

การพัฒนาเครื่องมือประเมินระดับสมรรถนะเพื่อการเปลี่ยนแปลงสำหรับผู้เรียนระดับ
มัธยมศึกษาตอนต้นแบบพหุมิติ โดยประยุกต์ใช้โมเดลเชิงโครงสร้างเป็นฐานการพัฒนา
Development of a Multidimensional Assessment Scale of Transformative
Competency Levels for Lower Secondary School Students Based on
Construct Modeling Approach

ชลันดา พจนานา^{1*} พัชรี จันทร์เพ็ง² นุชwana หลีอังกูร² และ ชัชวาล หนองนา¹
Chalunda Podjana^{1*}, Pucharee Junpeng², Nuchwana Luanganggoon² and Chatchawan Nongna¹

(Received: August 21, 2021; Revised: October 3, 2021; Accepted: October 3, 2021)

บทคัดย่อ

การวิจัยครั้งนี้เป็นส่วนหนึ่งของการพัฒนาเครื่องมือประเมินสมรรถนะเพื่อการเปลี่ยนแปลงผ่านแพลตฟอร์มดิจิทัลอัตโนมัติแบบเรียลไทม์ มีวัตถุประสงค์เฉพาะเพื่อพัฒนาและตรวจสอบคุณภาพเครื่องมือประเมินระดับสมรรถนะเพื่อการเปลี่ยนแปลงแบบพหุมิติ สำหรับผู้เรียนระดับมัธยมศึกษาตอนต้น การศึกษาครั้งนี้ใช้โมเดลเชิงโครงสร้างเป็นฐานการพัฒนา 4 ขั้นตอน คือ (1) การพัฒนาแผนที่โครงสร้างการประเมินสมรรถนะเพื่อการเปลี่ยนแปลง (2) พัฒนาคำถามเชิงสถานการณ์ตามแผนที่โครงสร้าง (3) การให้คะแนนผลลัพธ์การเรียนรู้ และ (4) การวิเคราะห์โมเดลเชิงโครงสร้าง โดยทั้ง 4 ขั้นตอน ดำเนินการภายใต้ระเบียบวิธีวิจัยการออกแบบ (design research) กลุ่มผู้สอบเพื่อนำมาใช้ตรวจสอบคุณภาพเครื่องมือประเมิน คือ ผู้เรียนระดับมัธยมศึกษาตอนต้น จำนวน 487 คน จากนั้นนำผลการตอบมาวิเคราะห์ข้อมูลวิเคราะห์ข้อมูลตามโมเดลการตอบสนองแบบพหุมิติของคะแนนคำตอบหลายค่าที่มีสัมประสิทธิ์แบบสุ่มจากการวัด (MRCML) ประกอบการพิจารณาร่วมกับ Wright map ผลการวิจัยมีดังนี้

(1) ผลการพัฒนาเครื่องมือประเมินระดับสมรรถนะเพื่อการเปลี่ยนแปลงสำหรับผู้เรียนระดับมัธยมศึกษาตอนต้นแบบพหุมิติ ตามแผนที่โครงสร้าง ประกอบด้วย (1) มิติการสร้างสิ่งใหม่ (2) มิติความรับผิดชอบ และ (3) มิติการอยู่ภายใต้ความขัดแย้ง ในแต่ละมิติประกอบด้วย 5 ระดับ ระดับสูงสุดคือ 5 ระดับ ขยายการคิด ไปจนถึงระดับต่ำสุด คือ ชั้นระดับตอบคำถามไม่สอดคล้องกับคำถามหรือตอบคำถามในเชิงลบ ได้พัฒนาเป็นข้อคำถามเชิงสถานการณ์แบบเลือกตอบ 5 ตัวเลือก จำนวน 30 ข้อ มีการให้คะแนนจาก 0-4 คะแนน

¹ นักศึกษาดุษฎีบัณฑิต หลักสูตรการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

² รองศาสตราจารย์ หลักสูตรการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

¹ Doctoral degree student, Educational Measurement and Evaluation Program, Faculty of Education, Khon Kaen University

² Associate Professor, Educational Measurement and Evaluation Program, Faculty of Education, Khon Kaen University

