียนระดับมัธยมศึกษา เนื่องจากว่านักเรียนมักขาดทักษะการรู้สารสนเทศ ผู้ศึกษาจึงอยากศึกษาการวัดและประเมินในเรื่องดังกล่าว เพื่อจัดลำดับ และทำการประเมินเพื่อพัฒนาผู้เรียนให้มีคุณลักษณะอันพึ่งประสงค์ โดยมีการกำหนดขอบเขตของเนื้อหาที่ประเมิน จากผลการเรียนรู้หน่วยการเรียนรู้เรื่องระบบเครือข่าย เมื่อนำผลการเรียนรู้ดังกล่าวมาทำกรอบการประเมิน 4 โครงสร้างการประเมิน (Building Blocks) โดยการสร้างแผนที่เชิงโครงสร้าง (Construct map) จะสร้างจากทฤษฎี และสามารถตรวจสอบจากผู้เชี่ยวชาญได้เช่น construct map ก่อนปรับ

**ตารางที่ 1** รายละเอียดความสามารถของนักเรียนแต่ละระดับ

| Level | Name | Description of Respondent | Examples |
|---|---|---|---|
| 4 | **ขยายการคิด**<br>Extended Thinking | **<u>การวางแผน การพัฒนา การคิดขั้นสูงและงานต้องซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง</u>**<br>- ใช้กระบวนการ วิธีการ หรือหลักการในการแก้ปัญหา ที่มากกว่า 1 วิธีในปัญหานั้นๆ<br>- วิเคราะห์จุดเด่น จุดดีของหลักการ วิธีการ เพื่อเลือกเอากระบวนการนั้นๆ มาปรับประยุกต์ให้เข้ากับวิธีการแก้ปัญหาของตัวเอง<br>- มีกระบวนการสร้างขั้นตอน วิธีการแก้ปัญหา และปรับปรุงเพิ่มเติมการแก้ปัญหา<br>- การวางแผนเพื่อการพัฒนา การออกแบบ การสร้างการเชื่อมโยงระหว่าง ความคิดรวบยอดและข้อเท็จจริง รวบรวม และ สังเคราะห์แนวคิดไปสู่ความคิดรวบยอดใหม่ | |
| | **ยุทธวิธีในการคิด**<br>Strategic Thinking | **<u>ให้เหตุผล การวางแผน การใช้หลักฐานเพื่อตัดสินใจอย่างสมเหตุสมผล อธิบายข้อเท็จจริงในรูปความคิด รวบยอด</u>**<br>- การสร้างข้อคาดเดา อธิบายข้อเท็จจริงในรูปความคิดรวบยอด ใช้ความคิดรวบยอดเพื่อแก้ปัญหา<br>- ระดับที่ต้องการให้นักเรียนให้เหตุผล การวางแผน การใช้หลักฐานเพื่อตัดสินใจ<br>- การแก้ปัญหามความซับซ้อนและเป็นนามธรรม กิจกรรมที่มีคำตอบ ที่เป็นไปได้มากกว่า 1 คำตอบ<br>- นักเรียน พิสูจน์ กิจกรรมลักษณะอื่นๆ ที่อยู่ในระดับ 3 คือ นักเรียนหาข้อสรุปจากการสังเกตหลักฐานและอ้างเหตุผลอย่างสมเหตุสมผล อธิบายข้อเท็จจริงในรูปความคิด รวบ | |

| Level | Name | Description of Respondent | Examples |
|---|---|---|---|
| | | ยอด และใช้ความคิดรวบยอดเพื่อแล้วปัญหา<br>- นักเรียนวิเคราะห์วิธีการ หลักการ อย่างเข้าใจและมีเหตุผล เพื่อใช้ประกอบการตัดสินใจในการเลือกวิธีที่เหมาะสม สำหรับแก้ปัญหา | |
| 2 | **ทักษะหรือความคิดรวบยอด**<br>Skill/Concept | **<u>การใช้กระบวนการ นักเรียนตัดสินใจ ด้วยตนเองว่าจะแก้ปัญหาอย่างไร อ่านข้อมูลและต้องทราบว่าจะนำข้อมูลที่ได้มาสรุปอย่างไร</u>**<br><br>- ระดับการใช้กระบวนการที่เกิดขึ้น ภายในสมอง ในระดับนี้มิความต้องการให้นักเรียนตัดสินใจ ด้วยตนเองว่าจะแก้ปัญหาอย่างไร<br>- นักเรียนสามาถเข้าใจในกระบวน หลักการ วิธีการ รวมถึงทุกขั้นตอนของวิธีที่กล่าวมา เพื่อนำมาตัดสินใจ วางแผน ในการแก้ปัญหา<br>- นักเรียนบอกการตัดสินใจในการเลือกใช้รูปแบบ ประบวนการ วิธีการนั้นๆ<br>- นักเรียนสามารถสรุปวิธีการ จาหเหตุผลหลายๆปัจจัยที่มีส่วนหรือผลต่อคำตอบ<br>- นักเรียนสามารถสรุปคำตอบออกมาเป็นภาษา รูปแบบมโนทัศน์ หรือรูปแบบอื่นๆที่เข้าใจ เป็นของตัวเอง | |
| 1 | **ความรู้ความจำ**<br>Recall | **<u>สามารถจำเนื้อหาการจำนิยาม รูปแบบ ข้อความจริง หรือการแสดงกระบวนการง่ายๆ</u>**<br><br>- นักเรียนสามารถจำเนื้อหา หรือหลักการนั้นๆมาตอบคำถามได้ เช่นบอกความหมาย บอกขั้นตอน บอกวิธีการ แต่ไม่เข้าใจในหลักการแต่ละขั้นตอน<br>- ไม่สามารถระบุหรือทราบถึงว่า สามารถเอาสิ่งที่จำได้นั้นไปใช้ทำอะไรเพิ่มเติมได้<br>- ใช้เพียงแค่ทำที่พบเห็นในเนื้อหา หรือเอกสารประกอบการสอนเท่านั้นมาตอบคำถาม | |

| Level | Name | Description of Respondent | Examples |
|---|---|---|---|
| | | - ไม่สามารถบอกเหตุผล หรือวิธีการที่เลือกเอากระบวนการนั้นๆมาใช้ หรือมาตอบ | |
| 0 | **ไม่มีการตอบ**<br>Non Response | **<u>ไม่มีการตอบหรือไม่ตรงประเด็น</u>**<br><br>ไม่ตอบคำถาม ไม่ส่งข้อสอบ หรือไม่ได้เขียนคำตอบที่เกี่ยวกับคำถาม | |

### 4.3 การให้คะแนนของผลลัพธ์การเรียนรู้ (Outcome Space)

**ตารางที่ 2** แนวทางการให้คะแนนตามแผนที่ความก้าวหน้า

| | | Item | | | | | | | | | Characteristics of Respondents |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1a | 1b | 1c | 2a | 2b | 2c | 3a | 3b | 3c | |
| 4 | **ขยายการคิด** Extended Thinking | | | | | | | | | | การวางแผน การพัฒนา การคิดขั้นสูงและงานต้อง ซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง |
| 3 | **ยุทธวิธีในการคิด** Strategic Thinking | | | ✘ | | | ✘ | | | ✘ | ให้เหตุผล การใช้หลักฐานเพื่อตัดสินใจ อย่างสมเหตุสมผล หรืออธิบายข้อเท็จจริงในรูปความคิดรวบยอด |
| 2 | **ทักษะ หรือความคิดรวบยอด** Skill/Concept | | ✘ | | | ✘ | ✘ | | ✘ | ✘ | การใช้กระบวนการ ตัดสินใจ ด้วยตนเองว่าจะแก้ปัญหาอย่างไร อ่านข้อมูลและต้องทราบว่าจะนำข้อมูลที่ได้มาสรุปอย่างไร |
| 1 | **ความรู้ความจำ** Recall | ✘ | ✘ | ✘ | ✘ | ✘ | ✘ | ✘ | ✘ | ✘ | สามารถจำเนื้อหาการจำนิยาม รูปแบบ ข้อความจริง หรือการแสดงกระบวนการง่ายๆ |
| 0 | **ไม่มีการตอบ** Non Response | ✘ | ✘ | ✘ | ✘ | ✘ | ✘ | ✘ | ✘ | ✘ | ไม่มีการตอบหรือไม่ตรงประเด็น |
| | | 1a | 1b | 1c | 2a | 2b | 2c | 3a | 3b | 3c | |

**4.4 โมเดลการวัด (Measurement Model/Wright Map)**

การประเมินในชั้นเรียนจะต้องมีมาตรฐานทั้งความตรง ความเชื่อถือได้และความยุติธรรม (High Quality Evidence) ดังนั้นในการศึกษาครั้งนี้จึงจำเป็นต้องมีวิธีการในการตรวจสอบคุณภาพของกรอบการประเมินและเครื่องมือที่ใช้ในการประเมินในด้านคุณภาพของข้อสอบรายข้อ และคุณภาพของแบบทดสอบทั้งฉบับ โดยเฉพาะย่างยิ่งการตรวจสอบความตรงเชิงโครงสร้างภายในของเครื่องมือ โดยพิจารณาจากผลการวิเคราะห์ จาก Wright Map ในการวิจัยครั้งนี้คณะผู้วิจัยมีการวิเคราะห์โดยใช้โมเดลราส์ช เป็นฐานสำคัญในการวิเคราะห์ที่เรียกว่าโมเดลโลจิสตามทฤษฎีการตอบสนองแบบพหุมิติ (Multidimensional Random Coefficients Multinomial Logit Model) ในที่นี้คณะผู้วิจัยใช้ชื่อย่อว่า “MRCLM” ที่พัฒนาโดย Adams, Wilson, & Wang ในปี ค.ศ. 1997 ซึ่งเป็นโมเดลที่แผ่ขยายมาจากโมเดลราสช์ (Rash Coefficients Multinomial Logit Model) ในที่นี้ผู้วิจัยใช้ชื่อย่อว่า “RCML” ซึ่งพัฒนาโดย Adams and Wilson ในปี ค.ศ. 1996 โมเดล RCML และ MRCMLM เป็นโมเดลตระกูลราส์ช (Rash-family Models) ซึ่งจะประกอบไปด้วย 2 เมตริกซ์ คือ เมตริกซ์การออกแบบ (Design Metrix) และเมตริกซ์การให้คะแนน (Scoring Metrix) ซึ่งโมเดลเหล่านี้จะมีลักษณะที่สอดคล้องคล้องกับบริบทจริงและง่ายต่อการแปลผลในทางปฏิบัติ (Wilson & Adams, 1995) เนื่องจากมีการประมาณค่าความพารามิเตอร์น้อยกว่าโมเดลประเภทอื่น เช่น โมเดลแบบ 2 และ 3 พารามิเตอร์ เป็นต้น จึงสามารถนำมาประยุกต์ใช้กับกลุ่มตัวอย่างไม่มาก เช่น 30-50 คนได้ ซึ่งสามารถนำมาใช้ในการประเมินในชั้นเรียนได้ คณะผู้วิจัยจึงนำมาประยุกต์ใช้ในการพัฒนากรอบการประเมินความก้าวหน้าในการศึกษาครั้งนี้

MRCMLM เป็นโมเดลที่นิยมใช้ทั่วไปและมีลักษณะยืดหยุ่นเหมาะสำหรับการนำไปใช้ทั้งในบริบทเฉพาะที่มีกลุ่มศึกษาไม่มากรวมถึงกลุ่มที่มีขนาดใหญ่ไปพร้อมกันได้ (Large Scale) (Draney, Yamada, & Xie, 2000) ซึ่งจะใช้เมตริกซ์การออกแบบเพื่อระบุความสัมพันธ์ระหว่างการตอบสนองที่มีต่อข้อสอบแต่ละข้อกับค่าพารามิเตอร์ที่ต้องการประมาณค่าซึ่งถือเป็นโครงสร้างภายใน (Structural Parameters) เพื่อระบุคุณภาพของโครงสร้างภายในของข้อสอบ

สำหรับเมตริกซ์การให้คะแนน (Scoring Metrix) แทนด้วยสัญลักษณ์ $\mathbf{b}_{ik}=(b_{ik1},\ldots,b_{ikV})'$ และทักจะแทนด้วยสัญลักษณ์ว่า B (Scoring Matrix **B**) เป็นเมตริกซ์แสดงความสัมพันธ์ระหว่างข้อสอบและมิติความสามารถที่แท้จริงซึ่งแสดงเป็นฟังก์ชันของการให้คะแนน เช่น การตอบสนองในระดับที่ j ในแผนที่ความก้าวหน้าในข้อที่ i ซึ่งถูกแทนด้วย เวคเตอร์ V ซึ่งเป็นเวคเตอร์แสดงความสามารถของผู้สอบในแต่ละมิติแทนด้วยสัญลักษณ์ vector $\boldsymbol{\theta} = (\boldsymbol{\theta_1, \ldots, \theta_V})$. เช่น ด.ช. บอย มีความสามารถในการแก้ปัญหาทางคณิตศาสตร์ 2 ด้าน คือ ด้านกระบวนการแก้ปัญหาในระดับปานกลาง และด้านการใช้กลยุทธ์ในการแก้ปัญหาระดับสูง แสดงว่า ความสามารถในการแก้ปัญหาครั้งนี้ประกอบด้วย 2 มิติ ซึ่งจะมีความสามารถที่แท้จริง 2 ค่า คือ $\boldsymbol{\theta_1}$ และ $\boldsymbol{\theta_2}$ ซึ่งจำนวนของค่าความสามารถที่แท้จริงของผู้สอบแต่ละคนจะขึ้นอยู่กับจำนวนมิติหรือจำนวนด้านที่ต้องการศึกษานั่นเอง

จำนวนค่าพาราเตอร์ของข้อสอบ แทนด้วย p แทนด้วยเวคเตอร์ $\boldsymbol{\xi}=(\boldsymbol{\xi_1, \ldots, \xi_p})$ ซึ่งแสดงความสัมพันธ์ระหว่างการตอบสนองข้อสอบกับค่าพารามิเตอร์ ซึ่งในที่นี้เรียกว่าเมตริกซ์การออกแบบ A (Design Matrix **A**) แสดงด้วยสัญลักษณ์ $\mathbf{a}_{ij}$ ในแต่ละค่าพารามิเตอร์ซึ่งเมื่อนำมาคูณด้วยเวคเตอร์ $\boldsymbol{\xi}$ ซึ่งเป็นการดำเนินการรวมกันเชิงเส้นตรงของค่าพารามิเตอร์สำหรับการตอบสนองข้อสอบระดับ j ข้อที่ i

ความน่าจะเป็นสำหรับการตอบสนองระดับที่ j ของข้อที่ i สามารถแสดงได้ดังสมการที่ 1

$$\mathrm{P}(X_{ij}=1|A,B,\xi|\theta) = \frac{\exp(b_{ij}\theta+a'_{ij}\xi)}{\sum_{k=1}^{J_i}\exp(b_{ik}\theta+a'_{ik}\xi)} \quad (1)$$

**5. กลุ่มผู้สอบ (Examinees)**

นักเรียนระดับชั้นมัธยมศึกษาตอนปลาย โรงเรียนประชารัฐวิทยาเสริม อ.พระยืน จ.ขอนแก่น เพศชายจำนวน 30 คน และเพศหญิงจำนวน 30 คน ในภาคเรียนที่ 2 ปีการศึกษา 2559

**6. การเก็บรวบรวมข้อมูล (Data Collection)**

1. กำหนดจุดมุ่งหมายในการสร้างแบบประเมิน
2. ศึกษาเอกสารต่าง ๆ และ หลักสูตรมัธยมศึกษาตอนปลายของโรงเรียนประชารัฐวิทยาเสริม อำเภอพระยืน จังหวัดขอนแก่น ในรายวิชาคอมพิวเตอร์ พุทธศักราช 2550 เพื่อกำหนดขอบเขตเนื้อหาในการสร้าง ตลอดจนศึกษาทฤษฎีและหลักการสร้าง ซึ่งการวัดการปฏิบัติเน้นวิธีการปฏิบัติงานและผลงานเท่า ๆ กัน
3. วิเคราะห์งานและจุดมุ่งหมายเชิงพฤติกรรม ตามเนื้อหา เอกสารหลักสูตรการเรียนการสอนวิชาคอมพิวเตอร์ คู่มือครู และหนังสือแบบเรียนวิชาคอมพิวเตอร์ ชั้นมัธยมศึกษาปีที่ 4-6 จากนั้นเขียนเป็นจุดมุ่งหมายเชิงพฤติกรรมและการปฏิบัติงาน
4. นำผลการวิเคราะห์งานมาเขียนขั้นตอนการปฏิบัติงานที่วัด แล้วสร้าง Progress map และดำเนินการให้ผู้เชี่ยวชาญตรวจสอบ Progress map เพื่อปรับแก้
5. ออกแบบข้อสอบ (Item Design) ให้สอดคล้องกับ Progress map แล้วให้ผู้เชี่ยวชาญตรวจสอบว่าข้อสอบดังกล่าวอยู่ระดับใดใน Progress map และภาษาที่อยู่ในข้อสอบมีความชัดเจนมากน้อยเพียงใด
6. นำข้อสอบที่ออกแบบ และปรับแก้ มา Try out ครั้งที่ 1 กับนักเรียน 3 คน เพื่อให้นักเรียน Think aloud ออกมาว่าตนคิดแบบใด ไม่เข้าใจส่วนใดในข้อสอบ
7. นำการคิดของนักเรียนมาปรับปรุง และลองตรวจข้อสอบ เพื่อปรับแก้การให้คะแนน (Outcome space)
8. นำข้อสอบที่ได้จากการปรับปรุงครั้งแรกมาใช้กับกลุ่มตัวอย่าง 20 คน แล้วตรวจคะแนนให้นักเรียน

**7. การวิเคราะห์ข้อมูล (Data Analysis)**

สำหรับการวิจัยในครั้งนี้คณะผู้วิจัยประยุกต์ใช้โมเดล MRCML (Adams, Wilson, & Wang, 1997) โดยผสมผสานร่วมกันระหว่าง Rasch Model (Rasch, 1960) และ Partial Credit Model (Masters, 1982) ดังที่ได้กล่าวมาแล้วข้างต้นในหัวข้อโมเดลการวัด ซึ่งนำมาใช้ในการตรวจสอบความตรงของแผนที่ความก้าวหน้า ซึ่งศึกษาใน **1** แผนที่ความก้าวหน้า โดยใช้โปรแกรม ACER ConQuest Version 2.0 (Wu, Adam, Wilson, & Haldane, 2007) โดยแยกผลการวิเคราะห์ตามวัตถุประสงค์การวิจัย ดังนี้

**7.1** การพัฒนาและตรวจสอบคุณภาพของกรอบการประเมินการความก้าวหน้าในการเรียนรู้ (Learning Progression) การประเมินทักษะการรู้สารสนเทศ สำหรับนักเรียนระดับมัธยมศึกษา จากทั้งการสัมภาษณ์เชิงลึกผู้ทรงคุณวุฒิ การถอดโปรโตคอลการคิดของผู้เรียน ผนวกกับข้อมูลเชิงปริมาณที่ได้จากการตรวจสอบคุณภาพของกรอบการประเมิน ซึ่งพิจารณารายละเอียดดังต่อไปนี้

**(1)** คุณภาพของข้อสอบรายข้อ โดยนำผลที่ได้สู่การปรับปรุงและพัฒนาข้อสอบรายข้อ ประกอบด้วย

**(1.1)** ค่าความยาก **(Estimate: b)** ซึ่งโดยทั่วไปค่าที่ยอมรับได้อยู่ในช่วงความยาก **-3** ถึง **+3**

**(1.2)** โค้งคุณลักษณะของข้อสอบแต่ละข้อ **(ICC)** เพื่อบ่งชี้ว่าข้อสอบแต่ละข้อเหมาะสมกับผู้สอบที่มีความสามารถอยู่ในระดับใด **(Modelled and Empirical Item Characteristic Curves** ตามคู่มือในหนังสือ **ConQuest 2.0** หน้า **27)**

**(1.3)** อำนาจจำแนกตามทฤษฎีการทดสอบแบบดั้งเดิม **(CTT)** โดยพิจารณาจากค่า **Point Biserial (r)** โดยปกติควรมีค่าตั้งแต่ **0.2** ขึ้นไป และมีนัยสำคัญ

**(1.4)** ประสิทธิภาพตัวลวงในกรณีที่เป็นข้อสอบแบบปรนัยตามทฤษฎีการทดสอบแบบดั้งเดิม **(CTT)** โดยพิจารณาจากจำนวนของผู้สอบในการเลือกตัวลวงแต่ละค่า และพิจารณาจากโค้ง **ICC** ของตัวลวง ซึ่งปรากฏอยู่ในไฟล์ **.itn** เพื่อพิจารณาว่าตัวลวงแต่ละตัวเหมาะสำหรับผู้สอบที่ความสามารถในระดับใด โดยพิจารณาจากกราฟ **Modelled and Empirical Category Characteristics Curves (MCC)** ประกอบ **(**สามารถดูตัวอย่างเพิ่มเติมในหน้าที่ **27** ตามคู่มือ **ConQuest 2.0)**

**(1.5)** ค่าสารสนเทศของข้อสอบรายข้อ

**(1.6)** ค่าความคลาดเคลื่อนในการประมาณค่า โดยพิจารณาจากค่า **MNSQ** ใน **Weight Fit** โดยมีค่าอยู่ในช่วง **CI** โดยควรมีค่าความเชื่อมั่นที่ยอมรับในระดับร้อยละ **95** ของแต่ละข้อ **(**ค่า **MNSQ** อยู่ในช่วง **CI)**

**(1.7)** พิจารณาค่าขนาดอิทธิพล **(Effect Size)** โดยพิจารณาจากค่า **MNSQ** ควรอยู่ในช่วง **0.75-1.33** ซึ่งเป็นช่วงที่ยอมรับได้

**(1.8)** พิจารณาค่า **T** ซึ่งเป็นค่าสะท้อนถึงความคลาดเคลื่อนที่ยอมรับได้โดยเปรียบเทียบกับโค้ง **ICC** ที่คาดหวังกับ โค้ง **ICC** ที่เก็บข้อมูลได้จริงจากข้อมูลเชิงประจักษ์ โดยค่าที่ยอมรับได้อยู่ในช่วง **-2** ถึง **+2**

**(2)** คุณภาพของแบบทดสอบทั้งฉบับ ประกอบด้วย

(2.1) ค่าความเที่ยงแบบสอดคล้องภายใน **(**สัมประสิทธิ์แอลฟาของครอนบาค**; $\alpha$)** ตามทฤษฎีการทดสอบแบบดั้งเดิม ควรมีค่าตั้งแต่ **0.7** ขึ้นไป