* Corresponding Author E-mail: chalundap@kkumail.com

(2) ผลการตรวจสอบคุณภาพของเครื่องมือเป็นไปตามเกณฑ์มาตรฐานการทดสอบทางการศึกษาและจิตวิทยา นั่นคือ (2.1) หลักฐานความตรง เครื่องมือประเมินมีความตรงตามหลักฐานความตรง ดังนี้ หลักฐานความตรงด้านเนื้อหาของรายการประเมินพบว่ารายการข้อคำถามครอบคลุมการประเมินสมรรถนะเพื่อการเปลี่ยนแปลง หลักฐานความตรงด้านกระบวนการตอบจากการสัมภาษณ์ผู้สอบหลังการทำแบบประเมินพบว่าผู้เรียนมีความเข้าใจคำถามตรงตามที่ผู้วิจัยตั้งคำถาม หลักฐานความตรงด้านโครงสร้างภายในและการกำหนดจุดตัด พบว่า มีความสอดคล้องกับข้อมูลเชิงประจักษ์ ($\chi^2 = 72.21, df = 5, p = .01$) เมื่อทดสอบด้วยสถิติทดสอบ Likelihood-Ratio พบว่า โมเดลพหุมิติสอดคล้องกลมกลืนกับข้อมูลมากที่สุด ($G^2 = 33741.91, AIC = 33993.91, BIC = 34080.54$) เนื่องจากมีค่า G^2, AIC และ BIC ต่ำกว่าโมเดลเอกมิติ ($G^2 = 33814.12, AIC = 34056.12, BIC = 34139.31$) (2.2) หลักฐานความเที่ยง พบว่า ค่าความเที่ยงแบบ EAP/PV ในมิติการสร้างสิ่งใหม่ มิติความรับผิดชอบ และมิติการอยู่ภายใต้ความขัดแย้ง เท่ากับ 0.80 0.83 และ 0.82 ตามลำดับ ค่าความเที่ยงแบบสอดคล้องภายในทั้งหมด เท่ากับ 0.87 และมีค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานของการวัด (SEM) ในมิติการสร้างสิ่งใหม่ มิติความรับผิดชอบและมิติการอยู่ภายใต้ความขัดแย้ง เท่ากับ 0.03, 0.03 และ 0.02 ตามลำดับ และ (2.3) ผลการวิเคราะห์ค่าสถิติความเหมาะสมรายข้อ (item fit) โดยพิจารณาค่า OUTFIT MNSQ และ INFIT MNSQ ของข้อคำถามจากการวิเคราะห์ด้วยโมเดลพหุมิติพบว่า ค่า OUTFIT MNSQ ของทั้งหมดมีค่าอยู่ระหว่าง 0.77 ถึง 1.73 ส่วนค่า INFIT MNSQ มีค่าอยู่ระหว่าง 0.86 ถึง 1.30 ข้อโดยคำถามเกือบทุกข้ออยู่ในเกณฑ์ที่ยอมรับได้

คำสำคัญ: สมรรถนะเพื่อการเปลี่ยนแปลง การพัฒนาเครื่องมือ โมเดลเชิงโครงสร้าง โมเดลการตอบสนองแบบพหุมิติของคะแนนคำตอบหลายค่าที่มีสัมประสิทธิ์แบบสุ่มจากการวัด

Abstract

This research is a part of the development of a scale for assessing transformative competencies using real-time digital automation platform. It specifically aimed to develop and validate the quality of a multidimensional assessment scale of transformative competency levels for lower secondary school students. This study used a construct modeling approach as the basis for four processes of development: (1) developing the construct maps of transformative competency assessment; (2) developing the situational questions based on the construct maps; (3) scoring the learning outcomes; and (4) analyzing the construct modeling. All of the four processes were conducted under the Design Research methodology. The group of test takers employed for the process of verifying the quality of the assessment scale consisted of 487 lower secondary school students. The data from their responses were analyzed and considered through the Multidimensional Random Coefficients Multinomial Logit Model (MRCML) and the Wright Map. The findings were as follows:

(1) The results of developing the multidimensional assessment scale of transformative competency levels for lower secondary school students based on the construct maps were of three dimensions: (1) creating new value; (2) taking responsibility; and (3) reconciling tensions and dilemmas. Each dimension consisted of five levels starting from the highest level of 5, or the level of extended thinking (Advanced Level), to the lowest one or the level of responding with inconsistent or negative answers (No-relevance Level). They were developed into 30 multiple choice-situational questions with 5 choices each and scoring scales of 0 to 4.

(2) The results of scale validation were acceptable according to the standards for educational and psychological testing. (2.1) In terms of the validity evidence, the assessment scale was consistent with the validity evidence as the content validity evidence of the assessment items indicated the list of questions covered the transformative competency assessment. The response process validity evidence from the interviews with the test takers after the assessment showed that they understood the questions exactly as the researcher asked them. The internal structure validity evidence and the determination of cut-off score were found to be conformed to the empirical data ($\chi^2 = 72.21$, $df = 5$, $p = .01$). When it was tested by the statistical test as in Likelihood-Ratio, the multidimensional model was found to be most consistent with the data ($G^2 = 33741.91$, $AIC = 33993.91$, $BIC = 34080.54$) because the values of G^2 , AIC and BIC were lower than the unidimensional model ($G^2 = 33814.12$, $AIC = 34056.12$, $BIC = 34139.31$). (2.2) For the reliability evidence, the EAP/PV reliability values in the dimensions of creating new value, taking responsibility, and reconciling tensions and dilemmas were 0.80, 0.83 and 0.82, respectively. The reliability of internal consistency of the total scale was 0.87; and the mean of the standard error of measurement (SEM) in the dimensions of creating new value, taking responsibility, and reconciling tensions and dilemmas were 0.03, 0.03, and 0.02, respectively. (2.3) The results of the item fit statistics analysis on the OUTFIT MNSQ and INFIT MNSQ values of the question items from the multidimensional model analysis indicated that the OUTFIT MNSQ of the scale were between 0.77 and 1.73, while the INFIT MNSQ ranged from 0.86 to 1.30. Almost all questions were within the acceptable criteria.

Keywords: transformative competencies; tool development; construct model;

Multidimensional Random Coefficients Multinomial Logit Model

บทนำ

สมรรถนะเพื่อการเปลี่ยนแปลงมีความจำเป็นต่อการเปลี่ยนแปลงของโลกอนาคตที่มีความซับซ้อนและมีความไม่แน่นอน ประเทศไทยจึงมีการกำหนดสมรรถนะหลักของผู้เรียนไว้ในหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน 2551 และมีปรับปรุงให้มีความทันสมัยและทันต่อความเปลี่ยนแปลงอย่างรวดเร็วของโลกปัจจุบัน ในการกำหนดกรอบการเรียนรู้สมรรถนะหลักเพื่อศึกษานำร่องในปี พ.ศ. 2562 มีจุดมุ่งหมายในการเตรียมพร้อมผู้เรียน สู่การเรียนรู้ในระดับที่สูงขึ้นและนำไปประกอบอาชีพในอนาคต (Office of the Education Council, 2019) อย่างไรก็ตามสมรรถนะหลักของผู้เรียนยังไม่มีควมเพียงพอต่อการเปลี่ยนแปลงของโลกในปัจจุบัน จึงมีความจำเป็นต้องอาศัยสมรรถนะที่เป็นตัวกลางในการขับเคลื่อนให้ผู้เรียนสามารถมีภูมิคุ้มกันภัยคุกคามและความเสี่ยงได้อย่างรู้เท่าทันการเปลี่ยนแปลง โดยการนำความรู้ ทักษะ และทัศนคติที่ผู้เรียนนำมาประยุกต์ใช้ รวมถึงพฤติกรรมต่างๆ ที่แสดงออกมามีภายใต้บทบาทหน้าที่และสถานการณ์ต่างๆ เพื่อให้ประสบความสำเร็จในการตอบสนองความท้าทายของศตวรรษที่ 21 ซึ่งจะช่วยให้ผู้เรียนสามารถที่จะช่วยกำหนดความผูกพันและความยั่งยืนสำหรับตนเอง ผู้อื่นและโลกได้