**(2.2)** ค่าฟังก์ชันสารสนเทศของแบบทดสอบ **(TCC)** เป็นค่าที่สะท้อนถึงความคงเส้นคงวาของการประมาณค่าของแบบทดสอบทั้งสอบทั้งฉบับ ว่าแบบทดสอบทั้งฉบับเหมาะสำหรับผู้สอบที่มีความสามารถที่แท้จริง ($\theta$) ที่ระดับใดจึงจะเหมาะสม ซึ่งตำแหน่งดังกล่าวจะเป็นตำแหน่งที่มีความคลาดเคลื่อนในการประมาณค่า (SE) ต่ำที่สุด โดยค่าฟังก์ชันสารสนเทศของแบบทดสอบจะสะท้อนถึงค่าความเที่ยง (Reliability) นั่นเอง

(2.3) ค่าความตรงเชิงโครงสร้าง โดยพิจารณาจาก Wright Map สะท้อนถึงความตรงเชิงโครงสร้างภายในของข้อสอบแต่ละข้อที่สัมพันธ์กับผลการตอบของนักเรียนและแผนที่ความก้าวหน้าที่พัฒนาขึ้นสอดคล้องกับข้อมูลเชิงประจักษ์หรือไม่ โดยแสดงให้เห็นถึงความน่าจะเป็นในการตอบข้อสอบจากระดับขั้นหนึ่งไปยังระดับขั้นที่สูงกว่า (ระดับ k-1 ไปยังระดับ k) ของข้อสอบข้อที่ **i** ซึ่งในที่นี้ เรียกว่า “Item Thresholds” ในกรณีที่มีการตรวจให้คะแนนมากกว่า 1 ค่า เช่น ข้อสอบอัตนัย ซึ่งในการวิจัยครั้งนี้เลือกใช้

โมเดลที่เรียกว่า “Partial Credit Model” ซึ่งระดับขั้นของการตอบ (Step) มีค่าเท่ากับ k-1 เมื่อ k คือระดับ (Category) ของแผนภาพความก้าวหน้าในการเรียนรู้ (Progress Map)

7.2 ประเมินการความก้าวหน้าในการเรียนรู้ (Learning Progression) การประเมินทักษะการรู้สารสนเทศ สำหรับนักเรียนระดับมัธยมศึกษา โดยการวิเคราะห์ความสามารถของนักเรียนแต่ละคน โดยนำเสนอทั้งการวิเคราะห์ข้อมูลเชิงคุณภาพจากการวิเคราะห์ผลการตอบเพื่อเรียน โดยการถอดโปรคอลตีความร่วมกับข้อมูลเชิงปริมาณที่ได้จากการพิจารณาค่าคะแนนดิบ และความสามารถที่แท้จริงของนักเรียน ($\boldsymbol{\theta}$) พร้อมทั้งการวินิจฉัยผู้เรียนเป็นรายบุคคลเพื่อนำไปสู่การปรับปรุงผู้เรียนในอนาคตโดยเทียบกับจุดตัดตามแผนที่ความก้าวหน้าภายในหลังการพัฒนากรอบการประเมินความก้าวหน้าแล้ว ซึ่งเป็นผลการนำเสนอเป็นทั้งรายบุคคลและภาพรวม

## 8. ผลการวิจัย (Results)

การนำเสนอผลการวิจัยในครั้งนี้ได้แบ่งการนำเสนอ แยกเป็น 2 ส่วนตามวัตถุประสงค์การวิจัย ส่วนแรกจะกล่าวถึงคุณภาพรายข้อและทั้งฉบับ พร้อมทั้งแนวทางปรับปรุงแก้ไขในแต่ละข้อ และส่วนที่สองเป็นการนำเสนอผลการประเมินผู้เรียนเป็นรายบุคคลและภาพรวมตามแผนภาพความก้าวหน้าภายหลังการปรับให้สมบูรณ์จากส่วนที่แรก โดยมีรายละเอียด ดังนี้

**วัตถุประสงค์ข้อที่** 1 เพื่อพัฒนาและตรวจสอบคุณภาพกรอบการประเมินการความก้าวหน้าในการเรียนรู้ (Learning Progression) ทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษาของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย โรงเรียนประชารัฐวิทยาเสริม เรื่องระบบเครือข่าย กลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี

### 8.1 คุณภาพของข้อสอบรายข้อ

#### 8.1.1 ค่าความยาก อำนาจจำแนก และการพิจารณาความเหมาะสมของข้อสอบ

**ตารางที่ 3** ค่าความยาก อำนาจจำแนก และการพิจารณาความเหมาะสมของข้อสอบ

| ข้อสอบ (Item) | **ค่าความยาก (Estimate)** | ค่าอำนาจจำแนกตามทฤษฎี CTT ($r_{pb}$) | Weight Fit (ปรากฏค่าในไฟล์ .shw) | | | แปลผล (อธิบายว่าข้อสอบ ยาก ปานกลาง หรือง่าย พร้อมพิจารณาว่าค่า MNSQ อยู่ในช่วงที่ยอมรับได้หรือไม่) | พิจารณาโดยรวม (ใส่เครื่องหมาย ✓) | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | MNSQ | CI | T | | ดี | แก้ไข/ปรับปรุง | ตัดทิ้ง |
| 1 | 0.119 | 0.67 | 1.02 | 0.56, 1.44 | 0.1 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | ✓ | | |
| 2 | 0.004 | 0.82 | 0.84 | 0.55, 1.45 | -0.7 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | ✓ | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 3 | 0.526 | 0.79 | 0.91 | 0.58, 1.42 | -0.4 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางยาก เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | ✓ | | |
| 4 | 0.306 | 0.67 | 0.94 | 0.58, 1.42 | -0.3 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | ✓ | | |
| 5 | -1.218 | 0.69 | 0.95 | 0.58, 1.42 | -0.2 | ข้อสอบมีค่าความยากระดับง่าย เหมาะสำหรับนักเรียนที่มีความสามารถระดับต่ำ ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | ✓ | | |
| 6 | 0.062 | 0.69 | 1.20 | 0.56, 1.44 | 0.9 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | ✓ | | |
| 7 | -0.087 | 0.70 | 0.91 | 0.55, 1.45 | -0.4 | ข้อสอบมีค่าความยากระดับง่าย เหมาะสำหรับนักเรียนที่มีความสามารถระดับต่ำ ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | ✓ | | |
| 8 | 0.048 | 0.80 | 0.92 | 0.53, 1.47 | -0.3 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | ✓ | | |
| 9 | 0.240 | 0.69 | 1.37 | 0.54, 1.46 | 1.5 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | ✓ | | |

จากตารางที่ *3* พบว่าข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ มีค่าความยาก (Estimate) อยู่ในช่วง (-1,+1) จำนวน 9 ข้อ รวมทั้ง MNSQ อยู่ในช่วง CI ที่ยอมรับได้ทั้งหมดจำนวน 9 ข้อ

### 8.1.1 ค่า Threshold ในแต่ละระดับขั้นของการตอบของแต่ละข้อใน Wright Map

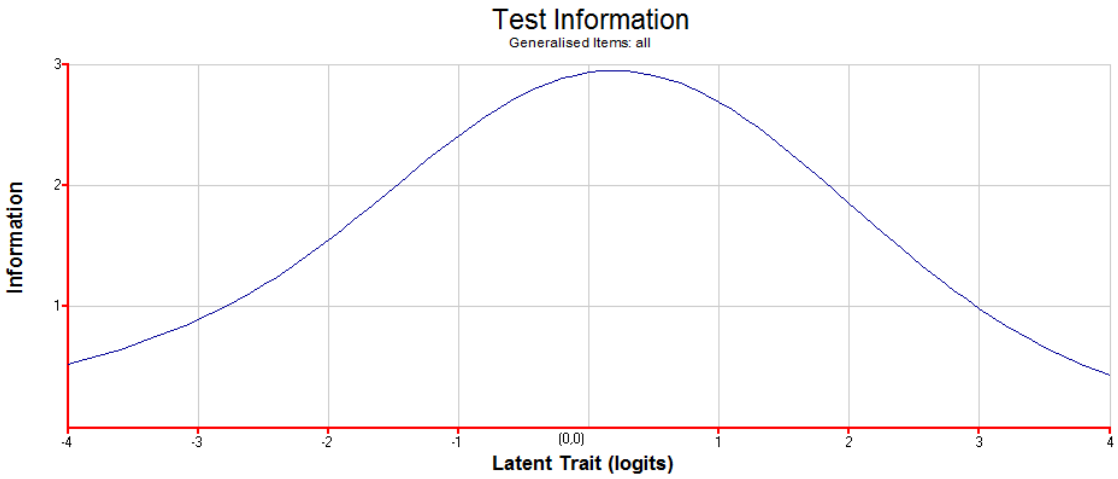
ตารางที่ 4 ค่า Threshold ในแต่ละระดับขั้นของการตอบของแต่ละข้อใน Wright Map

| ข้อสอบ (Item) | ค่าที่แปลงแล้ว (Recode) | ค่าคะแนนจริง | ระดับขั้นของการตอบ (Step) | Threshold (Estimate) | Weight Fit (ปรากฏค่าในไฟล์ .shw) | | | การพิจารณาโดยภาพรวม | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | MNSQ | CI | T | ดี | แก้ไข | ตัดทิ้ง |
| 1 | 0,1 | 0,1 | 0 | - | - | - | - | - | - | - |
| | | | 1 | 0.119 | 1.02 | 0.56, 1.44 | 0.1 | ✓ | | |
| 2 | 0,1,2 | 0,1,2 | 0 | - | 0.93 | 0.37, 1.63 | -0.1 | ✓ | | |
| | | | 1 | -1.135 | 0.94 | 0.73, 1.27 | -0.4 | ✓ | | |
| | | | 2 | 1.135 | 0.86 | 0.60, 1.40 | -0.7 | ✓ | | |
| 3 | 0,1,2 | 1,2,3 | 1 | - | 1.05 | 0.41, 1.59 | 0.2 | ✓ | | |
| | | | 2 | -1.516 | 0.96 | 0.75, 1.25 | -0.3 | ✓ | | |
| | | | 3 | 1.516 | 0.86 | 0.59, 1.41 | -0.7 | ✓ | | |
| 4 | 0,1 | 0,1 | 0 | - | - | - | - | - | - | - |
| | | | 1 | 0.306 | 0.94 | 0.58, 1.42 | -0.3 | ✓ | | |
| 5 | 0,1,2 | 0,1,2 | 0 | - | 0.98 | 0.00, 2.48 | 0.2 | ✓ | | |
| | | | 1 | -2.478 | 0.99 | 0.65, 1.35 | 0.0 | ✓ | | |
| | | | 2 | 2.478 | 0.97 | 0.61, 1.39 | -0.1 | ✓ | | |
| 6 | 0,1,2 | 1,2,3 | 1 | - | 0.88 | 0.36, 1.64 | -0.3 | ✓ | | |
| | | | 2 | -1.250 | 1.03 | 0.74, 1.26 | 0.2 | ✓ | | |
| | | | 3 | 1.250 | 1.26 | 0.60, 1.40 | 1.2 | ✓ | | |
| 7 | 0,1 | 0,1 | 0 | - | - | - | - | - | - | - |
| | | | 1 | -0.087 | 0.91 | 0.55, 1.45 | -0.4 | ✓ | | |
| 8 | 0,1,2 | 0,1,2 | 0 | - | 0.91 | 0.40, 1.60 | -0.3 | ✓ | | |
| | | | 1 | -0.850 | 1.10 | 0.71, 1.29 | 0.7 | ✓ | | |
| | | | 2 | 0.850 | 1.05 | 0.58, 1.42 | 0.3 | ✓ | | |
| 9 | 0,1,2 | 1,2,3 | 1 | - | 1.24 | 0.43, 1.57 | 0.8 | ✓ | | |
| | | | 2 | -0.803 | 1.05 | 0.71, 1.29 | 0.4 | ✓ | | |
| | | | 3 | 0.803 | 1.25 | 0.58, 1.42 | 1.2 | ✓ | | |

จากตารางที่ 4 พบว่าข้อสอบทั้งหมดอยู่ในระดับค่าความยากที่ยอมรับได้ และถือว่าใช้ได้ในทุกข้อ สามารถนำไปใช้งานได้จริงทั้งหมดจำนวน 9 ข้อ เนื่องจากทุกข้ออยู่ในระดับค่าความยากที่ยอมรับได้ และค่า MNSQ อยู่ในช่วง Cl ทุกข้อ ข้อสอบอยู่ในระดับขั้นตาม progress map ซึ่งถือว่าสอดคล้อง

## 8.2 คุณภาพของแบบทดสอบทั้งฉบับ

8.2.1 นำเสนอกราฟ Test Information พร้อมนำเสนอ ค่าความเที่ยงแบบสัมประสิทธิ์แอลฟาของครอนบาค และความเที่ยงแบบ EAP ประกอบ ใต้แผนภาพ พร้อมอธิบายแผนภาพประกอบค่าเหล่านั้น




ภาพที่ 5 กราฟ Test Information

EAP/PV RELIABILITY: 0.864

Coefficient Alpha: 0.88

8.2.2 นำเสนอระดับขั้นของการตอบ ที่สัมพันธ์กับการกำหนดจุดตัดใน Wright Map เพื่อตรวจสอบความตรงเชิงโครงสร้าง

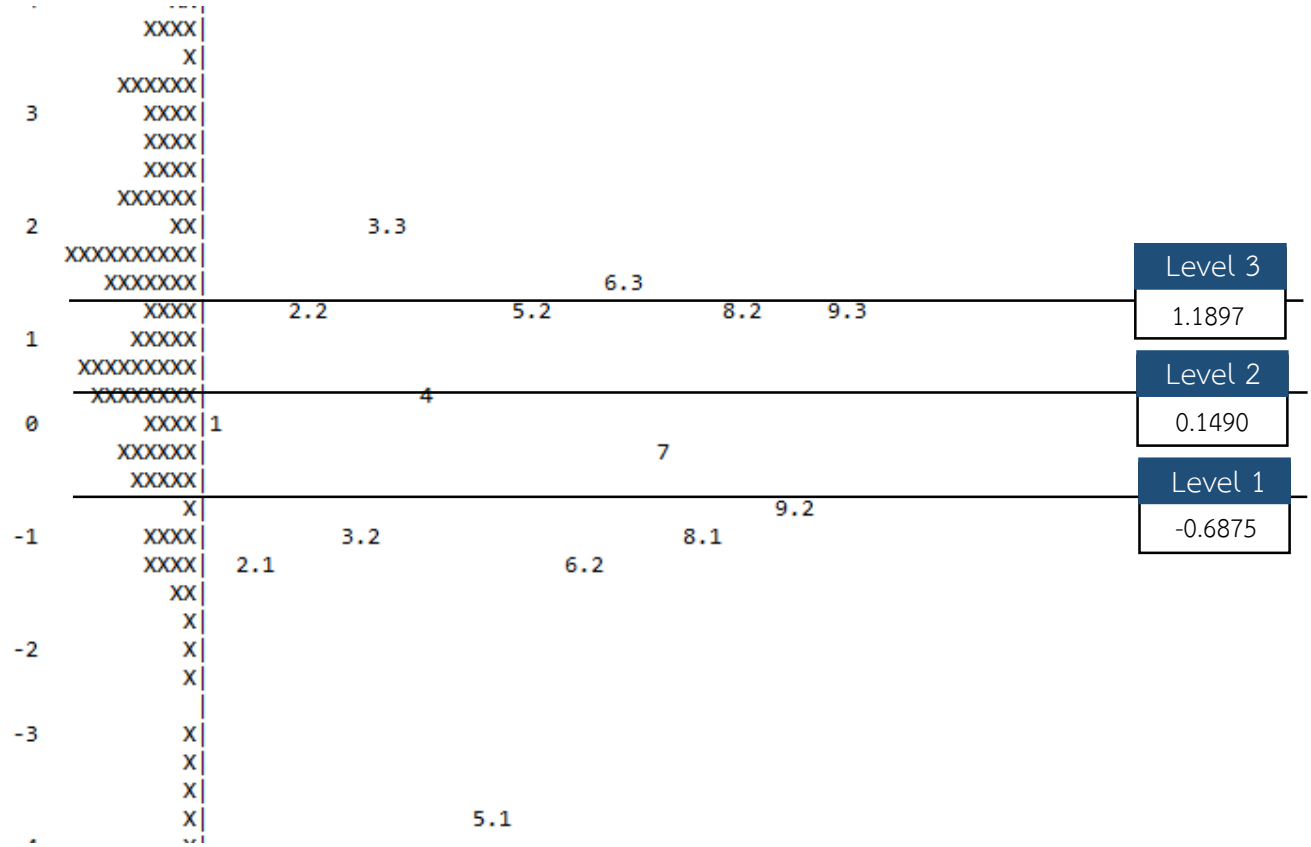
**ตารางที่ 5** ระดับขั้นของการตอบ ที่สัมพันธ์กับการกำหนดจุดตัดใน Wright Map เพื่อตรวจสอบความตรงเชิงโครงสร้าง

| ระดับขั้นของการตอบที่แท้จริง (Step) | ข้อสอบที่อยู่ในระดับขั้นของการตอบที่แท้จริงเดียวกัน | ค่า **Threshold** (Estimate) ในแต่ละข้อที่อยู่ใน **step** เดียวกัน | จำนวนข้อ | ค่าเฉลี่ยของ **Threshold**-จุดตัด (ผลรวมของ **Threshold**/ จำนวนข้อ) |
|---|---|---|---|---|
| Step 1 | 1, 2, 4, 5, 7, 8 | 0.119, -1.135, 0.306, -2.478, -0.087, -0.850 | 6 | -0.6875 |
| Step 2 | 2, 3, 5, 6, 8, 9 | 1.135, -1.516, 2.478, -1.250, 0.850, -0.803 | 6 | 0.149 |
| Step 3 | 3, 6, 9 | 1.516, 1.250, 0.803 | 3 | 1.1897 |

จากตารางที่ 5 พบว่าขั้นที่ 1 มีข้อสอบที่อยู่ระดับนี้จำนวน 6 ข้อ ค่าเฉลี่ยของ Threshold-จุดตัดอยู่ที่ -0.6875 ขั้นที่ 2 มีข้อสอบที่อยู่ระดับนี้จำนวน 6 ข้อ ค่าเฉลี่ยของ Threshold-จุดตัดอยู่ที่ 0.149 และขั้นที่ 3 มีข้อสอบที่อยู่ระดับนี้จำนวน 3 ข้อ ค่าเฉลี่ยของ Threshold-จุดตัดอยู่ที่ 1.1897

8.3.3 นำเสนอ Wright Map

การพิจารณา Wright Map




**แผนภาพที่** 6 Wright Map ระดับ Thresholds ของข้อสอบแต่ละข้อ
เพื่อตรวจสอบความตรงเชิงโครงสร้างตามโมเดล PCM

จากแผนภาพที่..... พบว่า ระดับ 1 มีจุดตัดอยู่ที่ต่ำกว่า -0.6875 ลงมาซึ่งได้แก่ข้อ 2.1, 3.2, 5.1, 6.2, 8.1 และ 9.2 ระดับ 2 มีจุดตัดอยู่ช่วงระหว่าง -0.6876 ถึง 0.148 ซึ่งได้แก่ข้อ 1, 4 และ 7 ระดับ 3 มีจุดตัดอยู่ในช่วง 0.149 ถึง 1.1896 ซึ่งได้แก่ข้อ 2.2, 5.2, 8.2 และ 9.3 และระดับ 4 จุดตัด 1.1898 ขึ้นไป ซึ่งได้แก่ข้อ 3.3 และ 6.3 ในส่วนของความยากนั้นข้อที่อยู่ระดับง่าย ได้แก่ 2.1, 3.2, 5.1, 6.2, 8.1 ข้อที่อยู่ระดับปานกลาง ได้แก่ 1, 4, 7, 2.2, 5.2, 8.2 และ 9.3 และข้อที่อยู่ระดับยาก ได้แก่ 3.3 และ 6.3

8.2.4 การปรับปรุงแก้ไขกรอบการประเมินความก้าวหน้าให้สมบูรณ์ก่อนนำไปประเมินผู้เรียนในชั้นเรียนจริง

**ตารางที่** 6 การปรับปรุงแก้ไขกรอบการประเมินความก้าวหน้าให้สมบูรณ์ก่อนนำไปประเมินผู้เรียนในชั้นเรียนจริง

| ระดับขั้นของการตอบที่แท้จริง **(Step)** | ข้อที่อยู่ใน **step** การตอบภายหลังการกำหนดคะแนนจุดตัด | ระดับใน **Progress Map** | ช่วงระดับความยาก **(ยาก,ปานกลาง,ง่าย)** | คำอธิบาย | การตีความแนวทางการพัฒนา **Progress Map** |
|---|---|---|---|---|---|
| 1 | -0.6875 | 1 | ง่าย | | |
| 2 | 0.149 | 2 | ปานกลาง | | |
| 3 | 1.1897 | 3 | ยาก | | |