จากรายงานผลการวิจัยของโครงการ DeSeCo ในปี ค.ศ. 2017 พบว่าสมรรถนะเพื่อการเปลี่ยนแปลงเป็นเป้าหมายที่ OECD Learning Compass 2030 ได้ระบุไว้ นอกเหนือจากเดิมที่มีการกำหนดสมรรถนะพื้นฐานจากการเรียนรู้ในหลักสูตร สมรรถนะเพื่อการเปลี่ยนแปลงจะเป็นสิ่งจำเป็นต่อการเจริญรุ่งเรืองในศตวรรษที่ 21 และเป็นแง่มุมที่สำคัญของสติปัญญาของมนุษย์ (OECD 2018; European Commission, 2007; OECD Better Life Index, 2018) แบ่งออกเป็น 3 มิติ คือ (1) การสร้างสิ่งใหม่ (Creating New Value: CNV) คือ การสร้างสรรค์สิ่งต่างๆ ในรูปแบบใหม่ หรือการพัฒนาความรู้ความเข้าใจแนวคิดกลยุทธ์และวิธีการแก้ปัญหา นำไปใช้กับปัญหาทั้งเก่าและใหม่ เมื่อผู้เรียนสร้างคุณค่าใหม่ พวกเขาจะสามารถตั้งคำถามกับสถานการณ์หรือสภาพแวดล้อมที่เกิดขึ้น โดยผ่านการทำงานร่วมกันกับผู้อื่นและพยายามคิด “นอกกรอบ” (2) การอยู่ภายใต้ความขัดแย้ง (Reconciling Tensions and Dilemmas: RTD) คือ การวางตัวอยู่ภายใต้สถานการณ์ความขัดแย้งที่ผู้เรียนต้องเผชิญ และสามารถพิจารณาผลลัพธ์ของความขัดแย้งนั้นจากมุมมองระยะสั้นและระยะยาว สมรรถนะนี้ผู้เรียนจะต้องหาคำตอบได้ว่าสิ่งใดเป็นเหตุเป็นผลที่ทำให้เกิดความขัดแย้งและสามารถจัดการสิ่งเหล่านั้นให้เป็นไปในทางที่ดีต่อตนเองและบุคคลอื่นได้ และสิ่งที่ตกผลึกจากเหตุการณ์นั้นนำมาพัฒนาตนเอง และ (3) ความรับผิดชอบ (Talking Responsibility: TRE) คือ การเชื่อมโยงความสามารถในการไตร่ตรองและประเมินการกระทำของตนเองในแง่ของประสบการณ์และความรู้ ปฏิบัติถูกต้องตามหลักคุณธรรมและตามอุดมการณ์ของตน ผู้เรียนต้องเข้าใจว่าการกระทำทุกสิ่งมีผลตามมาและกระทำการต่างๆ ภายใต้ศีลธรรมและความดีงาม รวมถึงมีความรับผิดชอบต่อสิ่งที่ตามมาด้วย (OECD, 2018; Leadbeater, 2017; Canto-Sperber and Dupuy, 2001; Haste, 2001; Steinberg, 2017; Nussbaum, 1997; Grayling, 2017)

สิ่งที่กล่าวมาเป็นเพียงกรอบการเรียนรู้ของผู้เรียน แต่ยังไม่มีความชัดเจนในแนวทางการประเมินสมรรถนะเพื่อการเปลี่ยนแปลงสำหรับโลกอนาคตที่ชัดเจน รวมถึงยังไม่มีเครื่องมือที่จะช่วยวัดและประเมินผลให้ครูผู้สอนและผู้มีส่วนเกี่ยวข้องนำไปใช้ในการพัฒนาการเรียนการสอนและพัฒนาผู้เรียน ดังนั้น สิ่งที่เข้ามามีส่วนช่วยให้สามารถวัดและประเมินผลได้อย่างมีความหมาย คือ การสร้างแผนที่โครงสร้าง (construct map)

มีความสำคัญและมีประโยชน์ในเชิงของการแปลความหมาย แสดงให้เห็นถึงโครงสร้างภายในของตัวแปรแฝงหรือตัวแปรคุณลักษณะการคิดเชิงคณิตศาสตร์เป็นระดับ อะไรบ้าง พร้อมคำอธิบายการแปลความหมายในแต่ละระดับที่มีลักษณะของพฤติกรรมที่ขาดจากกันในแต่ละระดับอย่างชัดเจน โดยมีแนวคิดพื้นฐานในการพัฒนาโมเดลเชิงโครงสร้าง (construct modeling) (Wilson, 2005) จากการสร้างแบบพิมพ์ ประกอบด้วย 4 ขั้นตอนหลักที่สำคัญประกอบด้วย (1) แผนทีโครงสร้างของระดับผลลัพธ์การเรียนรู้ เพื่อกำหนดแผนที่นำทางในการวัดสิ่งที่ต้องวัด (2) การพัฒนาคำถามตามแผนที่โครงสร้างของระดับผลลัพธ์การเรียนรู้ (3) การให้คะแนนของผลลัพธ์การเรียนรู้ ซึ่งเป็นเกณฑ์สำคัญในการให้คะแนนสามารถจำแนกระดับการเรียนรู้ของผู้เรียนเพื่อให้ทราบความสามารถที่แท้จริงว่าสิ่งใดที่ผู้เรียนรู้และสิ่งใดที่ผู้เรียนจำเป็นต้องปรับปรุงและพัฒนา และ (4) การวิเคราะห์โมเดลเชิงโครงสร้างและการแปลใช้ในการประเมินถึงคุณภาพของระบบการประเมินที่พัฒนาขึ้นโดยมีการตรวจสอบจากข้อมูลเชิงประจักษ์ จำเป็นต้องอาศัยโมเดลการวัดจากทฤษฎีทางด้านจิตมิติเข้ามาช่วยในการตรวจสอบร่วมกับการใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ เป็นวิธีการตรวจสอบโดยอาศัยจุดเด่นของวิธีการวิเคราะห์องค์ประกอบและทฤษฎีการตอบสนองข้อสอบไว้ด้วยกัน โดย Adams, Wilson & Wang (1997) ได้พัฒนาโมเดลการตอบสนองแบบพหุมิติของคะแนนคำตอบหลายค่าที่มีสัมประสิทธิ์แบบสุ่มจากการวัด Multidimensional Random Coefficients Multinomial Logit Model: MRCML ซึ่งอยู่ในตระกูลราล์ชที่สามารถประยุกต์ใช้กับแบบวัดที่ตรวจให้คะแนนมากกว่า 2 ค่า (polytomous item) ได้ ซึ่งการประยุกต์ใช้เทคนิคดังกล่าวสามารถนำมาใช้หาหลักฐานความตรง และหลักฐานความเที่ยงจากการวิเคราะห์ด้วยโปรแกรมร่วมกับการใช้ Partial Credit Model (PCM) สำหรับใช้ในการวิเคราะห์ข้อสอบหรือข้อคำถามที่มีกระบวนการตอบหลายลำดับขั้นซึ่งจำเป็นต้องมีการตรวจให้คะแนนการตอบถูกต้องหรือตอบถูกต้องบางส่วนในแต่ละลำดับขั้นของกระบวนการตอบ ซึ่งเป็นโมเดลที่พัฒนาขยายต่อจาก โมเดล 1 พารามิเตอร์ และใช้หลักการคำนวณความน่าจะเป็นของการตอบแต่ละระดับขั้นการตอบโดยตรงแบบขั้นตอนเดียว (direct IRT method) เพื่อให้เหมาะสมกับการตรวจให้คะแนนแบบหลายค่า โดยกำหนดคำถามวิจัยว่า เครื่องมือประเมินสมรรถนะเพื่อการเปลี่ยนแปลงสำหรับผู้เรียนระดับชั้นมัธยมศึกษาตอนต้นโดยการสร้างข้อคำถามเชิงสถานการณ์แบบเลือกตอบแบบพหุมิติมีคุณภาพเป็นอย่างไร

ความมุ่งหมายของการวิจัย

การวิจัยนี้มีความมุ่งหมายเพื่อพัฒนาและตรวจสอบคุณภาพเครื่องมือประเมินระดับสมรรถนะเพื่อการเปลี่ยนแปลงสำหรับผู้เรียนระดับชั้นมัธยมศึกษาตอนต้นแบบพหุมิติ

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ใช้ระเบียบวิธีวิจัยการออกแบบ (design research) (Vongvanich, 2020; Reeves, 2006) โดยแบ่งเป็น 4 ระยะ คือ (1) สำรองและพัฒนารอบการประเมินสมรรถนะเพื่อการเปลี่ยนแปลง (2) ออกแบบและสร้างเครื่องมือต้นแบบในการประเมินสมรรถนะ (prototype) ผ่านแพลตฟอร์มดิจิทัลอัตโนมัติแบบเรียลไทม์ (3) นำนวัตกรรมต้นแบบไปทดลองใช้เพื่อนำข้อมูลที่ได้มาวิเคราะห์ ปรับปรุงพัฒนา

ระบบให้สมบูรณ์และรายงานคุณภาพของนวัตกรรม และ (4) เสนอหลักการในการพัฒนานวัตกรรมต้นแบบ โดยการนำเสนอในบทความครั้งนี้เป็นการศึกษาในระยะที่ 2 มีรายละเอียด ดังนี้