**ตารางที่** 7 กรอบการประเมินความก้าวหน้าในการเรียนรู้ทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษาของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย

| **ระดับ** | **คะแนน** | **ชื่อระดับ** | **คำอธิบายคุณลักษณะ (What Students Know)** | **สิ่งที่นักเรียนมีความจำเป็นต้องปรับปรุงพัฒนา (What Students Need to Know) เพื่อให้มีความสามารถที่สูงขึ้น** |
|---|---|---|---|---|
| **0** | **0** | ไม่มีการตอบ Non Response | **<u>ไม่มีการตอบหรือไม่ตรงประเด็น</u>**<br>ไม่ตอบคำถาม ไม่ส่งข้อสอบ หรือไม่ได้เขียนคำตอบที่เกี่ยวกับคำถาม | **<u>สามารถจำเนื้อหาการจำนิยาม รูปแบบ ข้อความจริง หรือการแสดงกระบวนการง่ายๆ</u>**<br>- นักเรียนสามารถจำเนื้อหา หรือหลักการนั้นๆมาตอบคำถามได้ เช่น บอกความหมาย บอกขั้นตอน บอกวิธีการ แต่ไม่เข้าใจในหลักการแต่ละขั้นตอน<br>- ไม่สามารถระบุหรือทราบถึงว่าสามารถเอาสิ่งที่จำได้นั้นไปใช้ทำอะไรเพิ่มเติมได้<br>- ใช้เพียงแค่ทำที่พบเห็นในเนื้อหาหรือเอกสารประกอบการสอนเท่านั้นมาตอบคำถาม<br>- ไม่สามารถบอกเหตุผล หรือวิธีการที่เลือกเอากระบวนการนั้นๆมาใช้หรือมาตอบ |
| **1** | **1** | **ความรู้ความจำ Recall** | **<u>สามารถจำเนื้อหาการจำนิยาม รูปแบบ ข้อความจริง หรือการแสดง</u>** | **<u>การใช้กระบวนการ นักเรียนตัดสินใจ ด้วยตนเองว่าจะแก้ปัญหา</u>** |

| ระดับ | คะแนน | ชื่อระดับ | คำอธิบายคุณลักษณะ (What Students Know) | สิ่งที่นักเรียนมีความจำเป็นต้องปรับปรุงพัฒนา (What Students Need to Know) เพื่อให้มีความสามารถที่สูงขึ้น |
|---|---|---|---|---|
| | | | **<u>กระบวนการง่ายๆ</u>**<br>- นักเรียนสามารถจำเนื้อหา หรือหลักการนั้นๆ มาตอบคำถามได้ เช่น บอกความหมาย บอกขั้นตอน บอกวิธีการ แต่ไม่เข้าใจในหลักการแต่ละขั้นตอน<br>- ไม่สามารถระบุหรือทราบถึงว่า สามารถเอาสิ่งที่จำได้นั้นไปใช้ทำอะไรเพิ่มเติมได้<br>- ใช้เพียงแค่ทำที่พบเห็นในเนื้อหา หรือเอกสารประกอบการสอนเท่านั้นมาตอบคำถาม<br>- ไม่สามารถบอกเหตุผลหรือวิธีการที่เลือกเอากระบวนการนั้นๆมาใช้หรือมาตอบ | **<u>อย่างไร อ่านข้อมูลและต้องทราบว่าจะนำข้อมูลที่ได้มาสรุปอย่างไร</u>**<br>- ระดับการใช้กระบวนการที่เกิดขึ้นภายในสมอง ในระดับนี้มิความต้องการให้นักเรียนตัดสินใจ ด้วยตนเองว่าจะแก้ปัญหาอย่างไร<br>- นักเรียนสามาถเข้าใจในกระบวนหลักการ วิธีการ รวมถึงทุกขั้นตอนของวิธีที่กล่าวมา เพื่อนำมาตัดสินใจวางแผน ในการแก้ปัญหา<br>- นักเรียนบอกการตัดสินใจในการเลือกใช้รูปแบบ ประบวนการ วิธีการนั้นๆ<br>- นักเรียนสามารถสรุปวิธีการ จาหเหตุผลหลายๆปัจจัยที่มีส่วนหรือผลต่อคำตอบ<br>- นักเรียนสามารถสรุปคำตอบออกมาเป็นภาษา รูปแบบ มโนทัศน์ หรือรูปแบบอื่นๆที่เข้าใจ เป็นของตัวเอง |
| 2 | 2 | ทักษะหรือความคิดรวบยอด Skill/Concept | **<u>การใช้กระบวนการนักเรียนตัดสินใจ ด้วยตนเองว่าจะแก้ปัญหาอย่างไร อ่านข้อมูลและต้องทราบว่าจะนำข้อมูลที่ได้มาสรุปอย่างไร</u>**<br>- ระดับการใช้กระบวนการที่เกิดขึ้นภายในสมอง ในระดับนี้มิ | **<u>ให้เหตุผล การวางแผน การใช้หลักฐานเพื่อตัดสินใจ อย่างสมเหตุสมผล อธิบายข้อเท็จจริงในรูปความคิด รวบยอด</u>**<br>- การสร้างข้อคาดเดา อธิบายข้อเท็จจริงในรูปความคิดรวบยอด ใช้ความคิดรวบยอดเพื่อแก้ปัญหา<br>- ระดับที่ต้องการให้นักเรียนให้เหตุผล การวางแผน การใช้หลักฐานเพื่อตัดสินใจ |

| ระดับ | คะแนน | ชื่อระดับ | คำอธิบายคุณลักษณะ (What Students Know) | สิ่งที่นักเรียนมีความจำเป็นต้องปรับปรุงพัฒนา (What Students Need to Know) เพื่อให้มีความสามารถที่สูงขึ้น |
|---|---|---|---|---|
| | | | ความต้องการให้นักเรียนตัดสินใจ ด้วยตนเองว่าจะแก้ปัญหาอย่างไร<br>- นักเรียนสามาถเข้าใจในกระบวน หลักการ วิธีการ รวมถึงทุกขั้นตอนของวิธีที่กล่าวมา เพื่อนำมาตัดสินใจ วางแผนในการแก้ปัญหา<br>- นักเรียนบอกการตัดสินใจในการเลือกใช้รูปแบบ ประบวนการวิธีการนั้นๆ<br>- นักเรียนสามารถสรุปวิธีการ จาหเหตุผลหลายๆปัจจัยที่มีส่วนหรือผลต่อคำตอบ<br>- นักเรียนสามารถสรุปคำตอบออกมาเป็นภาษารูปแบบ มโนทัศน์ หรือรูปแบบอื่นๆที่เข้าใจ เป็นของตัวเอง | - การแก้ปัญหามความซับซ้อนและเป็นนามธรรม กิจกรรมที่มีคำตอบ ที่เป็นไปได้มากกว่า 1 คำตอบ<br>- นักเรียน พิสูจน์ กิจกรรมลักษณะอื่นๆ ที่อยู่ในระดับ 3 คือ นักเรียนหาข้อสรุปจากการสังเกตหลักฐานและอ้างเหตุผล อย่างสมเหตุสมผล อธิบายข้อเท็จจริงในรูปความคิด รวบยอด และใช้ความคิดรวบยอดเพื่อแล้วปัญหา<br>- นักเรียนวิเคราะห์วิธีการ หลักการอย่างเข้าใจและมีเหตุผล เพื่อใช้ประกอบการตัดสินใจในการเลือกวิธีที่เหมาะสม สำหรับแก้ปัญหา |
| **3** | **3** | ยุทธวิธีในการคิด Strategic Thinking | **<u>ให้เหตุผล การวางแผนการใช้หลักฐานเพื่อตัดสินใจ อย่างสมเหตุสมผล อธิบายข้อเท็จจริงในรูปความคิด รวบยอด</u>**<br><br>- การสร้างข้อคาดเดา อธิบายข้อเท็จจริงในรูปความคิดรวบยอด ใช้ความคิดรวบยอดเพื่อแก้ปัญหา | **<u>การวางแผน การพัฒนา การคิดขั้นสูงและงานต้อง ซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง</u>**<br><br>- ใช้กระบวนการ วิธีการ หรือหลักการในการแก้ปัญหา ที่มากกว่า 1 วิธีในปัญหานั้นๆ<br>- วิเคราะห์จุดเด่น จุดดีของหลักการ วิธีการ เพื่อเลือกเอากระบวนการนั้นๆ มาปรับประยุกต์ให้เข้ากับวิธีการแก้ปัญหาของตัวเอง<br>- มีกระบวนการสร้างขั้นตอน วิธีการ |

| ระดับ | คะแนน | ชื่อระดับ | คำอธิบายคุณลักษณะ (What Students Know) | สิ่งที่นักเรียนมีความจำเป็นต้องปรับปรุงพัฒนา (What Students Need to Know) เพื่อให้มีความสามารถที่สูงขึ้น |
|---|---|---|---|---|
| | | | - ระดับที่ต้องการให้นักเรียนให้เหตุผล การวางแผน การใช้หลักฐานเพื่อตัดสินใจ<br>- การแก้ปัญหามความซับซ้อนและเป็นนามธรรม กิจกรรมที่มีคำตอบ ที่เป็นไปได้มากกว่า 1 คำตอบ<br>- นักเรียน พิสูจน์กิจกรรมลักษณะอื่นๆ ที่อยู่ในระดับ 3 คือ นักเรียนหาข้อสรุปจากการสังเกตหลักฐานและอ้างเหตุผล อย่างสมเหตุสมผล อธิบายข้อเท็จจริงในรูปความคิดรวบยอด และใช้ความคิดรวบยอดเพื่อแล้วปัญหา<br>- นักเรียนวิเคราะห์วิธีการ หลักการ อย่างเข้าใจและมีเหตุผล เพื่อใช้ประกอบการตัดสินใจในการเลือกวิธีที่เหมาะสม สำหรับแก้ปัญหา | แก้ปัญหา และปรับปรุงเพิ่มเติมการแก้ปัญหา<br>- การวางแผนเพื่อการพัฒนา การออกแบบ การสร้างการเชื่อมโยงระหว่าง ความคิดรวบยอดและข้อเท็จจริง รวบรวมและ สังเคราะห์แนวคิดไปสู่ความคิดรวบยอดใหม่ |

**วัตถุประสงค์ข้อที่** 2 ประเมินการความก้าวหน้าในการเรียนรู้ (Learning Progression) ทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษาของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย โรงเรียนประชารัฐวิทยาเสริม เรื่องระบบเครือข่าย กลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี

**ตารางที่ 8** คะแนนดิบ คะแนนความสามารถที่แท้จริง เมื่อเทียบกับแผนที่ความก้าวหน้าในการเรียนรู้ พร้อมสิ่งที่นักเรียนแต่ละคนต้องการปรับปรุงพัฒนา

| เลขที่ | คะแนนดิบ (คะแนนเต็ม = 18 คะแนน) | คะแนนความสามารถที่แท้จริง (ปรากฏในไฟล์ EAP) | ระดับตาม Progress Map | คำอธิบายคุณลักษณะ (What Students Know) | สิ่งที่นักเรียนมีความจำเป็นต้องปรับปรุงพัฒนา (What Students Need to Know) เพื่อให้มีความสามารถที่สูงขึ้น |
|---|---|---|---|---|---|
| 1 | 18 | 4.06589 | 4 | การวางแผน การพัฒนา การคิดขั้นสูงและงานต้องซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง | - |
| 2 | 15 | 1.80800 | 4 | การวางแผน การพัฒนา การคิดขั้นสูงและงานต้องซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง | - |
| 3 | 13 | 0.99582 | 3 | ให้เหตุผล การวางแผน การใช้หลักฐานเพื่อตัดสินใจ อย่างสมเหตุสมผล อธิบายข้อเท็จจริงในรูปความคิด รวบยอด | การวางแผน การพัฒนา การคิดขั้นสูงและงานต้อง ซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง |
| 4 | 16 | 2.31075 | 4 | การวางแผน การพัฒนา การคิดขั้นสูงและงานต้องซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง | - |

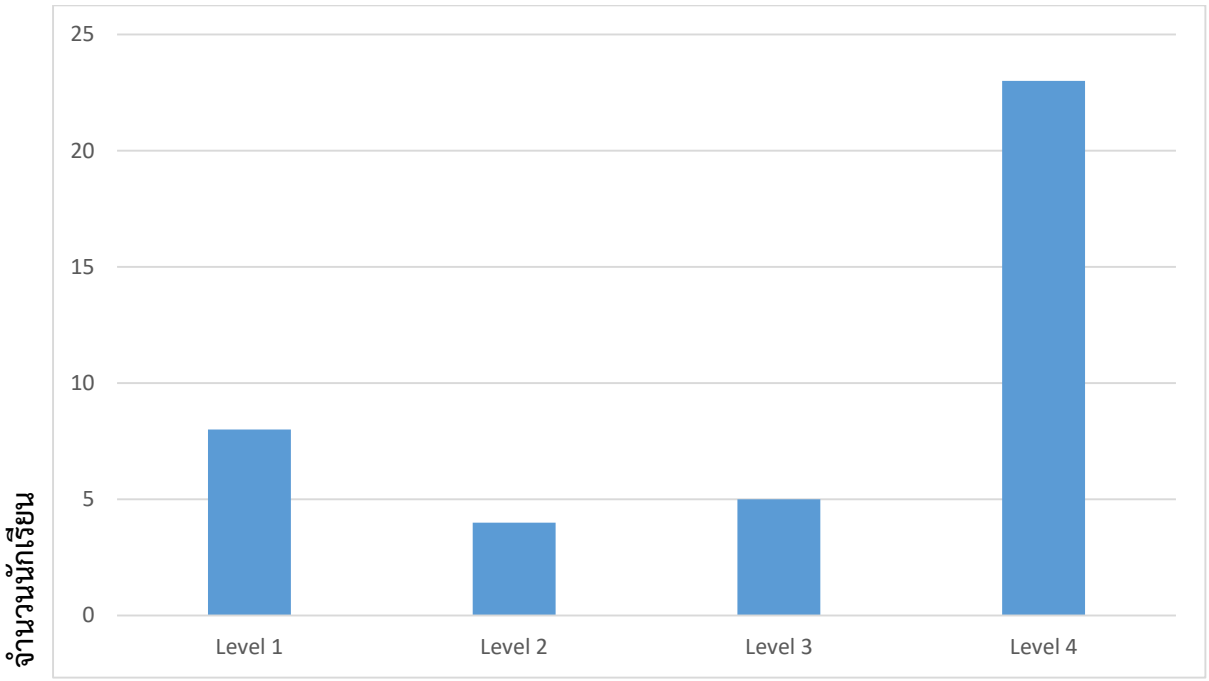
| เลขที่ | คะแนนดิบ (คะแนนเต็ม = 18 คะแนน) | คะแนนความสามารถที่แท้จริง (ปรากฏในไฟล์ EAP) | ระดับตาม Progress Map | คำอธิบายคุณลักษณะ (What Students Know) | สิ่งที่นักเรียนมีความจำเป็นต้องปรับปรุงพัฒนา (What Students Need to Know) เพื่อให้มีความสามารถที่สูงขึ้น |
|---|---|---|---|---|---|
| 5 | 7 | -1.08467 | 1 | สามารถจำเนื้อหาการจำนิยามรูปแบบ ข้อความจริง หรือการแสดงกระบวนการง่ายๆ | การใช้กระบวนการนักเรียนตัดสินใจ ด้วยตนเองว่าจะแก้ปัญหาอย่างไร อ่านข้อมูลและต้องทราบว่าจะนำข้อมูลที่ได้มาสรุปอย่างไร |
| 6 | 9 | -0.30951 | 2 | การใช้กระบวนการนักเรียนตัดสินใจด้วยตนเองว่าจะแก้ปัญหาอย่างไร อ่านข้อมูลและต้องทราบว่าจะนำข้อมูลที่ได้มาสรุปอย่างไร | ให้เหตุผล การวางแผนการใช้หลักฐานเพื่อตัดสินใจ อย่างสมเหตุสมผล อธิบายข้อเท็จจริงในรูปความคิดรวบยอด |
| 7 | 12 | 0.64754 | 3 | ให้เหตุผล การวางแผน การใช้หลักฐานเพื่อตัดสินใจ อย่างสมเหตุสมผล อธิบายข้อเท็จจริงในรูปความคิด รวบยอด | การวางแผน การพัฒนาการคิดขั้นสูงและงานต้อง ซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง |
| 8 | 11 | 0.33194 | 3 | ให้เหตุผล การวางแผน การใช้หลักฐานเพื่อตัดสินใจ อย่างสมเหตุสมผล อธิบายข้อเท็จจริงในรูปความคิด รวบยอด | การวางแผน การพัฒนาการคิดขั้นสูงและงานต้อง ซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง |

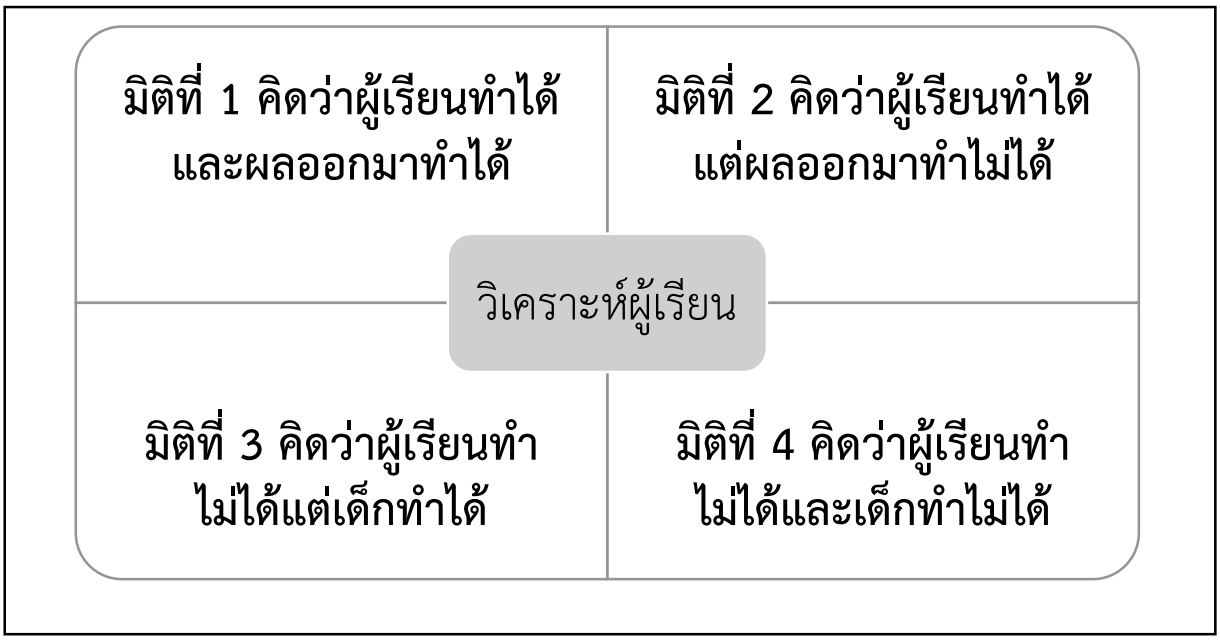
| เลขที่ | คะแนนดิบ (คะแนนเต็ม = 18 คะแนน) | คะแนนความสามารถที่แท้จริง (ปรากฏในไฟล์ EAP) | ระดับตาม Progress Map | คำอธิบายคุณลักษณะ (What Students Know) | สิ่งที่นักเรียนมีความจำเป็นต้องปรับปรุงพัฒนา (What Students Need to Know) เพื่อให้มีความสามารถที่สูงขึ้น |
|---|---|---|---|---|---|
| 9 | 17 | 2.99121 | 4 | การวางแผน การพัฒนา การคิดขั้นสูงและงานต้องซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง | - |
| 10 | 16 | 2.31075 | 4 | การวางแผน การพัฒนา การคิดขั้นสูงและงานต้องซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง | - |
| 11 | 10 | 0.02354 | 1 | สามารถจำเนื้อหาการจำนิยามรูปแบบ ข้อความจริง หรือการแสดงกระบวนการง่ายๆ | การใช้กระบวนการนักเรียนตัดสินใจ ด้วยตนเองว่าจะแก้ปัญหาอย่างไร อ่านข้อมูลและต้องทราบว่าจะนำข้อมูลที่ได้มาสรุปอย่างไร |
| 12 | 8 | -0.68229 | 1 | สามารถจำเนื้อหาการจำนิยามรูปแบบ ข้อความจริง หรือการแสดงกระบวนการง่ายๆ | การใช้กระบวนการนักเรียนตัดสินใจ ด้วยตนเองว่าจะแก้ปัญหาอย่างไร อ่านข้อมูลและต้องทราบว่าจะนำข้อมูลที่ได้มาสรุปอย่างไร |
| 13 | 17 | 2.99121 | 4 | การวางแผน การพัฒนา การคิดขั้นสูงและงานต้องซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง | - |

| เลขที่ | คะแนนดิบ (คะแนนเต็ม = 18 คะแนน) | คะแนนความสามารถที่แท้จริง (ปรากฏในไฟล์ EAP) | ระดับตาม Progress Map | คำอธิบายคุณลักษณะ (What Students Know) | สิ่งที่นักเรียนมีความจำเป็นต้องปรับปรุงพัฒนา (What Students Need to Know) เพื่อให้มีความสามารถที่สูงขึ้น |
|---|---|---|---|---|---|
| 14 | 14 | 1.38234 | 4 | การวางแผน การพัฒนา การคิดขั้นสูงและงานต้องซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง | - |
| 15 | 16 | 2.31075 | 4 | การวางแผน การพัฒนา การคิดขั้นสูงและงานต้องซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง | |

จากตารางที่ 8 พบว่าขั้นแรกผู้เรียนสามารถจำเนื้อหาการจำนิยาม รูปแบบ ข้อความจริง หรือการแสดงกระบวนการง่ายๆ ต่อมาการใช้กระบวนการ นักเรียนตัดสินใจ ด้วยตนเองว่าจะแก้ปัญหาอย่างไร อ่านข้อมูลและต้องทราบว่าจะนำข้อมูลที่ได้มาสรุปอย่างไร หลังจากนั้นสามารถให้เหตุผล การวางแผน การใช้หลักฐานเพื่อตัดสินใจ อย่างสมเหตุสมผล อธิบายข้อเท็จจริงในรูปความคิด รวบยอดและขั้นสุดท้ายสามารถวางแผน การพัฒนา การคิดขั้นสูงและงานต้อง ซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง

กราฟเปรียบเทียบ ระดับความสามารถ กับจำนวนผู้เรียนในแต่ละระดับความสามารถ



## การนำเสนอผลการประเมินผู้เรียนเป็นรายบุคคล

**Diagnostic Scoring Report | โรงเรียนประชารัฐวิทยาเสริม อ.พระยืน จ.ขอนแก่น**

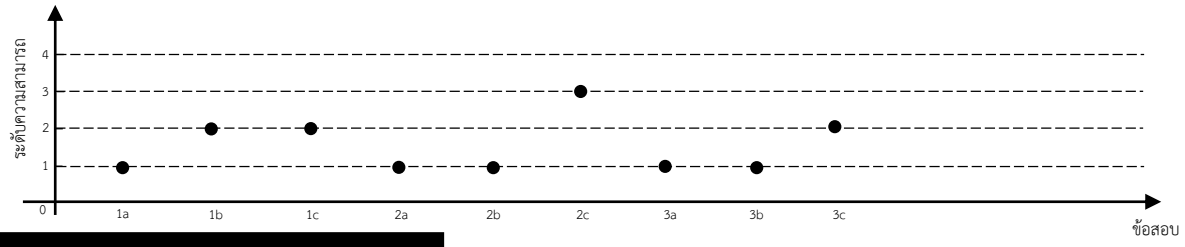
**ชื่อ**.....