1. กลุ่มผู้สอบ

การวิจัยครั้งนี้มุ่งเน้นการพัฒนาเครื่องมือประเมินสมรรถนะเพื่อการเปลี่ยนแปลงสำหรับผู้เรียน ระดับชั้นมัธยมศึกษาตอนต้น โดยใช้โมเดลพหุมิติ (MRCML) (Adams, Wilson, & Wang, 1997) ซึ่งให้ความสำคัญกับการกำหนดขนาดกลุ่มตัวอย่างและความครอบคลุมระดับความสามารถ (θ) ให้เพียงพอครอบคลุมทุกมิติที่ทำการศึกษา มากกว่ามุ่งเน้นการสุ่มตัวอย่างให้มีความเป็นตัวแทนของประชากร (Demars, 2010) การตรวจสอบคุณภาพเครื่องมือที่มีลักษณะเป็นสถานการณ์ปลายเปิดที่มีการตรวจให้คะแนนหลายค่า ควรใช้กลุ่มผู้สอบอย่างน้อย 250-500 คน จึงจะมีความเพียงพอในการประมาณค่าพารามิเตอร์ให้มีความถูกต้อง ลดความคลาดเคลื่อนมาตรฐานของการวัดและไม่ทำให้เกิดความแปรเปลี่ยนระหว่างกลุ่มผู้สอบ และข้อสอบ (Sass, Schmitt, & Walker, 2008) ดังนั้นการศึกษาครั้งนี้ได้กำหนดผู้สอบ จำนวน 487 คน จากผู้เรียนชั้นมัธยมศึกษาตอนต้น 10 โรงเรียน ในเขตภาคตะวันออกเฉียงเหนือ และภาคกลางที่ครอบคลุมในทุกระดับสมรรถนะเพื่อการเปลี่ยนแปลง 3 ระดับ คือ ต่ำ ปานกลาง และสูง โดยพิจารณาจาก (1) ผลการประเมินผู้เรียนในระดับชาติ ความสามารถทางด้านวิชาการ ดนตรี ศิลปะ กีฬา และคุณธรรมจริยธรรม (2) ขนาดโรงเรียน และ (3) ความพร้อมของโรงเรียนในด้านการใช้อุปกรณ์เพื่อทำการประเมินแบบออนไลน์

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือประเมินสมรรถนะเพื่อการเปลี่ยนแปลง อิงกรอบการประเมินที่พัฒนามาจากการศึกษานำร่องของ Podjana, Junpeng, Luangangoon, Nongna & Tang (in press) ในงานวิจัยในระยะที่ 1 ด้วยวิธีการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง ชีวประวัติของบุคคลที่ถูกยอมรับในสังคมที่มีคุณลักษณะตรงตามนิยามของสมรรถนะเพื่อการเปลี่ยนแปลง โดยมีเกณฑ์ในการคัดเลือก คือ เป็นบุคคลที่มีชื่อเสียงในระดับโลกที่มีผลงานเด่นชัดในแต่ละมิติ จำนวน 3 คน ประกอบด้วย บุคคลที่มีชื่อเสียงเรื่องของการสร้างนวัตกรรม 1 คน บุคคลที่มีความสามารถในการปรับตัวในสังคมได้อย่างมีความสุข (well-being) 1 คน และบุคคลที่สร้างแรงบันดาลใจให้กับคนอื่นตระหนักต่อสิ่งแวดล้อม 1 คน จากนั้นจึงสัมภาษณ์ผู้เรียนระดับชั้นมัธยมศึกษาตอนต้นและตอนปลายที่คาดว่าจะเป็นผู้มีสมรรถนะเพื่อการเปลี่ยนแปลง โดยเป็นบุคคลที่เป็นที่ยอมรับในโรงเรียนที่มีผลงานเชิงประจักษ์ เช่น ประธานนักเรียน ผู้เรียนที่ได้รับทุนความเป็นเลิศทางคณิตศาสตร์และวิทยาศาสตร์ หรือผู้เรียนที่มีความสามารถทางศิลปะ ดนตรี กีฬาจากการสอบถามครูผู้สอน จำนวน 5 คน และจากการตอบข้อคำถามเชิงสถานการณ์ปลายเปิดจำนวน 15 ข้อ ของผู้เรียนระดับชั้นมัธยมศึกษาตอนต้นจำนวน 120 คน ได้เป็นกรอบการประเมินสมรรถนะเพื่อการเปลี่ยนแปลง โดยการวิเคราะห์โมเดลรากลักษณะเอกมิติ เนื่องจากเมื่อวิเคราะห์ความสอดคล้องภายในเพื่อตรวจสอบความเหมาะสมของโมเดลการวัดเบื้องต้น พบว่า ข้อมูลมีความเหมาะสมกับการวิเคราะห์แบบเอกมิติ โดยพิจารณาเปรียบเทียบระหว่างการวิเคราะห์แบบเอกมิติ และพหุมิติ จากค่าสถิติดีไวเนียนซ์ ค่า AIC และ BIC ที่มีค่าน้อยกว่าโมเดลพหุมิติ ซึ่งประกอบด้วย 3 มิติ คือ มิติการสร้างสิ่งใหม่ มิติความรับผิดชอบ และมิติการอยู่ภายใต้ความขัดแย้ง ซึ่งผู้วิจัยได้นำกรอบการประเมิน