Review Your Answer

| Question | 1a | 1b | 1c | 2a | 2b | 2c | 3a | 3b | 3c |
|---|---|---|---|---|---|---|---|---|---|
| Full Score | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |
| Your Score | 1 | 2 | 2 | 1 | 1 | 3 | 1 | 1 | 2 |
| Difficulty | e | m | h | e | m | h | e | m | h |

Guide; e=Easy, m=Medium, h=Hard

**ระดับความสามารถของนักเรียนรายข้อ**




**วิเคราะห์ระดับความสามารถของนักเรียน**

**คะแนนรวมที่ได้** : 13/18 **คะแนน EAP**: 0.99582

**ระดับตาม Progress Map**: 3 (ยุทธวิธีในการคิด Strategic Thinking)

**คำอธิบายคุณลักษณะ:** นักเรียนให้เหตุผล การวางแผน การใช้หลักฐานเพื่อตัดสินใจ อย่างสมเหตุสมผล อธิบายข้อเท็จจริงในรูปความคิด รวบยอด

**สิ่งที่นักเรียนมีความจำเป็นต้องปรับปรุงพัฒนาเพื่อให้มีความสามารถที่สูงขึ้น**: การวางแผน การพัฒนา การคิดขั้นสูงและงานต้อง ซับซ้อนมาก นักเรียนจำเป็นต้องสร้างความเชื่อมโยง

**ภาพที่ 7 แบบรายงานคะแนนนักเรียนรายบุคคล**

**9. สรุปผลการวิจัยและอภิปรายผลการวิจัย (Conclusions and Discussions)**

9.1 จากการพัฒนาและตรวจสอบคุณภาพกรอบการประเมินการความก้าวหน้าในการเรียนรู้ (Learning Progression) ทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษา ทำให้เราได้กรอบการประเมินความก้าวหน้าในการเรียนรู้ทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษาที่มีคุณภาพและมีความถูกต้องเชื่อถือได้

9.2 จากการประเมินการความก้าวหน้าในการเรียนรู้ (Learning Progression) ทักษะการรู้สารสนเทศ (Information literacy Skill) สำหรับนักเรียนมัธยมศึกษาพบว่านักเรียนมีความก้าวหน้าในการเรียนรู้และเกิดการเรียน ทำให้ครูสามารถนำเอาวิธีการและหลักการประเมินนี้ไปใช้ เพื่อให้รู้จักระดับความสามารถของนักเรียนรายบุคคล และวางแผนพัฒนานักเรียนรายบุคคลได้

**10. ข้อเสนอแนะในการทำวิจัยและการนำไปใช้ให้เกิดขึ้นในบริบทจริง (Recommendations and Implications)**

1. เพื่อให้ความคลอบคลุมเนื้อหาสาระ ควรศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องให้มากยิ่งขึ้น
2. ควรมีการวิเคราะห์ผลการวิจัยเพื่อพัฒนาและประยุกต์ใช้กับงานอื่นๆเพิ่มเติม

**11. กิตติกรรมประกาศ (Acknowledgements)**

ขอขอบคุณท่าน ผู้ช่วยศาสตราจารย์ ดร. พัชรี จันทร์เพ็ง สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น ที่คอยให้คำแนะนำ คำปรึกษาตลอดการทำงานวิจัยครั้งนี้

ขอขอบคุณเพื่อน ๆ สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น รุ่นที่ 22 ทุกคนที่คอยให้คำแนะนำ ให้กำลังใจซึ่งกันและกันตลอดมา

**12. เอกสารอ้างอิง (References)**

พัชรี จันทร์เพ็ง. (2559ก). **เอกสารคำสอน รายวิชา 217 720 ระเบียบวิธีวิจัยทางการศึกษา (Research Methodology in Education)**. ขอนแก่น: คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น.

----------. (2559ข). **เอกสารคำสอน รายวิชา 217 920 ระเบียบวิธีวิจัยขั้นสูงทางการศึกษา (Advanced Research Methodology in Education).** ขอนแก่น: คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น.

----------. (2559ค). **การประเมินความก้าวหน้าในการเรียนรู้ (Learning Progression): โมเดลแผนที่เชิงโครงสร้าง (Construct Modelling).** ขอนแก่น: คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น.

----------. (2559ง). **การประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติเพื่อการวิจัย**. ขอนแก่น: คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น.

โครงการพัฒนาการคิดขั้นสูงทางคณิตศาสตร์ของนักเรียนในเขตภาคตะวันออกเฉียงเหนือ. (2559).**สารประชาสัมพันธ์ "โครงการพัฒนาการคิดขั้นสูงทางคณิตศาสตร์ของนักเรียนในเขตภาคตะวันออกเฉียงเหนือ"**. ขอนแก่น: โรงพิมพ์มหาวิทยาลัยขอนแก่น.

Adams, R. J., and Wilson, M. (1996). Formulating the Rasch model as a mixed coefficients multinomial logit. In G. Engelhard and M. Wilson (Eds.), **Objective measurement: Theory into practice** (Vol III; pp. 143-166). Norwood, NJ: Ablex.

Adams, R. J., Wilson, M., and Wang, W. (1997). The multidimensional random coefficients multinomial logit model. **Applied Psychological Measurement**, 21, 1-23.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational, & Psychological Testing (US). (2014). **Standards for educational and psychological testing**. Washington, DC: American Educational Research Association.

Camara, W. (2014). **Standards for Educational and Psychological Testing: Historical Notes**. [Public briefing]. Retrieved from http://www.aera.net/Portals/38/docs/Outreach/Standards_Hill_Briefing_Slides_FI NAL.pdf?timestamp=1410876719244

Creswell, J. W., & Plano Clark, V. L. (2011). **Designing and conducting mixed methods research** (2nd ed.). Thousand Oaks, CA: Sage.

Inprasitha, M. (2015). Preparing ground for the introduction of lesson study in Thailand. Lesson Study Challenges in Mathematics Education. **Editor by Inprasitha, M., Isoda, M., Wang-Iverson, P., & Yeap, B. H. 109-122 Singapore: World Scientific real world situations.**

Junpeng, P., & Chinjunthug, S. (2011). The Development of a model for Mathematics teachers development in classroom assessment of the northeast basic education level. **Journal of Social Science Research, Social Science Research Association of Thailand,** 143, 141-166.

Junpeng, P., & Wilson, M., (2016). **Validating the framework for assessing mathematical problem solving of Thai Students.** California: Graduate School of Education, University of California at Berkeley.

Junpeng, P., & Wilson, M., (inpress). **Modelling of the Open-ended Items for Assessing Multiple Proficiencies in Mathematical Problem Solving**. California: Graduate School of Education, University of California at Berkeley.

Junpeng, P., Wilson, M., Inprasite, M., & Marwiang M. (unpublished data). **Developing the framework for assessing cognitive process of Thai students' mathematical problem solving.** California: Graduate School of Education, University of California at Berkeley.

Junpeng. P. (2012). The development of classroom assessment system in Mathematics for basic education of thailand. **Procedia - Social and Behavioral Sciences**, 69, 1965-1972.

Marwiang, M., Junpeng, P., & Na Nakorn, N. (2014). The development of a model for mathematics classroom assessment: collaborative assessment pyramid, **Procedia – Social and Behavioral Sciences**, 143, 764–768.

National Research Council. (2014). **Developing Assessments for the Next Generation Science Standards**. Committee on Developing Assessments of Science Proficiency in K-12. Board on Testing and Assessment and Board on Science Education, J.W. Pellegrino, M.R. Wilson, J.A. Koenig, and A.S. Beatty, *Editors.* Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Wilson, M. (2005). **Constructing Measures: An Item Response Modelling Approach**. Mahwah, NJ: Lawrence Erlbaum Assoc.

Wilson, M., & Adams, R. J. (1995). Rasch models for item bundles. **Psychometrika**, 60, 181–198.

Wilson, M., & Scalise, K. (2015). Assessment of Learning in Digital Networks. In P. Griffin and E. Care (Eds.). ***Assessment and teaching of 21st century skills: Methods and approach***. Educational Assessment in an Information Age. Dordrecht, Springer: 57-81

Wilson, M., & Sloane, K. (2000). From Principles to Practice: An Embedded Assessment System. **Applied Measurement in Education**, 13 (2), 181–208.

Wilson, M., Beja, I., Scalise, K.,Templin, J., Wiliam, D., & Irribarra, D. T. (2012). Perspectives on Methodological Issues. In Griffin, P., McGaw, B., & Care, E. (Eds.).

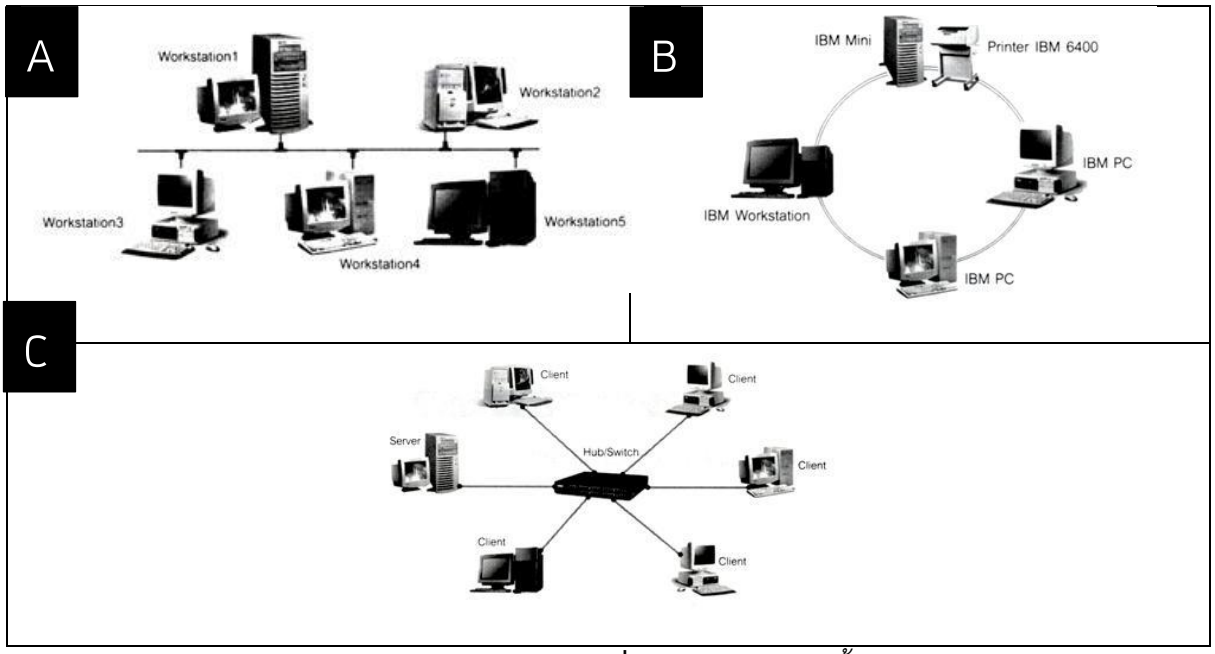
**Assessment and Teaching of 21st Century Skills springer New York 2012**, Springer: 67-190

Wu, M. L., Adams, R. J., Wilson, M. R., Haldane, S.A. (2007). **ACERConQuest Version 2: Generalised item response modelling software.** Camberwell: Australian Council for Educational Research.

**13. ภาคผนวก**

13.1 เครื่องมือฉบับสมบูรณ์ที่นำไปเก็บจริงในสถานศึกษา พร้อมกระดาษคำตอบ

Item set 1




**1a. ให้นักเรียนบอกว่าแต่ละรูปแบบเป็นการเชื่อมต่อแบบใด** *(วัดขั้นความรู้ความจำ Recall)*

A.....

B.....

C.....

**1b. หากต้องการความเร็วในการส่งสัญญาณ และมีความคงทนสูง โดยหากสายเคเบิ้ลบางโหนดเสียหาย จะไม่กระทบต่อโหนดอื่นๆ ควรเลือกการเชื่อมต่อในแบบใด และจะเลือกอุปกรณ์เครือข่ายใด เพื่อให้ทำงานร่วมกันได้ดี พร้อมบอกเหตุผลที่เลือกรูปแบบ และอุปกรณ์เครือข่ายนั้น** *(วัดขั้นทักษะ หรือความคิดรวบยอด Skill/Concept)*

.....

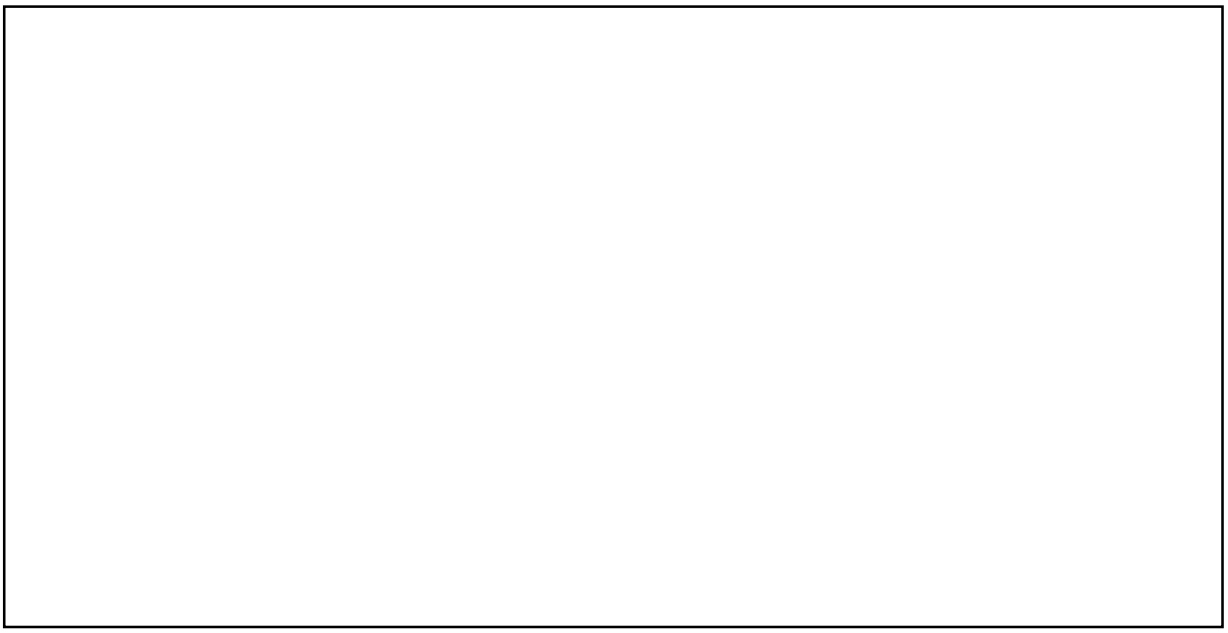
.....

.....

.....

.....

**1c. ให้นักเรียนออกแบบระบบเครือข่ายสำหรับห้องเรียนคอมพิวเตอร์ ที่มีจำนวนเครื่อง 30 เครื่อง พร้อมทั้ง (1) อธิบายรูปแบบการเขื่อมต่อ (2) เหตุผลในการเลือกรูปแบบและอุปกรณ์ (3) ระบุข้อด้อย หรือข้อดีของระบบที่ออกแบบ** *(วัดขั้นยุทธวิธีในการคิด Strategic Thinking)*



.....

.....

.....

.....

.....

.....

.....

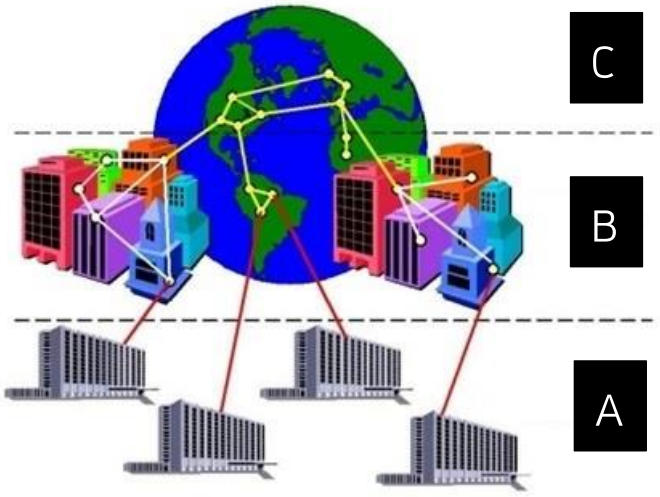
.....

.....

.....

.....

Item set 2




**2a. การเชื่อมต่อในแต่ละระดับ คือการเชื่อมต่อเครือข่ายแบบใด** *(วัดขั้นความรู้ความจำ Recall)*

.....

.....

.....

.....

**2b. ให้นักเรียนอธิบายลักษณะการเชื่อมต่อเครือข่ายในแต่ละแบบ รวมทั้งอธิบายว่าในแต่ละรูปแบบมีความสัมพันธ์กันอย่างไร** *(วัดขั้นทักษะ หรือความคิดรวบยอด Skill/Concept)*

.....

.....

.....

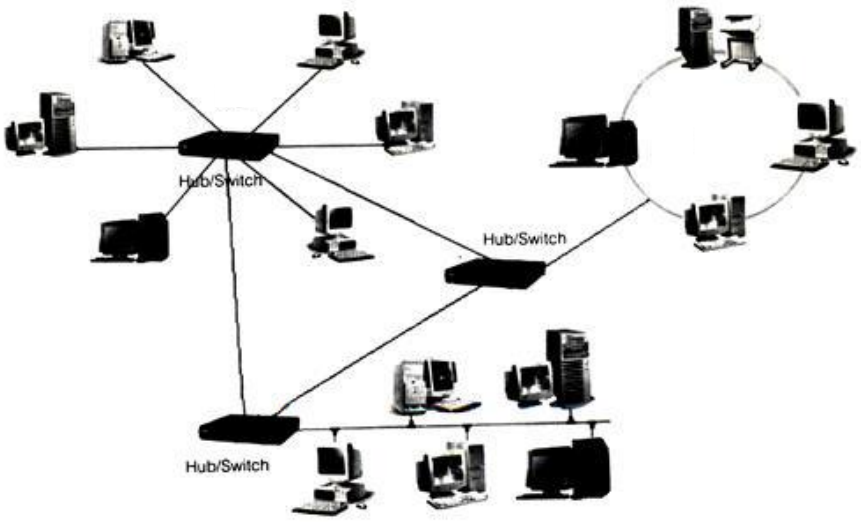
.....

.....

.....

**2c. ให้นักเรียนให้เหตุผลว่าของเครือข่ายจากภาพข้างต้นมีข้อดี หรือข้อเสียอย่างไร หากนักเรียนต้องการแบ่งใหม่ นักเรียนจะเลือกใช้เกณฑ์อะไร เพราะเหตุผลใด** *(วัดขั้นยุทธวิธีในการคิด Strategic Thinking)*

Item set 3




**3a. จากภาพที่กำหนดให้มีการเชื่อมกี่รูปแบบ อะไรบ้าง และรวมเรียกว่าอะไร** *(วัดขั้นความรู้ความจำ Recall)*

.....

.....

.....

.....

.....

.....

**3b. ให้นักเรียนอธิบายว่าเพราะเหตุใดจึงมีรูปแบบการเชื่อมต่อดังกล่าวเกิดขึ้นมา** *(วัดขั้นทักษะหรือความคิดรวบยอด Skill/Concept)*

.....

.....

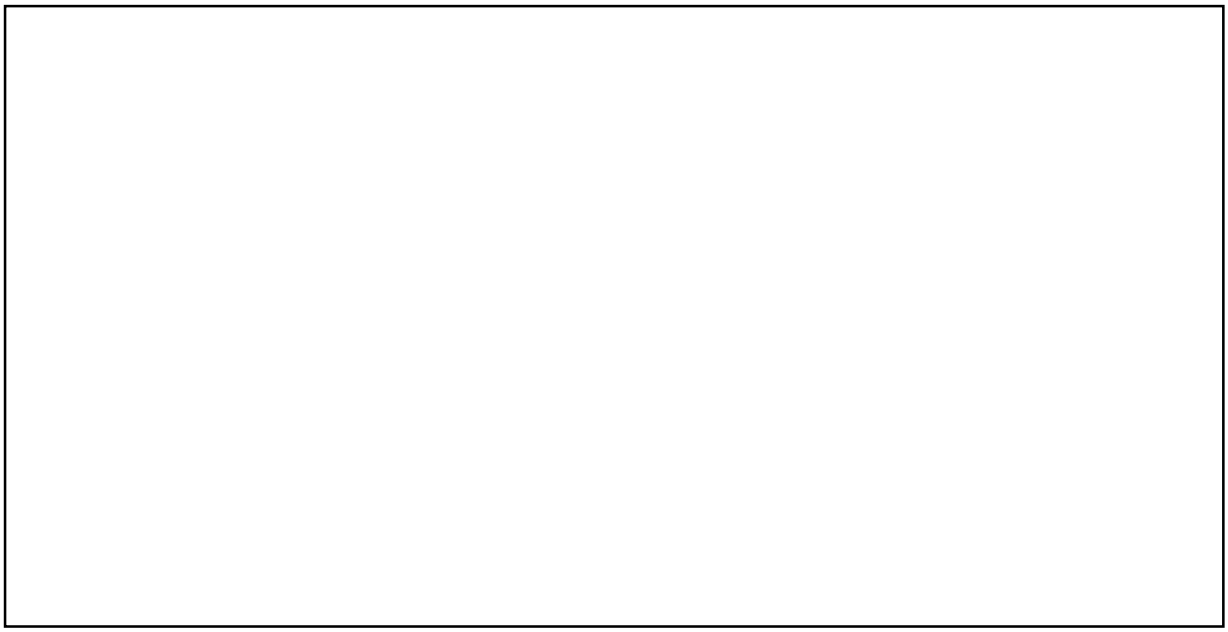
.....

.....

.....

.....

**3c. จากภาพข้างต้น ให้นักเรียนตรวจสอบว่าระบบนี้มีความเหมาะสมกับงานใด และมีข้อบกพร่องหรือข้อด้อยในด้านใด หากเป็นนักเรียนจะออกแบบอย่างไร** *(วัดขั้นยุทธวิธีในการคิด Strategic Thinking)*



.....

.....

.....

.....

.....

.....

.....

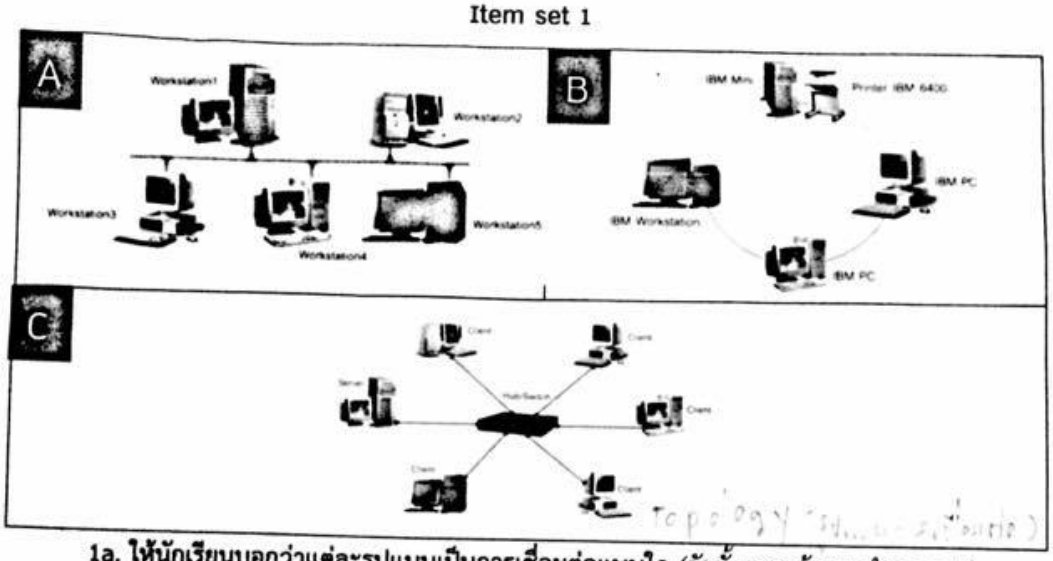
.....

.....

.....