ดังกล่าวมาเป็นฐานเพื่อพัฒนาเป็นข้อคำถามเชิงสถานการณ์ 5 ตัวเลือกแบบพหุมิติ ในการวิจัยระยะที่ 2 โดยข้อคำถามที่สร้างขึ้นเกี่ยวข้องกับสถานการณ์ในชีวิตประจำวันเพื่อให้แก้ไขปัญหาเกี่ยวข้องกับตนเอง ครอบครัว โรงเรียน และชุมชน ซึ่งการสร้างข้อคำถามในการวิจัยนี้ได้พัฒนาตามแผนที่โครงสร้างประกอบไปด้วย 5 ระดับ คือระดับสูงสุดคือ ระดับ 5 ขยายการคิด (advanced Level) ไปจนถึงระดับต่ำสุด คือ ขั้่นระดับตอบคำถามไม่สอดคล้องกับคำถามหรือตอบคำถามในเชิงลบ (no-relevance Level) ปรากฏดังภาพ 1



ภาพ 1 แผนที่โครงสร้างของระดับสมรรถนะเพื่อการเปลี่ยนแปลง

ที่มา: Podjana, et.al (in press)

3. การเก็บรวบรวมข้อมูล

การพัฒนาเครื่องมือประเมินสมรรถนะเพื่อการเปลี่ยนแปลงมีการเก็บข้อมูลทั้งเชิงปริมาณ โดยอาศัยโมเดลเชิงโครงสร้าง (construct modeling) (Wilson, 2005) เป็นฐานการพัฒนา 4 ขั้นตอน ดังนี้

3.1 การใช้กรอบการประเมินจากการพัฒนาแผนที่โครงสร้างเพื่อประเมินระดับสมรรถนะเพื่อการเปลี่ยนแปลงแบบพหุมิติ (construct maps) ผู้วิจัยได้ใช้กรอบการประเมินจากการศึกษานำร่องที่พัฒนาขึ้นจาก Podjana, et. al. (in press) ประกอบด้วย 3 มิติ เพื่อนำไปใช้ในการประเมินสมรรถนะเพื่อการเปลี่ยนแปลง นั่นคือ (1) มิติการสร้างสิ่งใหม่ (CNV) (2) มิติการรับผิชอบ (TRV) และ (3) มิติการอยู่ภายใต้ความขัดแย้ง (RTD) โดยในแต่ละมิตินี้มี 5 ระดับ ระดับต่ำสุด คือ ระดับการตอบคำถามไม่สอดคล้องกับคำถามหรือการตอบคำถามในเชิงลบ และระดับสูงสุดคือ ขั้่นขยายการคิด โดยมีการให้คะแนนตั้งแต่ 0-4 คะแนน เช่น มิติความรับผิดชอบมีระดับสูงสุด ซึ่งเป็นระดับที่ 5 คือ ผู้เรียนแก้ปัญหาของตนเองต่อเหตุการณ์ต่าง ๆ โดยยึดหลักคุณธรรมจริยธรรมและเคารพสิทธิของตนเองและผู้อื่น ซึ่งนำไปสู่การกำหนดคะแนนเท่ากับ 4 คะแนน ปรากฏดังตาราง 1