13.2 ตัวอย่างผลการตอบผู้เรียนที่น่าสนใจ

**Item set 1**




**1a. ให้นักเรียนบอกว่าแต่ละรูปแบบเป็นการเชื่อมต่อแบบใด** *(วัดขั้นความรู้ความจำ Recall)*

A. โทโปโลยี... บัส (bus topology)

B. โท... โอยี...แบบวงแหวน (ring topology)

C. โทโปโลยีแบบดาวแหวน (star Network)

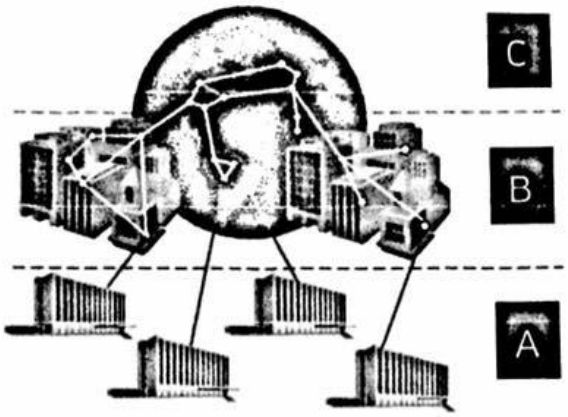
**1b. หากต้องการความเร็วในการส่งสัญญาณ และมีความคงทนสูง โดยหากสายเคเบิ้ลบางโหนดเสียหาย จะไม่กระทบต่อโหนดอื่นๆ ควรเลือกการเชื่อมต่อในแบบใด และจะเลือกอุปกรณ์เครือข่ายใด เพื่อให้ทำงานร่วมกันได้ดี พร้อมบอกเหตุผลที่เลือกรูปแบบ และอุปกรณ์เครือข่ายนั้น** *(วัดขั้นทักษะ หรือความคิดรวบยอด Skill/Concept)*

การสร้างเครือข่ายแบบดาว (star network) เป็นเครือข่ายที่เชื่อมต่อคอมพิวเตอร์เข้ากับอุปกรณ์กลางเป็นศูนย์กลางของเครือข่าย โดยการทำสถานการณ์ต่างๆ มาต่อรวมกัน สังเกตสายของการเกิดข้อผิดพลาด: เช่น สถานที่ๆ: การทำให้ได้ด้วยการ คลิกต่อทางการส่งสัญญาณของ ตัวเชื่อมศูนย์กลางของการเกิดข้อของการเชื่อมต่อ: เช่น สถานการณ์ต่างๆ

(ข้อดี) สะดวกเชื่อมต่อเครื่องใหม่ จะทำให้ใช้งานไม่กระทบกับเครื่องอื่นในระบบ

(ข้อเสีย) ค่าใช้จ่ายสูง และมีผลการเสียหายของคอมพิวเตอร์ทำให้ระบบจะไม่ใช้งานได้

**Item set 2**




**2a. การเชื่อมต่อในแต่ละระดับ คือการเชื่อมต่อเครือข่ายแบบใด** *(วัดขั้นความรู้ความจำ Recall)*

A LAN (Local Area Network) ระบบเครือข่ายระดับท้องถิ่น

B MAN (Metropolitan Area Network) ระบบเครือข่ายระดับเมือง

C WAN (Wide Area Network) ระบบเครือข่ายระดับประเทศ

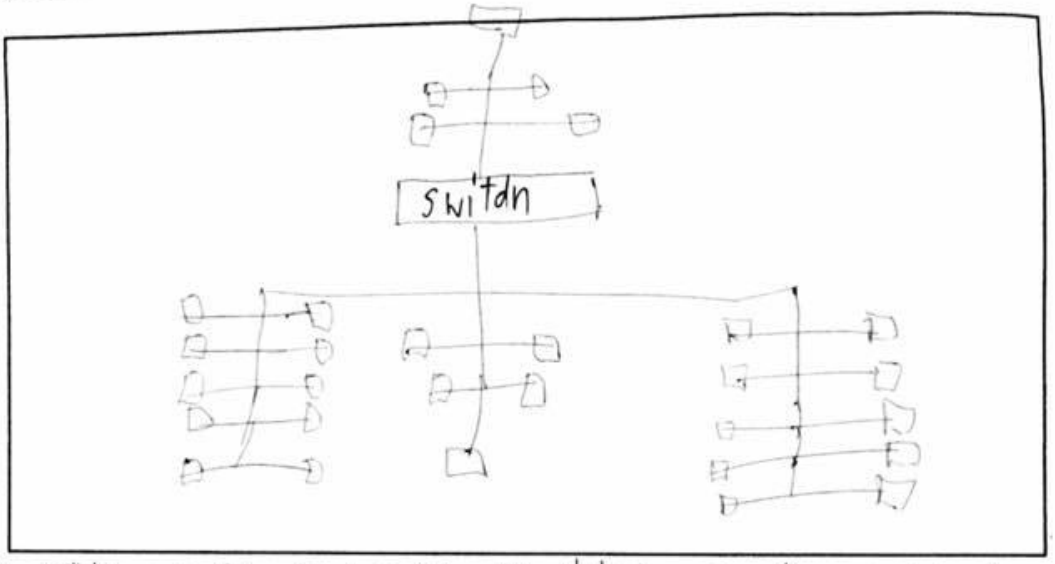
**2b. ให้นักเรียนอธิบายลักษณะการเชื่อมต่อเครือข่ายในแต่ละแบบ รวมทั้งอธิบายว่าในแต่ละรูปแบบมีความสัมพันธ์กันอย่างไร** *(วัดขั้นทักษะ หรือความคิดรวบยอด Skill/Concept)*

Lan เป็นเครือข่ายระดับท้องถิ่น เป็นระบบไม่ซับซ้อนอยู่ในตึกตึกเดียวกัน

Man เป็นเครือข่ายระดับเมือง เป็นระบบที่ซับซ้อนอยู่ระหว่าง Lan และ Wan เป็นเครือข่ายที่ใช้ภายในเมืองบริษัทใดบริษัทหนึ่งเท่านั้น

Wan เป็นเครือข่ายระดับประเทศใช้ติดต่อในบริเวณต่างๆ

**1c. ให้นักเรียนออกแบบระบบเครือข่ายสำหรับห้องเรียนคอมพิวเตอร์ ที่มีจำนวนเครื่อง 30 เครื่อง พร้อมทั้ง (1) อธิบายรูปแบบการเชื่อมต่อ (2) เหตุผลในการเลือกรูปแบบและอุปกรณ์ (3) ระบุข้อด้อย หรือข้อดีของระบบที่ออกแบบ** *(วัดขั้นยุทธวิธีในการคิด Strategic Thinking)*



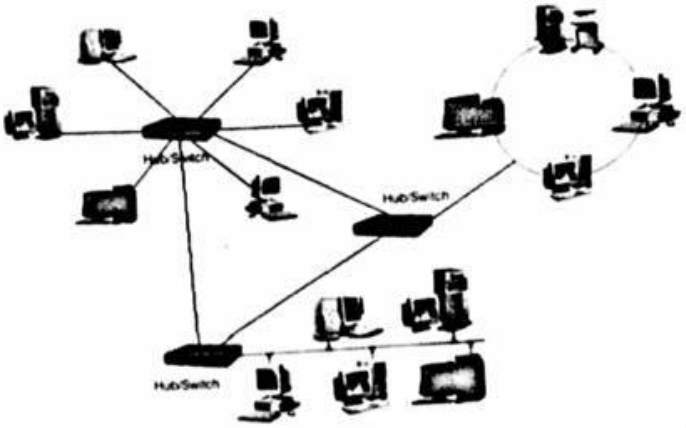

1. เครือข่ายแบบบัส (bus topology) เป็นเครือข่าย ที่เชื่อมต่อคอมพิวเตอร์แต่ละเครื่องเข้ากับสายเคเบิลเส้นเดียวกัน

2. โครงสร้างเครือข่ายคอมพิวเตอร์แบบวงแหวน (ring topology) เป็นเครือข่ายที่ต่อคอมพิวเตอร์ด้วยการยึดปลายสายแต่ละเส้นเข้ากับอุปกรณ์ถัดไป

3. โครงสร้างเครือข่ายคอมพิวเตอร์แบบดาว (star network) เป็นเครือข่ายที่เชื่อมต่อคอมพิวเตอร์เข้ากับอุปกรณ์ที่เป็นจุดศูนย์กลางของเครือข่าย

4. โครงสร้างเครือข่ายแบบผสม (Hybrid Topology) เป็นเครือข่ายที่ผสมผสานในระหว่างแบบดาวและแบบบัส

Item set 3

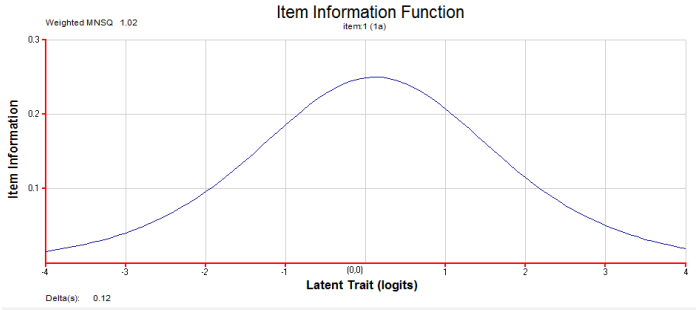



**3a. จากภาพที่กำหนดให้มีการเชื่อมกี่รูปแบบ อะไรบ้าง และรวมเรียกว่าอะไร** *(วัดขั้นความรู้ความจำ Recall)*

**3b. ให้นักเรียนอธิบายว่าเพราะเหตุใดจึงมีรูปแบบการเชื่อมต่อดังกล่าวเกิดขึ้นมา** *(วัดขั้นทักษะหรือความคิดรวบยอด Skill/Concept)*

13.3 ตารางการตรวจสอบคุณภาพของข้อสอบแต่ละข้อเทียบกับ Progress Map ตัวอย่างการตอบของนักเรียนแต่ละระดับของ Progress Map พร้อมกราฟ ICC และ Item Information

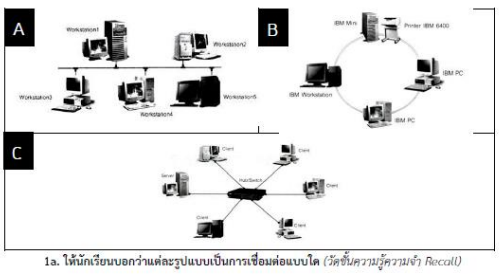
**ข้อที่ 1**

<table>
<tr><th>คุณภาพของข้อสอบ</th><th>Progress Map (ก่อนปรับ)</th><th>Progress Map (หลังปรับ)</th><th>คำอธิบาย</th><th>ตัวอย่างคำตอบของนักเรียน</th></tr>
<tr><td>ค่าความยาก (b)</td><td>0.119</td><td>0.119</td><td>ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย</td><td></td></tr>
<tr><td>ค่าอำนาจจำแนก</td><td>0.67<br>MNSQ= 0.89<br>CI= 0.56, 1.44<br>T=-0.5</td><td>0.67<br>MNSQ= 0.89<br>CI= 0.56, 1.44<br>T=-0.5</td><td>เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้</td><td></td></tr>
<tr><td>Step ของการตอบแต่ละระดับ</td><td>0 -> 1</td><td>0 -> 1</td><td>-</td><td></td></tr>
<tr><td>กราฟ ICC<br>กราฟ Item Information</td><td colspan="4"> 
</td></tr>
</table>

**ตีความผลการวิเคราะห์คุณภาพของข้อสอบพร้อมแนวทางการปรับปรุงแก้ไข**

ข้อสอบข้อ 1a ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่ายเหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง

**ข้อสอบที่แก้ไขแล้ว เพื่อจะเก็บไปใช้ครั้งต่อไป**



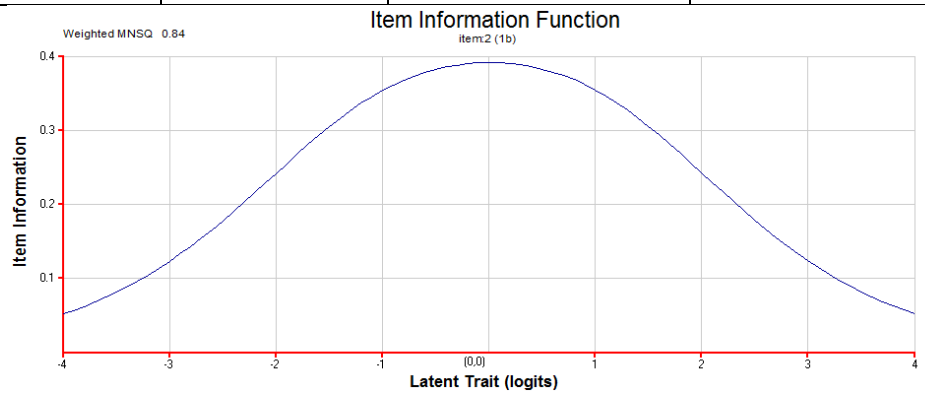

1a. ให้นักเรียนบอกว่าแต่ละรูปแบบเป็นการเชื่อมต่อแบบใด *(วัดขั้นความรู้ความจำ Recall)*

A.....

B.....

C.....

**ข้อที่ 2**

| คุณภาพของข้อสอบ | Progress Map (ก่อนปรับ) | Progress Map (หลังปรับ) | คำอธิบาย | ตัวอย่างคำตอบของนักเรียน |
|---|---|---|---|---|
| ค่าความยาก (b) | 0.004 | 0.004 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย | |
| ค่าอำนาจจำแนก | 0.82<br>MNSQ= 0.84<br>CI= 0.55, 1.45<br>T=-0.7 | 0.82<br>MNSQ= 0.84<br>CI= 0.55, 1.45<br>T=-0.7 | เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | |
| Step ของการตอบแต่ละระดับ | 0 -> 1 -> 2 | 0 -> 1 -> 2 | - | |
| กราฟ ICC<br>กราฟ Item Information |  | | | |

**ตีความผลการวิเคราะห์คุณภาพของข้อสอบพร้อมแนวทางการปรับปรุงแก้ไข**

ข้อสอบข้อ 1a ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่ายเหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง

**ข้อสอบที่แก้ไขแล้ว เพื่อจะเก็บไปใช้ครั้งต่อไป**

**1b. หากต้องการความเร็วในการส่งสัญญาณ และมีความคงทนสูง โดยหากสายเคเบิ้ลบางโหนดเสียหาย จะไม่กระทบต่อโหนดอื่นๆ ควรเลือกการเชื่อมต่อในแบบใด และจะเลือกอุปกรณ์เครือข่ายใดเพื่อให้ทำงานร่วมกันได้ดี พร้อมบอกเหตุผลที่เลือกรูปแบบ และอุปกรณ์เครือข่ายนั้น** *(วัดขั้นทักษะ หรือความคิดรวบยอด Skill/Concept)*

.....

.....

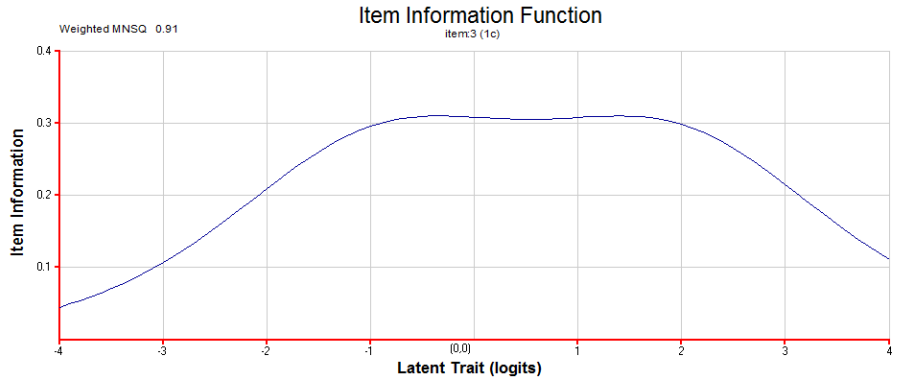
.....

.....

.....

.....

**ข้อที่ 3**

| คุณภาพของข้อสอบ | Progress Map (ก่อนปรับ) | Progress Map (หลังปรับ) | คำอธิบาย | ตัวอย่างคำตอบของนักเรียน |
|---|---|---|---|---|
| ค่าความยาก (b) | 0.526 | 0.526 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย | |
| ค่าอำนาจจำแนก | 0.79<br>MNSQ= 0.91<br>CI= 0.58, 1.42<br>T=-0.4 | 0.79<br>MNSQ= 0.91<br>CI= 0.58, 1.42<br>T=-0.4 | เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | |
| Step ของการตอบแต่ละระดับ | 0 -> 1-> 2-> 3 | 0 -> 1 -> 2-> 3 | - | |
| กราฟ ICC<br>กราฟ Item Information |  | | | |

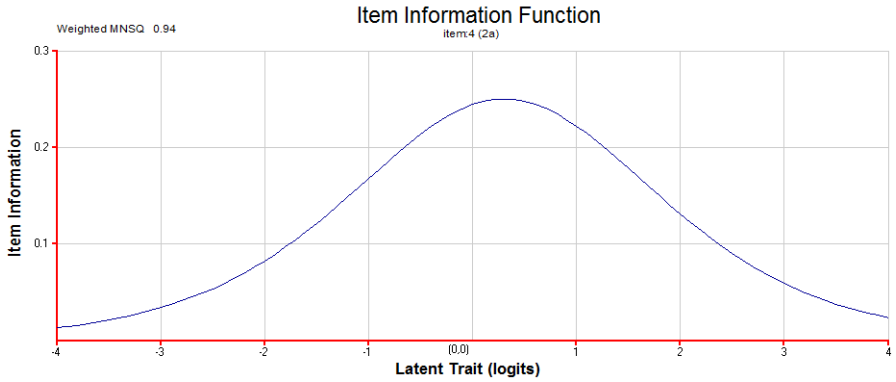
**ตีความผลการวิเคราะห์คุณภาพของข้อสอบพร้อมแนวทางการปรับปรุงแก้ไข**

ข้อสอบข้อ 1a ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่ายเหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง

**ข้อสอบที่แก้ไขแล้ว เพื่อจะเก็บไปใช้ครั้งต่อไป**

1c. ให้นักเรียนออกแบบระบบเครือข่ายสำหรับห้องเรียนคอมพิวเตอร์ ที่มีจำนวนเครื่อง 30 เครื่อง พร้อมทั้ง (1) อธิบายรูปแบบการเชื่อมต่อ (2) เหตุผลในการเลือกรูปแบบและอุปกรณ์ (3) ระบุข้อด้อยหรือข้อดีของระบบที่ออกแบบ *(วัดขั้นปฏิบัติในการคิด Strategic Thinking)*

**ข้อที่ 4**

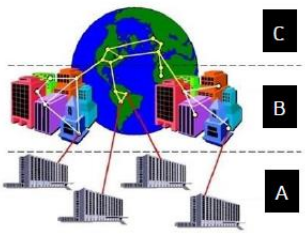
| คุณภาพของข้อสอบ | Progress Map (ก่อนปรับ) | Progress Map (หลังปรับ) | คำอธิบาย | ตัวอย่างคำตอบของนักเรียน |
|---|---|---|---|---|
| ค่าความยาก (b) | 0.306 | 0.306 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย | |
| ค่าอำนาจจำแนก | 0.67<br>MNSQ= 0.94<br>CI= 0.58, 1.42<br>T=-0.3 | 0.67<br>MNSQ= 0.94<br>CI= 0.58, 1.42<br>T=-0.3 | เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | |
| Step ของการตอบแต่ละระดับ | 0 -> 1 | 0 -> 1 | - | |
| กราฟ ICC<br>กราฟ Item Information |  | | | |

**ตีความผลการวิเคราะห์คุณภาพของข้อสอบพร้อมแนวทางการปรับปรุงแก้ไข**

ข้อสอบข้อ 1a ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่ายเหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง

**ข้อสอบที่แก้ไขแล้ว เพื่อจะเก็บไปใช้ครั้งต่อไป**

Item set 2

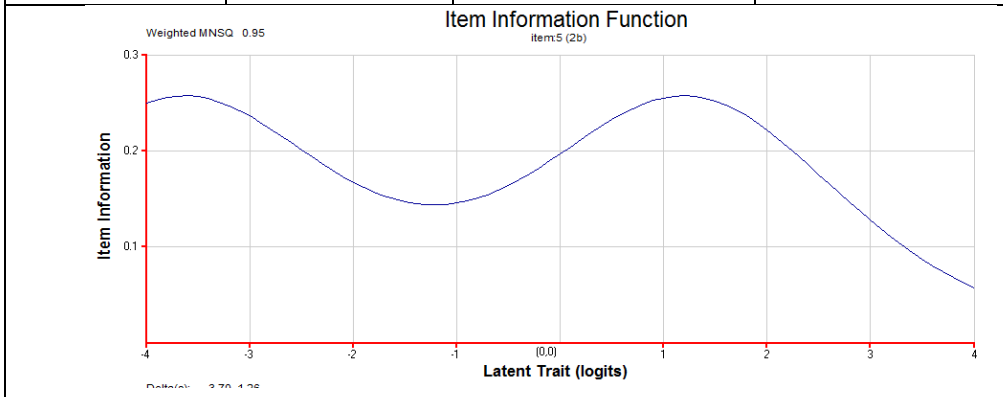



2a. การเชื่อมต่อในแต่ละระดับ คือการเชื่อมต่อเครือข่ายแบบใด *(วัดขั้นความรู้ความจำ Recall)*

.....

.....

**ข้อที่ 5**

| คุณภาพของข้อสอบ | Progress Map (ก่อนปรับ) | Progress Map (หลังปรับ) | คำอธิบาย | ตัวอย่างคำตอบของนักเรียน |
|---|---|---|---|---|
| ค่าความยาก (b) | -1.218 | -1.218 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย | |
| ค่าอำนาจจำแนก<br><br>Step ของการตอบแต่ละระดับ | 0.69<br>MNSQ= 0.95<br>CI= 0.58, 1.42<br>T=-0.2 | 0.69<br>MNSQ= 0.95<br>CI= 0.58, 1.42<br>T=-0.2 | เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | |
| | 0 -> 1-> 2 | 0 -> 1 -> 2 | - | |
| กราฟ ICC<br><br>กราฟ Item Information |  | | | |

**ตีความผลการวิเคราะห์คุณภาพของข้อสอบพร้อมแนวทางการปรับปรุงแก้ไข**

ข้อสอบข้อ 1a ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่ายเหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง

**ข้อสอบที่แก้ไขแล้ว เพื่อจะเก็บไปใช้ครั้งต่อไป**

**2b. ให้นักเรียนอธิบายลักษณะการเชื่อมต่อเครือข่ายในแต่ละแบบ รวมทั้งอธิบายว่าในแต่ละรูปแบบมีความสัมพันธ์กันอย่างไร** *(วัดขั้นทักษะ หรือความคิดรวบยอด Skill/Concept)*

.....

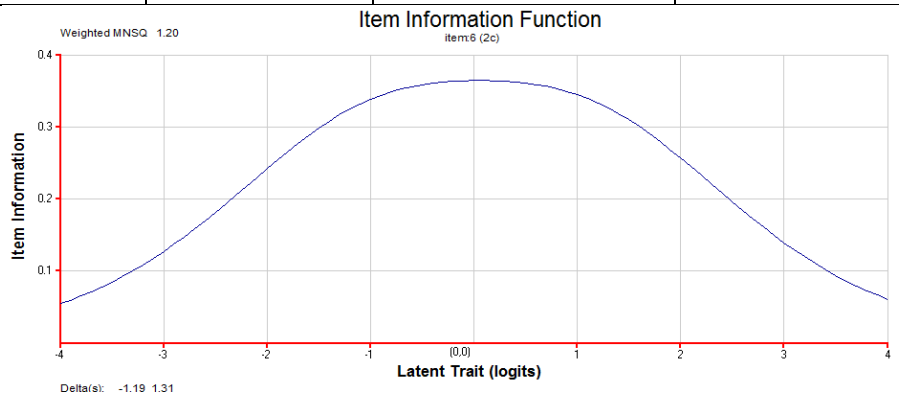
.....

.....

.....

.....

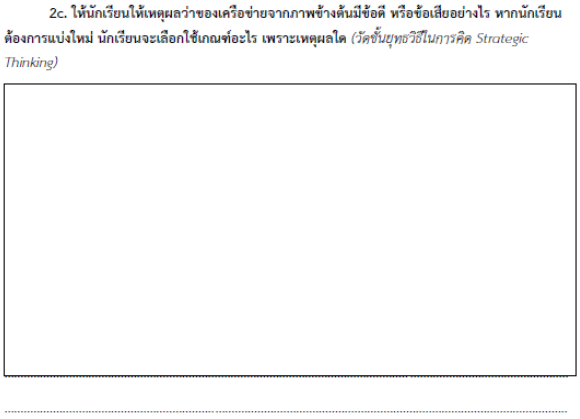
**ข้อที่ 6**

| คุณภาพของข้อสอบ | Progress Map (ก่อนปรับ) | Progress Map (หลังปรับ) | คำอธิบาย | ตัวอย่างคำตอบของนักเรียน |
|---|---|---|---|---|
| ค่าความยาก (b) | 0.062 | 0.062 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย | |
| ค่าอำนาจจำแนก | 0.69<br>MNSQ= 1.20<br>CI= 0.56, 1.44<br>T=0.9 | 0.69<br>MNSQ= 1.20<br>CI= 0.56, 1.44<br>T=0.9 | เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | |
| Step ของการตอบแต่ละระดับ | 0 -> 1-> 2-> 3 | 0 -> 1 -> 2-> 3 | - | |
| กราฟ ICC<br>กราฟ Item Information |  | | | |

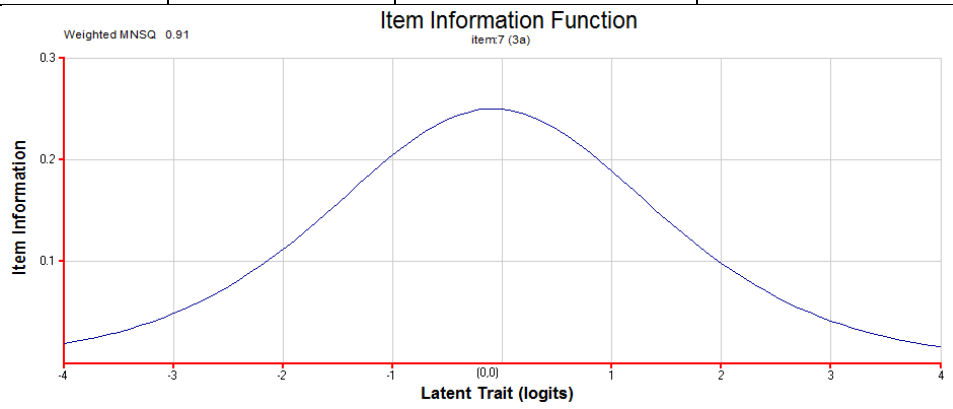
**ตีความผลการวิเคราะห์คุณภาพของข้อสอบพร้อมแนวทางการปรับปรุงแก้ไข**

ข้อสอบข้อ 1a ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่ายเหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง

**ข้อสอบที่แก้ไขแล้ว เพื่อจะเก็บไปใช้ครั้งต่อไป**



**ข้อที่ 7**

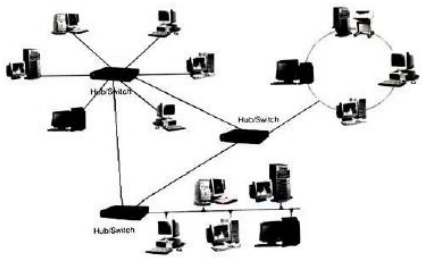
| คุณภาพของข้อสอบ | Progress Map (ก่อนปรับ) | Progress Map (หลังปรับ) | คำอธิบาย | ตัวอย่างคำตอบของนักเรียน |
|---|---|---|---|---|
| ค่าความยาก (b) | -0.087 | -0.087 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย | |
| ค่าอำนาจจำแนก | 0.70<br>MNSQ= 0.91<br>CI= 0.55, 1.45<br>T=-0.4 | 0.70<br>MNSQ= 0.91<br>CI= 0.55, 1.45<br>T=-0.4 | เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | |
| Step ของการตอบแต่ละระดับ | 0 -> 1 | 0 -> 1 | - | |
| กราฟ ICC<br>กราฟ Item Information |  | | | |

**ตีความผลการวิเคราะห์คุณภาพของข้อสอบพร้อมแนวทางการปรับปรุงแก้ไข**

ข้อสอบข้อ 1a ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่ายเหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง

**ข้อสอบที่แก้ไขแล้ว เพื่อจะเก็บไปใช้ครั้งต่อไป**

Item set 3

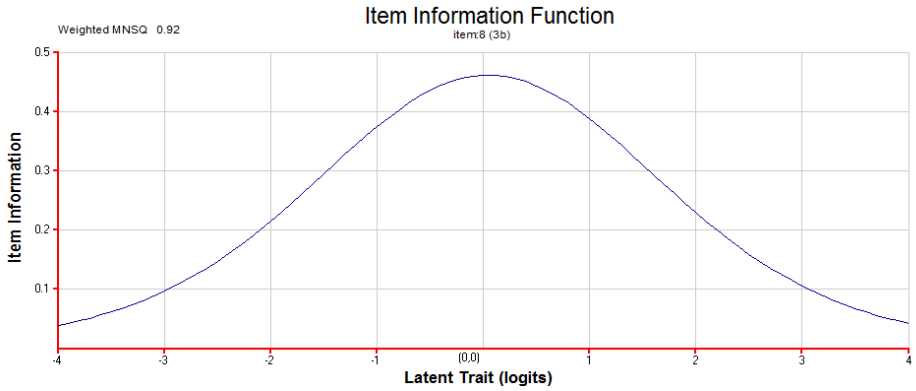


3a. จากภาพที่กำหนดให้มีการเชื่อมกี่รูปแบบ อะไรบ้าง และรวมเรียกว่าอะไร *(วัดขั้นความรู้ความจำ Recall)*

.....

.....

**ข้อที่ 8**

| คุณภาพของข้อสอบ | Progress Map (ก่อนปรับ) | Progress Map (หลังปรับ) | คำอธิบาย | ตัวอย่างคำตอบของนักเรียน |
|---|---|---|---|---|
| ค่าความยาก (b) | 0.048 | 0.048 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย | |
| ค่าอำนาจจำแนก | 0.80<br>MNSQ= 0.92<br>CI= 0.53, 1.47<br>T=-0.3 | 0.80<br>MNSQ= 0.92<br>CI= 0.53, 1.47<br>T=-0.3 | เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | |
| Step ของการตอบแต่ละระดับ | 0 -> 1-> 2 | 0 -> 1 -> 2 | - | |
| กราฟ ICC<br>กราฟ Item Information |  | | | |

**ตีความผลการวิเคราะห์คุณภาพของข้อสอบพร้อมแนวทางการปรับปรุงแก้ไข**

ข้อสอบข้อ 1a ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่ายเหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง

**ข้อสอบที่แก้ไขแล้ว เพื่อจะเก็บไปใช้ครั้งต่อไป**

**3b. ให้นักเรียนอธิบายว่าเพราะเหตุใดจึงมีรูปแบบการเชื่อมต่อดังกล่าวเกิดขึ้นมา** *(วัดชั้นทักษะหรือความคิดรวบยอด Skill/Concept)*

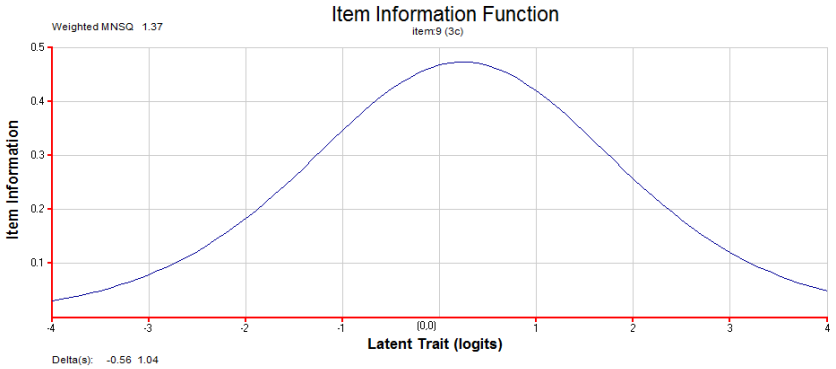
.....

.....

.....

.....

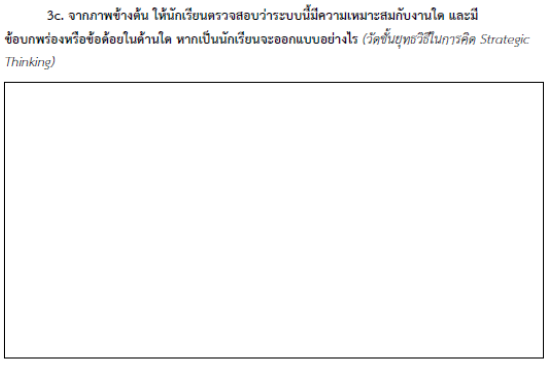
**ข้อที่ 9**

| คุณภาพของข้อสอบ | Progress Map (ก่อนปรับ) | Progress Map (หลังปรับ) | คำอธิบาย | ตัวอย่างคำตอบของนักเรียน |
|---|---|---|---|---|
| ค่าความยาก (b) | 0.240 | 0.240 | ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่าย | |
| ค่าอำนาจจำแนก | 0.69<br>MNSQ= 1.37<br>CI= 0.54, 1.46<br>T=1.5 | 0.69<br>MNSQ= 1.37<br>CI= 0.54, 1.46<br>T=1.5 | เหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง ค่า MNSQ อยู่ในช่วงที่ยอมรับได้ | |
| Step ของการตอบแต่ละระดับ | 0 -> 1-> 2-> 3 | 0 -> 1 -> 2-> 3 | - | |
| กราฟ ICC<br>กราฟ Item Information |  | | | |

**ตีความผลการวิเคราะห์คุณภาพของข้อสอบพร้อมแนวทางการปรับปรุงแก้ไข**

ข้อสอบข้อ 1a ข้อสอบมีค่าความยากระดับปานกลางค่อนไปทางง่ายเหมาะสำหรับนักเรียนที่มีความสามารถระดับกลาง

**ข้อสอบที่แก้ไขแล้ว เพื่อจะเก็บไปใช้ครั้งต่อไป**



13.4 ผลการวิเคราะห์จากโปรแกรม ConQuest2.0

13.4.1 ไฟล์ .cqc

```
Title Partial Credit Model: What happened last night;
data data.dat;
format name 1-2 responses 4-12;
labels << mps.lab;
codes 0,1,2,3,4;
recode (0,1) (0,1)                !items(1,4,7);
recode (0,1,2) (0,1,2)              !items(2,5,8);
recode (1,2,3) (0,1,2)              !items(3,6,9);
model item + item*step;
estimate;
show !estimates=latent >> data.shw;
itanal >> data.itn;
plot expected! gins=2;
plot icc! gins=2;
plot ccc! gins=2;
show cases !estimates=eap >> data.eap;
reset;
Title Partial Credit Model: What happened last night;
data data.dat;
format name 1-2 responses 4-12;
labels << mps.lab;
codes 0,1,2,3,4;
recode (0,1) (0,1)                !items(1,4,7);
recode (0,1,2) (0,1,2)            !items(2,5,8);
recode (1,2,3) (0,1,2)            !items(3,6,9);
model item + item*step;
estimate;
show !estimates=latent >> data.shw;
itanal >> data.itn;
plot expected! gins=2;
plot icc! gins=2;
plot ccc! gins=2;
show cases !estimates=eap >> data.eap;
```

13.4.2 ไฟล์ .dat

```
01 123123123
02 012123123
03 122112112
04 123113122
05 112011001
06 001121013
07 011122122
08 122011112
09 123123113
```

10 112123123
11 112112011
12 001013003
13 123113123
14 012023123
15 123123112
16 123123123
17 012012001
18 001111001
19 112012112
20 123122023
21 123123123
22 123123123
23 113122121
24 012013112
25 012012013
26 122122112
27 001011001
28 112122122
29 123123123
30 122112123
31 123012122
32 112013101
33 123123123
34 123122013
35 001013122
36 122112123
37 001001012
38 123123123
39 122123123
40 012012001

### 13.4.3 ไฟล์ .lab

===> item

1 1a
2 1b
3 1c
4 2a
5 2b
6 2c
7 3a
8 3b
9 3c

13.5 สะท้อนผลการวิจัย

Brown, Nathaniel J. S. , Nagashima Sam O. , Fu, Alice, Timms, Michael and Wilson, Mark. (2010). 'A Framework for Analyzing Scientific Reasoning in Assessments', Educational Assessment, 15:3, 142-174.

การศึกษาระบบการประเมินที่ได้รับการออกแบบและพัฒนาการสร้างระบบโดยใช้แบบจำลองที่เกี่ยวข้องกับการออกแบบการดำเนินงานและการประเมินผลของสี่กลุ่ม (The four building box) คือ Construct Maps, Item Design, Outcome Space และ Measurement Model.

โดยในประเด็นที่สนใจคือ Construct Maps ในส่วนที่เป็นการสร้างโครงสร้างทางความคิดสำหรับการคิดที่มีความซับซ้อน หลักการสำคัญของการสร้างแบบจำลองการสร้างคือการออกแบบการประเมิน ที่ต้องมีการพัฒนาโครงสร้างและแบบแผนเครื่องมือการประเมินที่สามารถตีความที่เกี่ยวข้องกับรูปแบบของกระบวนที่ทำให้มีความรู้ความเข้าใจ โดยรูปแบบส่วนใหญ่แล้วจะยังไม่มีความชัดเจนเท่าที่ควร ยังไม่มีแนวทางที่ชัดเจนเกี่ยวกับวิธีการสร้างแบบประเมินโครงสร้างความคิดที่ตรงตามโครงสร้างของข้อมูลเชิงลึกของผู้เชี่ยวชาญเชิงเนื้อหา แนวความคิดที่ซับซ้อนที่ว่าคือ แนวความคิดความพยายามที่จะจับการคิดด้านโดเมนของนักเรียน

ในประเด็นนี้หากเราสามรถที่จะออกแบบประเมินผลทางด้านความคิดของนักเรียน โดยใช้กระบวนการนี้ได้ โดยอาศัยความเข้าใจในการสร้างแบบจำลอง รูปแบบหรือโครงสร้างที่ซับซ้อนของการเข้าใจของนักเรียน เพื่อให้เกิดความชัดเจน มีรูปแบบการประเมินผลที่เป็นกระบวนการ วิธีการที่ยอมรับ ได้ผลเป็นข้อมูลเชิงลึก ก็จะทำให้เราได้เข้าใจถึงกระวนการคิดที่ซับซ้อน ส่งผลให้เกิดกระบวนการพัฒนาในรูปแบบต่างๆ ที่ตรงจุดมาขึ้น และมีความสอดคล้องมากขึ้นด้วย

ในความคิดแล้วหากเราสามารถสร้างรูปแบบหรือวิธีการ เพื่อให้ได้มาซึ่งวิธีหรือเครื่องมือซึ่งวัดประเมินความคิดที่ซับซ้อน โครงสร้างทางความคิดหรือวิธีการทำความเข้าใจของนักเรียนแล้ว ก็จะส่งผลให้เกิดรูปแบบการเรียนรู้ การจัดการเรียนการสอน รวมไปถึงวิธีการวัดประเมินผลความเข้าใจของนักเรียนได้ตรงจุดด้วย

Wilson, Mark. (2005). Construct Modeling: The “Four Building Blocks” Approach. In Kerry Breen (Eds.). *Constructing Measures: an item response modeling approach* (pp. 3-21). New Jersey, USA: Lawrence Erlbaum Associates.

ในบทที่ 1 จะเป็นการนำเสนอเกี่ยวกับภาพรวมและแนวความคิดเกี่ยวกับ Four Building Blocks ได้อธิบายถึง Construct Modeling จาก The four Building blocks. โดยเริ่มต้นจากการอธิบายในภาพรวมและแนวคิดหลักก่อน จากนั้นอธิบายความหมายของการวัด ว่าต้องมีการกำหนดประเภทของการสังเกตคุณลักษณะของประชากรเองส่งผลต่อคุณลักษณะของการวัดด้วย การกำหนดคุณลักษณะของกลุ่มประชากรจะต้องสอดคล้องกับความจริงที่ต้องการวัด เพื่อให้ได้กระวนการวัดที่สมบูรณ์ วัถตุประสงค์หลักของการวัด ก็เพื่อ เป็นการอธิบาย เตรียมการอย่างมีเหตุผล และเป็นการสรุปผลลักษณะของคนอย่างรวดเร็วทั้งทางด้านความสำเร็จ ทัศนะคติ

สิ่งที่ได้หรือข้อสรุปจากแนวคิดจะมีความเข้าใจพื้นฐาน ลักษณะเฉพาะของเครื่องมือที่สามารถนำมาใช้ออกแบบการวัดได้หลายแบบ ทั้งแบบสำรวจ ข้อสอบ แบบสอบถาม ก็สามารถออกแบบการวัดได้หลากหลายลักษณะ หากต้องการประสบความสำเร็จในการวัดควรมีตัวอย่างที่กว้างและหลากหลายครอบคลุม

ส่วน Construct map เป็นการหาสิ่งที่ต้องการวัดและหารูปแบบหรือแนวทางที่จะนำไปสู่การออกแบบเครื่องการวัด ที่จีมีหลากหลายลักษณะ โดยจะต้องเลือกลักษณะที่เหมาะสม การให้ค่ามาตรา

ส่วน Item Design เป็นขั้นต่อมา เพื่อศึกษาทฤษฎี โครงสร้าง การพิสูจน์จากสภาพจริง เพื่อนำมาซึ่งการสร้างมาตรในการวัด การออกแบบแบบรายข้อ เพื่อใช้เป็นเครื่องมือ

ส่วน Outcome Space การตัดสินคำตอบจากเกณฑ์ ที่กำหนดขึ้น หรือเป็นการกำหนดเกณฑ์การให้คะแนนจากคำตอบ ของแต่ละประเภทที่แตกต่างกัน

ส่วนของ โมเดลการวัดผล เป็นส่วนที่จะต้องทำการตรวจสอบค่า ตรวจสอบกระบวนการวัดผลทั้งหมดรวมทั้งการหาคุณภาพของโมเดลการวัดนี้ด้วย

ในความคิดของผมคิดว่าในบทที่ 1 นี้ต้องการให้ผู้อ่านได้รู้ถึงภาพรวมและหลักแนวคิดของ Four Building Blocks รวมทั้งหลักการและภาพรวมโดยสรุปของทั้ง 4 ด้าน โดยเน้นที่จะอธิบายความหมาย วิธีการดำเนินงานในขั้นตอนนั้น โดยสรุปสาระสำคัญ แนวทางปฏิบัติเพื่อเป็นแนวทางและนำร่องไปสู่บทที่ 2

โดยสรุปแล้วในบทนี้ต้องการจะให้รู้จักถึง Four Building Blocks รวมถึงองค์ประกอบ ความหมาย หลักการ แนวทางการทำในแต่ละด้าน เพื่อเป็นการทบทวนหรือเพื่อให้เข้าใจตรงกัน

13.6 บทความวิชาการ

การจัดการศึกษามีความจำเป็นที่จะต้องกำหนดจุดมุ่งหมายที่มีความชัดเจน เพื่อช่วยในการกำหนดคุณลักษณะที่ต้องการของผู้เรียน เพื่อใช้เป็นแนวทางในการวางแผนเพื่อใช้การจัดการเรียนรู้ จุดมุ่งหมายยังมีส่วนช่วยในการกำหนดทิศทางของการศึกษา การจัดประสบการณ์การเรียนรู้หรือกิจกรรมการจัดการเรียนรู้ที่เหมาะสม การวัดและการประเมินผลจึงเข้ามามีบทบาทสำคัญในการตัดสินผลการเรียนรู้ของผู้เรียนว่าสำเร็จตรงตามจุดมุ่งหมายที่กำหนดไว้หรือไม่ รวมทั้งยังช่วยให้สารสนเทศเพื่อช่วยในการปรับปรุงละพัฒนาการเรียนรู้ของผู้เรียนและการจัดกิจกรรมการเรียนรู้ของผู้สอน เมื่อการวัดและการประเมินผลเข้ามามีบทบาทสำคัญ ทำให้ทฤษฎีการทดสอบเข้ามามีบทบาทสำคัญในการวัดและประเมินผล เนื่องจากการการวัดและประเมินผลทางการศึกษานั้นเป็นสิ่งที่อยู่ภายในไม่สามารถวัดออกมาได้โดยตรง เพื่อทำความเข้าใจในสิ่งที่ต้องการวัดตามคุณลักษณะที่มุ่งวัด โครงสร้างของการวัด และการพัฒนาเครื่องมือ จึงจำเป็นต้องอาศัยทฤษฎีการทดสอบทดสอบเข้ามาช่วย โดยทฤษฎีการทดสอบนั้นหากแบ่งประเภทเป็น 2 ประเภทหลักใหญ่ๆ ได้แก่ ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) และทฤษฎีการทดสอบแนวใหม่ (Modern Test Theory)

ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) จะเน้นไปทางการตรวจสอบความสัมพันธ์ระหว่างคะแนนที่สังเกตได้กับคะแนนที่แท้จริง และวิเคราะห์คุณภาพโดยส่วนรวมของข้อสอบและแบบสอบ โดยใช้การวิเคราะห์คุณภาพของข้อสอบและแบบทดสอบ โดยใช้โมเดลคะแนนจริงแบบดั้งเดิม การใช้ทฤษฎีความเที่ยงและความคลาดเคลื่อน การใช้ความตรงเพื่อตรวจสอบความตรงต่างๆ หลักการการสร้างแบบทดสอบ และใช้หลักการวิเคราะห์ข้อสอบ ทฤษฎีการทดสอบแบบดั้งเดิมยังมีข้อด้อยของข้อตกลงเบื้องต้นมีข้อจำกัดของฐานความเชื่อเกี่ยวกับคะแนนความคลาดเคลื่อนและค่าพารามิเตอร์ของข้อสอบและแบบทดสอบ จึงส่งผลให้มีนักทฤษฎีการทดสอบหลายท่านได้ก่อการปฏิรูปแนวคิดเกิดขึ้น สู่ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theory) เพื่อคลายข้อตกลงเบื้องต้น และแก้ไขจุดอ่อนบางประการ

ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theory) จะมุ่งเน้นการขยายแนวคิดของทฤษฎีการทดสอบแบบดั้งเดิมให้มีขอบเขตที่กว้างขวาง และมีความน่าเชื่อถือมากยิ่งขึ้น โดยปรับข้อตกลงให้มีความสมเหตุสมผล ตามความจริงมากขึ้น เกี่ยวกับความคลาดเคลื่อนของการวัด ยอมให้มีความคลาดเคลื่อนจากการวัดนั้นมีความแตกต่างกันออกไปตามสถานการณ์การวัด เพื่อศึกษาความเที่ยงทั่วไป ภายใต้เงื่อนไขของการวัดตามแนวของ Generalizability Theory: G-Theory รวมทั้งศึกษาความคลาดเคลื่อนจากการวัดที่มีความแตกต่างกันไปตามลักษณะของข้อสอบและแบบสอบ ขึ้นกับระดับความสามารถของผู้สอบ และของข้อสอบและแบบสอบ ทั้งยังพยายามวัดคุณลักษณะภายในหรือความสามารถที่แท้จริงของบุคคลตามแนว Item Response Theory (IRT) นอกจากนี้ความคลาดเคลื่อนจากการวัดของบุคคลเองก็ส่งผลต่อการตอบข้อสอบมากกว่า 1 องค์ประกอบ ดังนั้นเพื่อวัดความสามารถของบุคคลจึงมี 2 องค์ประกอบขึ้นไป เพื่อให้ความสามารถหลายมิติของบุคคลสอดคล้องกับข้อมูลได้ดียิ่งขึ้นและส่งผลต่อความน่าจะเป็นของการตอบถูก ตามแนวของ Multidimensional Item Response Theory (MIRT). ทั้งทฤษฎีการทดสอบแบบดั้งเดิมทฤษฎีการทดสอบ

แนวใหม่ ต่างก็ต้องการวัดที่เน้นผลผลิต และพฤติกรรมที่เกิดจากการนำความรู้และการนำทักษะไปใช้ในบริบทหรือสถานการณ์ที่เป็นชีวิตประจำวันจริงและด้วยการใช้สถานการณ์ที่เลียนแบบชีวิตจริง

KEYWORDS: Classical Test Theory, Modern Test Theory

**พัฒนาการทฤษฎีการทดสอบ**

พัฒนาการของทฤษฎีการทดสอบนั้นมีความสัมพันธ์กันอย่างมากกับศาสตร์ทางด้านจิตวิทยา โดยเริ่มต้นจากนักจิตวิทยาเองที่ต้องการทดสอลและหาวิธีการแก้ปัญหาหารวัดต่างๆ เพื่อพัฒนาจนเป็นศาสตร์การวัดและตรวจสอบที่มีความน่าเชื่อถืออย่างที่เราพบเห็นและใช้กันในปัจจุบัน

ศาสตร์ทางด้านการวัดและประเมินผลนั้นมีมายาวนาน แต่ตอนปลายศตวรรษที่ 19 ที่นักวิทยาศาสตร์และนักสถิติได้ให้ความสำคัญ มีการพัฒนาทฤษฎีการทดสอบอย่างมากและมีความรวดเร็วและกว้างขวาง รวมทั้งมีความเป็นระบบมาก เนื่องมาจากการรวมตัวกันของนักวัดผลเพื่อแลกเปลี่ยนความคิดผลงานวิชาการกันจนทำให้เกิดเป็นแนวคิดใหม่ๆ ทฤษฎีการทดสอบอย่างมากมายจนเป็นที่ยอมรับและสนใจอย่างกว้างขวาง

**ทฤษฎีการทดสอบแบบดั้งเดิม (CTT)** ทฤษฎีการทดสอบแบบดั้งเดิมเป็นทฤษฎีที่ถือปฏิบัติใช้กันมานานแล้ว ส่วนใหญ่จะรู้จักกันในชื่อ *ทฤษฎีคะแนนจริงแบบดั้งเดิม* ซึ่งองค์ความรู้ทั่วไปจะเน้นเกี่ยวกับการตรวจสอบความสัมพันธ์ระหว่างคะแนนที่สังเกตได้กับคะแนนจริง รวมทั้งวิเคราะห์คุณภาพโดยส่วนรวมของข้อสอบและแบบสแบสำหรับผู้สอบเฉพาะกลุ่ม

ทฤษฎีคะแนนจริงแบบดั้งเดิมนี้จะตั้งอยู่บนพื้นฐานของ โมเดลการวัดและข้อตกลงเบื้องต้นที่สำคัญหลายประการ โดยทฤษฎีนี้หลักๆ จะเน้นเพื่อการพัฒนาหลักการสร้างข้อสอบ การวิเคราะห์คุณภาพของแบบสอบ และการวิเคราะห์คุณภาพของข้อสอบ โมเดลคะแนนจริงแบบดั้งเดิม (X) เกิดจากองค์ประกอบที่สังเกตไม่ได้ 2 ส่วน คือคะแนนจริง (T) และคะแนนความคลาดเคลื่อน (E) ซึ่งมีความสัมพันธ์ดังสมการ

$$X = T + E$$

ซึ่งโมเดลการวัดตั้งอยู่บนข้อตกลงเบื้องต้น (Model Assumption) จำนวน 6 ข้อ ดังนี้

**ข้อตกลงเบื้องต้นข้อที่ 1** คะแนนที่ได้จากการวัด มีความสัมพันธ์แบบเส้นตรงและเชิงบวกกับคะแนนจริงและคะแนนความคลาดเคลื่อน

**ข้อตกลงเบื้องต้นข้อที่ 2** คะแนนจริงมีสภาวะคงที่ ซึ่งเท่ากับค่าเฉลี่ยของคะแนนที่ได้จากการวัดซ้ำๆ หลายๆครั้ง

**ข้อตกลงเบื้องต้นข้อที่ 3** คะแนนความ คลาดเคลื่อนไม่มีความสัมพันธ์กับคะแนนจริง

**ข้อตกลงเบื้องต้นข้อที่ 4** ความคลาดเคลื่อนส่วนบุคคล ไม่มีความสัมพันธ์ซึ่งกันและกัน

**ข้อตกลงเบื้องต้นข้อที่ 5** การทดสอบ 2 ฉบับจะเป็นแบบทดสอบคู่ขนาน เมื่อคะแนนจริงของผู้สอบมีค่าเท่ากันทั้งสองฉบับ

**ข้อตกลงเบื้องต้นข้อที่ 6** แบบทดสอบ 2 ฉบับ จะถือว่าเป็นแบบทดสอบทัดเทียมกัน

**ความเที่ยง (Reliability)**

เมื่อใดก็ตามที่มีการนำแบบทดสอบไปใช้ แบบทดสอบจำเป็นต้องมีคุณลักษณะสำคัญอีกประการที่เรียกว่า ความเที่ยง ซึ่งความเที่ยงนี้มีอยู่หลายประเภท การประมาณค่าสัมประสิทธิ์ความเที่ยงของแบบทดสอบเองก็สามารถทำได้หลายวิธีเช่นเดียวกัน

ทฤษฎีความเที่ยงนั้นก็ได้ประยุกต์มาจากโมเดลของสเปียร์แมน โดยได้ตั้งข้อสังเกตว่า คะแนนจริงของผู้สอบกับคะแนนที่วัดได้มีความสัมพันธ์ใกล้ชิดกันเพียงใด เพื่อเป็นการบอกจึงมีค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจริงกับคะแนนที่วัดได้ หรือเรียกว่า ดัชนีความเที่ยง นั่นเอง

$$\sigma_X^2 = \sigma_T^2 + \sigma_E^2$$

การประมาณค่าตวามเที่ยงนั้นสามารถกระทำได้หลายวิธี

**1. ความเที่ยงแบบคงที่** มักจะใช้วิธีการสอบซ้ำ แล้วประมาณค่าโดยคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน

**2. ความเที่ยงแบบสมมูล** มักจะวัดเวลาเดียวกันโดยใช้แบบทดสอบแบบสมมูลแล้วประมาณค่าโดยคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน

**3. ความเที่ยงแบบคงที่และสมมูล** มักจะใช้วิธีการวัดเวลาเดียวกันแบบสอบซ้ำโดยใช้แบบ ทดสอบแบบสมมูลค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน

**4. ความเที่ยงแบบสอดคล้องภายใน** จะใช้กันอยู่หลายวิธีด้วยกัน ได้แก่ วิธีแบ่งครึ่งข้อสอบ วิธีคูเดอร์-ริชาร์ดสัน วิธีสัมประสิทธิ์แอลฟาของครอนบาค และวิธีวิเคราะห์ความแปรปรวนของฮอยท์ เป็นต้น

แต่ว่าความเที่ยงนั้นยังมีปัจจัยหลายประการที่ส่งผงต่อสัมประสิทธิ์ความเที่ยงนี้ ปัจจัยที่มีอิธผลและส่งผลสัมประสิทธิ์ความเที่ยง ได้แก่ ความเป็นเอกพันธ์ของกลุ่มผู้สอบ (Group Homogeneity) ความยาวของแบบทดสอบ (Test Length) ความสัมพันธ์ระหว่างข้อสอบ (Inter item Correlation) เวลาที้ในการทำข้อสอบ (Time Limit) และวิธีที่ใช้การประมาณค่าความเที่ยง (Methods of Estimating Reliabilities)

ปัจจัยเหล่านี้ควรนำมาพิจารณาเมื่อทำการศึกษาความเที่ยงของแบบสอบ โดยทั่วไปแล้วค่าสัมประสิทธิ์สหาสัมพันธ์จะสูงที่สุดเท่าที่จะเป็นไปได้ของแบบทดสอบนั้น

ความคลาดเคลื่อนมาตรฐานของการวัด (SEM) เป็นตัวบ่งชี้ถึงระดับการกระจายของคะแนนที่สังเกตได้จากคะแนนจริง การประมาณค่าความจริง สามารถกระทำได้โดยการประมาณเป็นช่วงด้วยการใช้ความคลาดเคลื่อนมาตรฐาน และประมาณเป็นจุดด้วยการประยุกต์การวิเคราะห์การถดถอย

**ความตรง (Validity)**

ความตรงเป็นคุณสมบัติที่สำคัญที่สุดของแบบสอบหรือของเครื่องมือวัดผล ซึ่งจะเกี่ยวข้องกับคุณภาพด้านความถูกต้องของผลที่ได้จากการวัด ทำให้เราสามารถนำคะแนนที่ได้ไปแปลความหมายถึงสิ่งที่มุ่งวัดได้อย่างเหมาะสม การตรวจสอบความตรง เป็น *“กระบวนการรวบรวมและวิเคราะห์หลักฐานเพื่อการสนับสนุนความเหมาะสมและความถูกต้องของการนำคะแนนจากเครื่องมือวัดไปสรุปอ้างอิงถึงคุณลักษณะที่มุ่งวัด”* การตรวจสอบความตรงสามารถจำแนกประเภทตามเป้าหมายสำคัญได้ 3 ประเภทหลักๆ ซึ่งได้แก่

1) การตรวจสอบความตรงตามเนื้อเรื่อง (Content validation)

2) การตรวจสอบความตรงตามเกณฑ์สัมพันธ์ (Criterion-related validation) และ

3) การตรวจสอบความตรงเชิงทฤษฎี (Construct validation)

**การตรวจสอบความตรงตามเนื้อเรื่อง (Content validation)** วิธีการตรวจสอบมักจะให้ผู้เชี่ยวชาญตรวจสอบความเหมาะสมของนิยาม และขอบเขตของเนื้อเรื่องหรือประสบการณ์ที่มุ่งวัด โดยหากต้องการตรวจสอบความสอดคล้องของข้อสอบกับจุดมุ่งหมายของการวัด มักจะใช้ IOC (Item- Objective Congruence) หากเป็นต้องการตรวจ สอบความสอดคล้องของข้อสอบกับเนื้อเรื่องที่มุ่งวัดมักจะใช้ CVR (Content Validity Ratio)

**การตรวจสอบความตรงตามเกณฑ์สัมพันธ์ (Criterion-related validation)**

ความตรงตามเกณฑ์สัมพันธ์เป็นคุณสมบัติด้านความสอดคล้องสัมพันธ์กันระหว่างคะแนนจากแบบสอบกับเกณฑ์ภายนอก (Criterion) ที่สามารถใช้วัดลักษณะที่ต้องการนั้นได้ ความตรงตามเกณฑ์สัมพันธ์สามารถจำแนกได้เป็น 2 ประเภทตามลักษณะของเกณฑ์ที่นำมาใช้หาความสัมพันธ์ของข้อสอบ ดังนี้

**1. ความตรงตามสภาพหรือความตรงร่วมสมัย (Concurrent Validity)** เป็นความตรงตามเกณฑ์ซึ่งสัมพันธ์กับสมรรถนะการดพเนินงานในสภาพปัจจุบัน โดยวิธีการประมาณค่าความตรงตามสภาพนั้นเราสามารถหาได้โดยการคำนวณจากค่าสัมประสิทธิ์สหาสัมพันธ์ระหว่างคะแนนจากแบบทดสอบกับเกณฑ์

**2. ความตรงเชิงทำนาย (Predictive Validity)** เป็นความตรงตามเกณฑ์ที่สัมพันธ์กับสมรรถนะการดำเนินงานในอนาคต การประมาณค่าความตรงเชิงทำนายของแบบทดสอบเราสามารถหาได้จากการคำนวณค่าสัมประสิทธิ์สัมพันธ์ระหว่าง คะแนนจากแบบทดสอบกับคะแนนเกณฑ์

จะเห็นได้ว่าความตรงเชิงทำนายกับความตรงตามสภาพ มีความแตกต่างกันที่ช่วงเวลาของการเก็บข้อมูลคะแนนเกณฑ์ ถ้าแบบทดสอบสามารถให้คะแนนสอดคล้องกับสภาพปัจจุบัน ถือว่ามีความตรงตามสภาพ แต่ถ้าสามารถให้คะแนนสอดคล้องกับสภาพที่เกิดขึ้นในอนาคตก็จะเป็นความตรงเชิงทำนาย ดังนั้นการมีความตรงเชิงทำนายย่อมมีความตรงเชิงสภาพ แต่ถ้าแบบสอบมีความตรงตามสภาพแล้ว ไม่จำเป็นต้องมีความตรงเชิงทำนายเสมอไป การประมาณค่าความตรงตามเกณฑ์สัมพันธีนั้นยังมีปัจจัยหลายปัจจัยที่ส่งผลต่อและเป็นข้อจำกัด เช่น ปัญหาเกี่ยวกับการเลือกเกณฑ์ ความเที่ยงขแงคะแยยสอบและคะแนนเกณฑ์และช่วงจำกัดของคะแนนสอบและคะแนนเกณฑ์

**การตรวจสอบความตรงเชิงทฤษฎี (Construct validation)** การตรวจสอบด้วยวิธีนี้จะแตกต่างจาก 2 วิธีแรก คือ การตรวจสอบความตรงตามเนื้อเรื่องและการตรวจสอบความตรงตามเกณฑ์สัมพันธ์ มุ่งที่จะทรายถึงความถูกต้องและเหมาะสมในการนำคะแนนสอบไปใช้ตามจุดมุ่งหมาย แต่ในขณะเดียวกันความตรงทั้ง 2 ประเภทนี้ต่างมีข้อจำกัดที่สำคัญหลายประการดังกล่าวมาแล้ว

ในกรณีที่ต้องการแปลงคะแนนเพื่อสรุปอ้างอิงลักษณะทั่วไปของบุคคล ซึ่งอาจเป็นลักษณะทางจิตวิทยา คุณลักษณะเหล่านี้ถือว่าเป็นโครงสร้างทางความคิด เป็นภาวะสันนิษฐาน ซึ่งเป็นนามธรรม ไม่สามารถสังเกตได้โดยตรง แบบทดสอบเหล่านี้จึงมีความจำเป็นต้องมีความตรงเชิงทฤษฎี

ความตรงเชิงทฤษฎี จัดว่าเป็นคุณสมบัติที่สำคัญที่สุดของเครื่องมือวัดลักษณะที่เป็นนามธรรม ที่ต้องทำการวัดทางอ้อม จึงจำเป็นต้องใช้การพิจารณาลักษณะในบริบทของทฤษฎี โดยอาศัย แนวคิดเชิงทฤษฎี

(Theoretical concepts) สำหรับการนิยามลักษณะที่มุ่งวัด เสนอโครงสร้างการวัด แลพกำหนดแนวทางการตั้งสมมติฐาน

การตรวจสอบความตรงเชิงทฤษฎี เป็นกระบวนการศึกษาขอบเขตความเหมาะสมของการนำคะแนนจากแบบทดสอบไปแปลผลถึงลักษณะทางจิตวิทยาที่มุ่งวัดได้ดีเพียงใด ดังนั้นการตรวจสอบความตรงเชิงทฤษฎียังคงไม่สามารถยืนยันและสรุปได้จากการศึกษาผลการทำนายเพียงครั้งเดียว แต่ต้องการผลการตรวจสอบในแง่ต่างๆ ตามโครงสร้างทฤษฎีของลักษณะที่มุ่งวัดนั้นจนได้หลักฐานเชิงประจักษ์ที่ชี้ชัดและเชื่อถือได้อย่างคงเส้นคงวา

วิธีการตรวจสอบความตรงเชิงทฤษฎี จำเป็นต้องอาศัยหลักฐานความสอดคล้องกับทฤษฎี สำหรับวิธีการหาหลักฐานเชิงประจักษ์เพื่อสนับสนุนความตรงเชิงทฤษฎีหลายวิธี วิธีที่นิยมใช้กันมาก ได้แก่ วิธีตัดสินโดยผู้เชี่ยวชาญ วิธีเปรียบเทียบคะแนนระหว่างกลุ่มที่ทราบ วิธีการเปรียบเทียบคะแนนจากการทดลอง วิธีการวิเคราะห์เมทริกซ์พหุลักษณะ-พหุวิธี วิธีวิเคราะห์ตัวประกอบ

**การประยุกต์ใช้ Confirmatory Factor Model: CFM สำหรับตรวจสอบความตรงตามเนื้อเรื่องและความเท่าเทียมกันของเนื้อเรื่อง**

เนื่องด้วยการตรวจสอบความตรงตามเนื้อเรื่องมักประสบปัญหาหลากหลายอย่าง ในการสร้างแบบสอบคู่ขนาน จำเป็นที่แบบสอบเหล่านั้นจะต้องมีความตรงจามเนื้อเรื่อง และความเท่าเทียมกันของเนื้อเรื่อง ซึ่งสามารถประยุกต์โมเดล CFM มาช่วยในการตรวจสอบได้ตามขั้นตอนดังนี้

1) ตรวจสอบความตรงตามโครงสร้างของเนื้อเรื่องที่มุ่งวัดของแบบทดสอบแต่ละฉบับ

2) ตรวจสอบความเท่าเทียมกันของเนื้อเรื่องระหว่างแบบสอบ

**ก่อนเกิดเป็นทฤษฎีการทดสอบแนวใหม่**

ทฤษฎีการทดสอบแบบดั้งเดิม Classical Test Theory (CTT) ถูกใช้เป็นแนวคิดพื้นฐานสำหรับการพัฒนาแบบทดสอบทางการศึกษาและจิตวิทยาเป็นส่วนใหญ่ ตลอดช่วงศตวรรษที่ 20 ที่ผ่านมา ในปัจจุบันก็ยังมีการใช้กันอย่างกว้างขวาง แต่ด้วยข้อจำกัดหลายประการของการวัดตามทฤษฎีการทดสอบแบบดั้งเดิม ทำให้มีนักทฤษฎีทางการทดสอบหลายท่านได้ก่อกระแสการวัดคุณลักษณะภายในแนวใหม่ เพื่อที่จะแก้ไขจุดอ่อนเกี่ยวกับข้อตกลงเบื้องต้นของทฤษฎีการทดสอบแบบดั้งเดิม เริ่มต้นด้วยเทอร์สโตน (L.L. Thurston, 1927, 1928) เป็นผู้บุกเบิกการวัดคุณลักษณะภายในบุคคล และพัฒนาเทคนิคการวิเคราะห์ตัวประกอบสำหรับศึกษาคุณลักษณะทางจิตวิทยา อันเป็นการวางรากฐานความคิดที่สำคัญเกี่ยวกับ ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theory: MTT) จากผลงานของ ครอนบาคและคณะ (Cronbach et.al., 1963, 1972) ได้เขียนบทความเรื่อง “The Dependability of Behavioral Measurement: Theory of Generalizability of Score and Profiles” ได้เสนอแนวคิดเกี่ยวกับโมเดลความเที่ยงทั่วไปภายใต้เงื่อนไขแบบต่างๆของการทดสอบ รวมทั้งผลงานของ ลอร์ดและโนวิค (Load and Novice, 1968) เรื่อง “Statistical Theories of Mental Test Score” ซึ่งได้นำเสนอหลักการและแนวคิดสำคัญที่ช่วยผลักดันกระแสการปฏิรูประบบความคิดของการวัดสู่ทฤษฎีการทดสอบแนวใหม่ รวมทั้งการมีส่วนร่วมของนักทฤษฎีการทดสอบในยุคใหม่อีกหลาย

ท่าน จนได้พัฒนาการทดสอบแนวใหม่ซึ่งสามารถจำแนกออกเป็น 2 แนวทางสำคัญๆ ได้แก่ ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของการวัด (Generalizability Theory) หรือ G-Theory และทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) หรือ IRT แต่ละแนวทางมีแนวความคิดและความเป็นมาที่สำคัญดังนี้

**ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory)**

ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) หรือ G-Study ได้มีการเสนอโมเดลของการศึกษาวิเคราะห์ความน่าเชื่อถือของผลการวัด หรือความเที่ยงในสถานการณ์หรือเงื่อนไขต่างๆ ของการวัด ทำให้เราทราบและสามารถควบคุมแหล่งความคลาดเคลื่อนได้ตรงประเด็น เพื่อผลการวัดมีความน่าเชื่อถือ หรือมีความเที่ยงสูงถึงระดับที่ต้องการ

จากแนวคิดทฤษฎีการทดสอบแบบดั้งเดิมมีความคิดพื้นฐานที่เชื่อว่า ความผันแปรของคะแนนที่สังเกตได้ประกอบด้วย 2 ส่วน คือ ความผันแปรของคะแนนจริง และความผันแปรของคะแนนความคลาดเคลื่อน ซึ่งไม่สามารถระบุหรือแบ่งแยกออกได้ นอกจากนี้ยังไม่ได้สนใจต่อเงื่อนไขการวัดและปฏิสัมพันธ์ระหว่างเงื่อนไขของการวัดที่สามรถส่งผลต่อความน่าเชื่อถือของการวัด ในขณะที่ G-Theory ได้เสนอแนวคิดในการแยกส่วนความคลาดเคลื่อน (Error) จากหลายแหล่ง (Multiple error sources) ประกอบด้วยความคลาดเคลื่อนอย่างเป็นระบบ (Systematic source) และความคลาดเคลื่อนแบบสุ่ม (Random source)

**แนวความคิดและข้อตกลงเบื้องต้น**

G-Theory เป็นทฤษฎีทางสถิติของการวิเคราะห์ความน่าเชื่อถือของผลการวัดในสถานการณ์ของการวัดผลลักษณะต่างๆ ที่เป็นเป้าหมายของการนำเครื่องมือไปใช้ ความถูกต้องความน่าเชื่อถือของการวัด คือ ความถูกต้องของการสรุปอ้างอิง (Generalization) จากคะแนนที่สังเกตได้ไปยังคะแนนจริงของบุคคล

การศึกษาความน่าเชื่อถือของผลการวัดตาม G-Theory ตั้งอยู่บนพื้นฐานของข้อตกลงเบื้องต้นดังนี้

1. คุณลักษณะที่มุ่งวัดของบุคคล เป็นค่าที่อยู่ในสภาวะคงที่ (Steady state)

2. ผู้สอบคนเดียวกันได้คะแนนต่างกันจากการวัดในแต่ละสถานการณ์ หรือเงื่อนไขของการวัดเนื่องมาจากความคลาดเคลื่อนที่เป็นระบบอย่างน้อย 1 แหล่ง

3. เมื่อพิจารณาผู้สอบทั้งกลุ่ม ความแปรปรวนที่สังเกตได้ ประกอบด้วย ความแปรปรวนของคะแนนจริง ซึ่งเป็นความแตกต่างที่แท้จริงระหว่างบุคคล

ผลการวิเคราะห์ด้วย G-Theory จะให้ค่าสัมประสิทธิ์สัมพันธ์เชิงสรุปที่แสดงถึงระดับความน่าเชื่อถือของคะแนนที่ได้จากการวัด (Level of dependability) เรียกว่า สัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของการวัด (Generalizability Coefficient) ซึ่งคล้ายกับสัมประสิทธิ์ความเที่ยง (Reliability Coefficient) ในทฤษฎีแบบดั้งเดิม

การตรวจสอบความน่าเชื่อถือของผลการวัดโดยใช้ G-Theory ทำให้ผู้บริหารการทดสอบสามารถทำการตัดสินความน่าเชื่อถือของผลการวัดได้ 2 ลักษณะ ได้แก่

1. ความน่าเชื่อถือของคะแนนในการนำผลไปใช้เชิงสัมพัทธ์ (Relative)

2. ความน่าเชื่อถือของคะแนนในการนำผลไปใช้สัมบูรณ์(Absolute)

**สัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient)**

สำหรับทฤษฎี G-Theory สัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient or $\rho_i^2$) เป็นสัดส่วนระหว่างความแปรปรวนของคะแนนเอกภพกับความแปรปรวนของค่าคาดหมายของคะแนนที่สังเกตได้ สัมประสิทธิ์การสรุปอ้างอิงมี 2 ประเภท คือ

1. สัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ ($\rho^2_{Abs}$)

2. สัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ ($\rho^2_{Rel}$)

**การประมาณค่าสัมประสิทธิ์สรุปอ้างอิงสำหรับการออกแบบการวัดที่มีฟาเซทเดียว (G-Coefficient for Single-Facet Design)**

ในการศึกษาคุณภาพของแบบทดสอบ นักวัดผลต้องการทราบว่าสถานการณ์ของการใช้แบบสอบ หรือเงื่อนไขทดสอบ 1 ปัจจัย เช่น ความยาวข้อสอบ (I) หรือ จำนวนครั้ง (O) หรือจำนวนผู้ตรวจ (R) จะมีผลต่อความน่าเชื่อถือของแบบสอบเพียงใด

POPULATION: Examinees or Person

UNIVERSE: All possible combinations of conditions of the facet.

FACET: Item or Occations

DESIGN: p x i design

**การประมาณค่าสัมประสิทธิ์สรุปอ้างอิงสำหรับการออกแบบการวัดที่มีสองฟาเซท (G-Coefficient for Two Facet Design)**

**1.) การออกแบบ TWO FACET CROSSED DESIGN**

ในการศึกษาคุณภาพของแบบทดสอบ นักวัดผลต้องการทราบว่าสถานการณ์ของการใช้แบบสอบ หรือเงื่อนไขทดสอบ 2 ปัจจัย ได้แก่ ความยาวข้อสอบ (I) และ จำนวนครั้ง (O) จะมีผลต่อความน่าเชื่อถือของแบบสอบเพียงใด

POPULATION: Examinees or Person

UNIVERSE: All possible combinations of conditions of the facet.

FACET 1: Item (I)

FACET 2: Occations (O)

DESIGN: p x i x o design

**2.) การออกแบบ TWO FACET NESTED DESIGN**

ในการศึกษาคุณภาพของแบบทดสอบ นักวัดผลต้องการทราบว่าสถานการณ์ของการใช้แบบสอบ หรือเงื่อนไขทดสอบ 2 ปัจจัย ได้แก่ ความยาวข้อสอบ (I) และ จำนวนครั้ง (O) ของการทดสอบ (เมื่อการสอบแต่ละครั้งใช้ข้อสอบต่างกัน) จะมีผลต่อความน่าเชื่อถือของแบบสอบเพียงใด

POPULATION: Examinees or Person

UNIVERSE: FACET 1: Item (I)

FACET 2: Occations (O) และ i : o

DESIGN: p x (i : o) design

**การประมาณค่าสัมประสิทธิ์สรุปอ้างอิงสำหรับการออกแบบการวัดที่มีสามฟาเซท (G-Coefficient for Three Facet Design)**

ในการศึกษาคุณภาพของแบบทดสอบ นักวัดผลต้องการทราบว่าสถานการณ์ของการใช้แบบสอบ หรือเงื่อนไขทดสอบ 1 ปัจจัย เช่น ผู้ประเมิน (R) จำนวนครั้ง (O) และแหล่งของผู้ประเมิน (R) จะมีผลต่อความน่าเชื่อถือของแบบสอบเพียงใด

**ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT)**

ทฤษฎีการตอบสนองข้อสอบ หรือ IRT สามารถจำแนกออกได้เป็น 2 ลักษณะ ได้แก่ ทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า (Dichotomous IRT) และทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous IRT)

**แนวคิดพื้นฐานของ IRT**

โมเดลการตอบสนองข้อสอบ (IRT Models) เป็นระบบความสัมพันธ์ระหว่างตัวแปรอิสระที่รวมกันสำหรับทำนายตัวแปรตาม สำหรับโมเดลการตอบสนองข้อสอบ ตัวแปรอิสระจะประกอบด้วย ตัวแปรแฝง คือ ความสามารถที่แท้จริงของผู้สอบ ($\theta$) และคุณลักษณะของผู้สอบ (B) sหรือค่าพารามิเตอร์ข้อสอบ (a, b, c) ส่วนตัวแปรอิสระเป็นตัวแปรที่สังเกตได้ คือ โอกาสการตอบข้อสอบได้ถูกต้อง

โมเดลการตอบสนองข้อสอบจะประกอบด้วยค่าพารามิเตอร์และค่าคงที่ดังนี้

1. พารามิเตอร์ของผู้สอบ

$\theta$ = ระดับความสามารถผู้สอบ โดยที่ค่า $\theta$ มีพิสัยอยู่ระหว่าง $-\infty$ ถึง $+\infty$

2. พารามิเตอร์ของข้อสอบ

$a_i$ = ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i (Discrimination parameter) ในทางทฤษฎีแล้วมีค่าระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัตินิยมใช้ $a_i$ ระหว่าง +0.50 ถึง +2.50

$b_i$ = ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i (Difficulty parameter) ในทางทฤษฎีแล้วมีค่าระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัตินิยมใช้ $b_i$ ระหว่าง -2.50 ถึง +2.50

$c_i$ = ค่าพารามิเตอร์โอกาสในการเดาข้อสอบถูก (Guessing parameter) ในทางทฤษฎีแล้วมีค่าระหว่าง 0 ถึง 1 แต่ในทางปฏิบัตินิยมใช้ $c_i$ ไม่เกิน 0.30

3. ค่าคงที่

e = 2.71828

D = 1.70

**ข้อตกลงของทฤษฎีการตอบข้อสอบ**

1. ความเป็นมิติเดียว (Unidimensionality: One trait)

2. ความเป็นอิสระ (Independence: Local Independent)

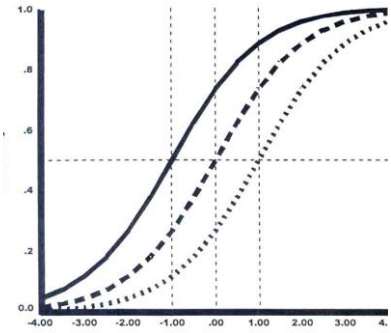
3. โค้งคุณลักษณะข้อสอบ (Item Characteristic Curve: Item Response Model)

4. ข้อสอบที่ใช้ต้องไม่เป็นข้อสอบประเภทความเร็ว

**โมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ (One parameter model or Rasch Model)**

โมเดลแรกที่มีความสำคัญที่สุดที่นำมาใช้ในโมเดลทฤษฎีการตอบข้อสอบ โค้งลักษณะข้อสอบสำหรับโมเดลโลจิสติกแบบ 1 พารามิเตอร์ แสดงในรูปฟังก์ชันทางคณิตศาสตร์ ดังนี้

$$P_i = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \text{ ; I =1,2,...,n}$$



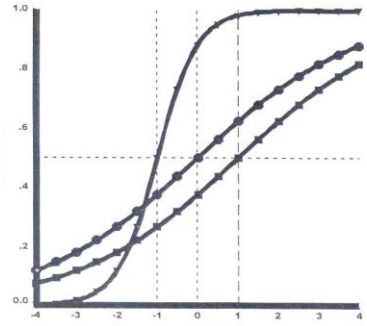

**รูปที่ 1** ลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์

โค้งลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ มีลักษณะที่สำคัญ คือ $b_i$ มีค่าแปรเปลี่ยนตามลักษณะข้อสอบแต่ละข้อ $a_i$ มีค่าคงที่ และ $c_i$ เท่ากับ 0

**โมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ (Two parameter model)**

โมเดลแรกที่มีความสำคัญที่สุดที่นำมาใช้ในโมเดลทฤษฎีการตอบข้อสอบ โค้งลักษณะข้อสอบสำหรับโมเดลโลจิสติกแบบ 2 พารามิเตอร์ แสดงในรูปฟังก์ชันทางคณิตศาสตร์ ดังนี้

$$P_i = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \text{ ; I =1,2,...,n}$$



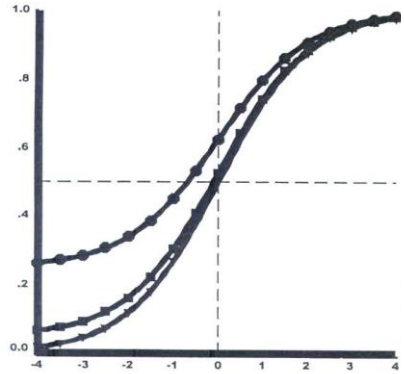

**รูปที่ 2** ลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์

โค้งลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ มีลักษณะที่สำคัญ คือ $b_i$ เป็นค่าที่แสดงตำแหน่งของ ICC ณ จุด $\theta$ ที่มีโอกาสตอบข้อสอบถูก 0.50 ส่วน $a_i$ เป็นค่าความชันของ ICC ณ จุด $b_i$ ส่วน $c_i$ เท่ากับ 0 และ e = 2.718, D = 1.7

**โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ (Three parameter model)**

โมเดลแรกที่มีความสำคัญที่สุดที่นำมาใช้ในโมเดลทฤษฎีการตอบข้อสอบ โค้งลักษณะข้อสอบสำหรับโมเดลโลจิสติกแบบ 3 พารามิเตอร์ แสดงในรูปฟังก์ชันทางคณิตศาสตร์ ดังนี้

$$P_i = \boldsymbol{C_i} + (1-c_i)\frac{e^{\boldsymbol{Da_i(\theta-b_i)}}}{\boldsymbol{1+e^{Da_i(\theta-b_i)}}} \; ; \; I = 1,2,\ldots,n$$

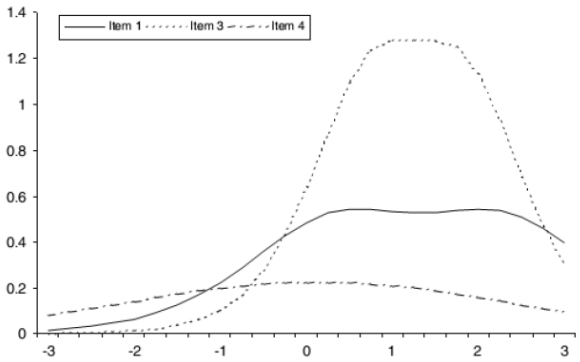



**รูปที่ 3** ลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์

โค้งลักษณะข้อสอบ (ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ มีลักษณะที่สำคัญ คือ $b_i$ เป็นค่าที่แสดงตำแหน่งของ ICC ณ จุด $\theta$ ที่มีโอกาสตอบข้อสอบถูก $\frac{1+c_i}{2}$ ส่วน $a_i$ เป็นค่าความชันของ ICC ณ จุด $b_i$ ส่วน Ci คือค่าพารามิเตอร์โอกาสการเดาข้อสอบได้ถูก e = 2.718, D = 1.7

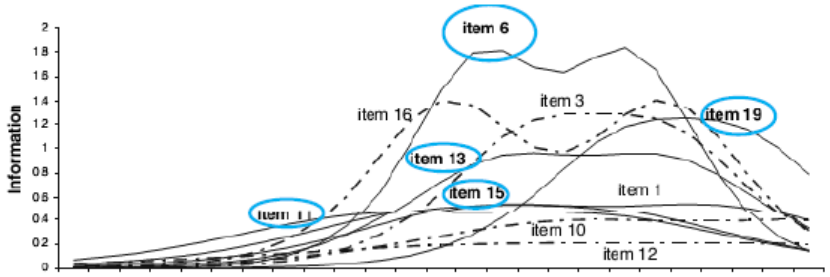
**ฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ (Item and Test Information)**

**ฟังก์ชันสารสนเทศของข้อสอบ (Item Information)** เป็นดัชนีผสมที่สร้างจากัชนีคุณลักษณะของข้อสอบหลายลักษณะ ประกอบด้วย ค่าพารามิเตอร์อำนาจจำแนก ค่าพารามิเตอร์ความยาก ค่าความแปรปรวนคะแนนรายข้อ เพื่อใช้บ่งชี้คุณภาพของข้อสอบ




**รูปที่ 4** ลักษณะกราฟฟังก์ชันสารสนเทศของข้อสอบ (Item Information)

**ฟังก์ชันสารสนเทศของแบบสอบ (Test Information)** เป็นค่าฟังก์ชันสารสนเทศของแบบสอบอันเกิดจากผลรวมเชิงพีชคณิตของฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อเข้าด้วยกันทั้งฉบับ ณ ตำแหน่ง $\theta$ เดียวกัน




**รูปที่ 5** ลักษณะกราฟฟังก์ชันสารสนเทศของแบบสอบ (Test Information)

**ความคลาดเคลื่อนมาตรฐานของการประมาณค่า (Standard Error of Estimation)**

ความคลาดเคลื่อนมาตรฐานของการประมาณค่า [SE($\theta$)] เป็นส่วนเบี่ยงเบนมาตรฐานของการแจกแจงความน่าจะเป็น ของค่าประมาณความสามารถที่แท้จริง ($\theta$) ซึ่งเป็นค่าสัดส่วนผกผันกับความถูกต้องแม่นยำของการประมาณค่าความสามารถ หรือค่าฟังก์ชันสารสนเทศของแบบสอบ

**โมเดลการตอบสนองข้อสอบแบบหลายมิติ (Multidimensional Item Response Theory)**

แนวคิดของการพัฒนาโมเดลการตอบสนองข้อสอบแบบหลายมิติ หรือ MIRT เป็นโมเดลที่มีคุณลักษณะแฝงมากกว่า 1 องค์ประกอบหรือหลายมิติ Multidimensional Item Response Theory ได้ขยายให้รองรับ $\theta$ หลายองค์ประกอบ

โมเดลการตอบสนองข้อสอบแบบหลายมิติถือว่าคุณลักษณะแฝงของบุคคลมากกว่า 1 องค์ประกอบส่งผลต่อการตอบข้อสอบ ดังนั้นพารามิเตอร์ของผู้สอบจึงมีตั้งแต่ 2 พารามิเตอร์ขึ้นไป การพิจารณาถึงความสามารถหลายมิติของบุคคล น่าจะช่วยทำให้โมเดลมีความสอดคล้องกับข้อมูลได้ดียิ่งขึ้น

จากการศึกษาเอกสารงานวิจัยที่ได้กล่าวไว้ข้างต้นแล้ว การวัดทางการศึกษาและจิตวิทยาเป็นการวัดคุณลักษณะภายในของมนุษย์ ซึ่งไม่สามารถสังเกตได้โดยตรง แต่มีความสำคัญและจำเป็นต้องศึกษา เพราะการวัดคุณลักษณะภายในจะทำให้เข้าใจการเกิดพฤติกรรมภายนอกของมนุษย์ที่สามารถสังเกตได้โดยตรง อันจะนำไปสู่การทำนายควบคุม และพัฒนาพฤติกรรมมนุษย์ การวัดคุณลักษณะภายในของมนุษย์จำเป็นต้องอาศัย ทฤษฎีการทดสอบ เพื่อทำความเข้าใจคุณลักษณะของสิ่งที่มุ่งวัด โครงสร้างของการวัด และการพัฒนาเครื่องมือสำหรับทดสอบ

ทฤษฎีการทดสอบเป็นองค์ความรู้ที่มีนัยทั่วไปเกี่ยวกับการทดสอบ วิธีการแก้ปัญหา การทดสอบและพัฒนาเครื่องมือการทดสอบ ซึ่งจะช่วยให้นักวัดผลสามารถทำการสร้างและพัฒนาแบบสอบให้มีคุณภาพ สามารถแปลความหมายผลการวัดได้อย่างถูกต้องและสามารถนำสารสนเทศไปใช้สำหรับ การตัดสินใจได้อย่างเหมาะสม ประเภททฤษฎีการทดสอบสามารถแบ่งได้เป็น 2 ประเภทหลักๆ ได้แก่ ทฤษฎีการทดสอบแบบ

ดั้งเดิม ซึ่งจะเน้นการตรวจสอบความสัมพันธ์ระหว่างคะแนนที่สังเกตได้กับคะแนนที่แท้จริง โดยการวิเคราะห์คุณภาพรวมของข้อสอบและแบบสอบ ส่วนทฤษฎีการทดสอบแนวใหม่ มุ่งเน้นการขยายแนวคิดของทฤษฎีการทดสอบแบบดั้งเดิม เพื่อให้ได้แนวทางการวัดที่น่าเชื่อถือยิ่งขึ้นซึ่งมีพัฒนาการที่สำคัญ 2 แนวทาง ได้แก่ ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) และทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT)

ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theory) จะมุ่งเน้นการขยายแนวคิดของทฤษฎีการทดสอบแบบดั้งเดิมให้มีขอบเขตที่กว้างขวาง และมีความน่าเชื่อถือมากยิ่งขึ้น โดยปรับข้อตกลงให้มีความสมเหตุสมผล ตามความจริงมากขึ้น เกี่ยวกับความคลาดเคลื่อนของการวัด ยอมให้มีความคลาดเคลื่อนจากการวัดนั้นมีความแตกต่างกันออกไปตามสถานการณ์การวัด เพื่อศึกษาความเที่ยงทั่วไป ภายใต้เงื่อนไขของการวัดตามแนวของ Generalizability Theory: G-Theory รวมทั้งศึกษาความคลาดเคลื่อนจากการวัดที่มีความแตกต่างกันไปตามลักษณะของข้อสอบและแบบสอบ ขึ้นกับระดับความสามารถของผู้สอบ และของข้อสอบและแบบสอบ ทั้งยังพยายามวัดคุณลักษณะภายในหรือความสามารถที่แท้จริงของบุคคลตามแนว Item Response Theory (IRT)

ทฤษฎีการทดสอบแนวใหม่จึงน่าจะให้ผลการวัดที่ชัดเจนตรงประเด็นมากขึ้น อันจะเป็นประโยชน์ต่อการพัฒนาข้อสอบและแบบสอบให้สามารถนำไปใช้ในสถานการณ์ต่างๆของการทดสอบได้อย่างมีประสิทธิภาพ สามารถระบุนัยทั่วไปของคะแนนจริง คุณของแบบสอบตามเงื่อนไขของการทดสอบ รวมทั้งการประเมินความสามารถที่แท้จริงของผู้สอบ และการบรรยายพฤติกรรมการตอบสนองข้อสอบของผู้สอบได้เป็นอย่างดี

13.7 สะท้อนผลในสิ่งที่ตนเองได้เรียนรู้ในรายวิชาทั้งหมด โดยมี 2 ประเด็น

(1) What students know.

ทำให้ได้เรียนรู้และศึกษาแนวความคิดเกี่ยวกับ Four Building Blocks ได้อธิบายถึง Construct Modeling จาก The four Building blocks. โดยเริ่มต้นจากการอธิบายในภาพรวมและแนวคิดหลักก่อน จากนั้นอธิบายความหมายของการวัด ว่าต้องมีการกำหนดประเภทของการสังเกต คุณลักษณะของประชากรเองส่งผลต่อคุณลักษณะของการวัดด้วย การกำหนดคุณลักษณะของกลุ่มประชากรจะต้องสอดคล้องกับความจริงที่ต้องการวัด เพื่อให้ได้กระวนการวัดที่สมบูรณ์ วัถตุประสงค์หลักของการวัด ก็เพื่อ เป็นการอธิบายเตรียมการอย่างมีเหตุผล

ได้ฝึกกระบวนการสร้างแบบประเมินทักษะควบคู่กับการเรียนทฤษฎีที่เกี่ยวข้องกับการวัดและการประเมินผลการศึกษา

(2) What students need to know.

กระบวนการหรือหลักการสร้างแบบประเมินความก้าวหน้าหรืออย่างอื่น ที่เป็นวิธีการใหม่ๆ ตอบสนองต่อสถานการณ์ที่หลากหลาย