ตาราง 1 กรอบการประเมินสมรรถนะเพื่อการเปลี่ยนแปลง ในแต่ละมิติ

ระดับ (คะแนน)	ระดับสมรรถนะ	มิติการสร้างสิ่งใหม่	มิติความรับผิดชอบ	มิติการอยู่ภายใต้ความขัดแย้ง
5(4)	ระดับขยายการคิด	ผู้เรียนอธิบายเหตุผลและบอกถึงสาเหตุของสภาพปัญหาในเชิงลึกที่ต้องการแก้ไขได้อย่างรอบด้านทั้งต่อตนเองและผู้อื่น เชื่อมโยงความรู้และประสบการณ์ เพื่อคิดสร้างสิ่งใหม่ในการแก้ไขปัญหาที่เกิดขึ้น โดยแสดงวิธีที่หลากหลาย และมีความคล่องแคล่วเป็นระบบ และสามารถร่วมมือกับผู้อื่นได้เพื่อแก้ปัญหาได้อย่างเป็นรูปธรรมและมีความเป็นไปได้อย่างสร้างสรรค์	ผู้เรียนคำนึงถึงผลกระทบต่อตนเองและส่วนรวมก่อนจะกระทำสิ่งใดเสมอ มีความรับผิดชอบต่อสิ่งที่ตนเองกระทำ โดยยึดหลักคุณธรรมจริยธรรมและเคารพสิทธิของตนเองและผู้อื่นอย่างเต็มใจและพร้อมจะขอโทษเมื่อตนเองทำผิด และแก้ไขอย่างถูกต้อง	ผู้เรียนสามารถจัดการความขัดแย้งโดยวิธีการให้ความร่วมมือ ทำความเข้าใจระหว่างคู่ขัดแย้ง รับฟังความต้องการหรือสื่อสารกับผู้อื่นหรือกลุ่มอื่นอย่างจริงจัง และร่วมกันระดมความคิดหาทางออกที่สามารถตอบสนองความต้องการของทุกฝ่าย บนพื้นฐานการแก้ปัญหาอย่างสร้างสรรค์
4(3)	ระดับมียุทธวิธีในการคิด	ผู้เรียนสามารถระบุปัญหาที่เกิดขึ้นได้อย่างมีเหตุผล สามารถประยุกต์ใช้ความรู้ในการสร้างและพัฒนาสิ่งต่างๆ เพื่อแก้ปัญหาโดยวิธีที่เป็นไปได้ เป็นขั้นตอนได้อย่างเป็นรูปธรรม	- ผู้เรียนรับผิดชอบต่อการกระทำของตนเองแม้ว่าจะขัดกับสิ่งที่ตนเองคิดก็ตาม แต่ต้องทำตามสิ่งที่ถูกต้องตามบรรทัดฐานของสังคม	ผู้เรียนแก้ปัญหาความขัดแย้งโดยวิธีการประนีประนอม รับมือกับความขัดแย้งโดยใช้วิธีที่ทั้งสองฝ่ายลดความต้องการของตนเอง เพื่อให้ความขัดแย้งนั้นยุติลง แต่อาจไม่ใช้วิธีที่ดีในระยะยาว เนื่องจากการยอมเสียตัวตนบางส่วนไป อาจไม่ได้เกิดจากความเต็มใจของทั้งสองฝ่าย
3(2)	ระดับทักษะ/ความคิดรวบยอด	ผู้เรียนสามารถแก้ปัญหาได้จากวิธีการทั่วไปที่ผู้อื่นก็ทำด้วยวิธีนั้น หรือเป็นวิธีที่โดยการเลียนแบบสิ่งที่มีอยู่แล้วในการคิดสิ่งต่างๆ เพื่อแก้ปัญหาได้อย่างสอดคล้องกัน	ผู้เรียนแสดงความรับผิดชอบต่อเพราะได้รับความกดดันจากรอบข้างมากกว่าความรู้สึกว่าต้องรับผิดชอบต่อตัวเอง	- ผู้เรียนสามารถระบุสถานการณ์ปัญหาความขัดแย้งได้ และแก้ปัญหาโดยการโอนอ่อนผ่อนตามโดยไม่คำนึงถึงความรู้สึกของตนเองหรือข้อเท็จจริงที่เกิดขึ้น เพื่อให้ปัญหาความขัดแย้งหมดไป
2(1)	ระดับการรับรู้	ผู้เรียนรับรู้ปัญหาแต่ระบุวิธีการแก้ปัญหาในแนวคิดใหม่ไม่ได้หรือไม่สามารถระบุแนวทางการแก้ปัญหาได้เลย	เมื่อเกิดปัญหาจากการกระทำของตนเอง หรือเป็นผู้มีส่วนเกี่ยวข้อง ผู้เรียนจะวางเฉยต่อสิ่งนั้นและไม่แสดงความรับผิดชอบใดๆ	ผู้เรียนรับรู้สภาพปัญหาว่าเกิดความขัดแย้ง หลีกเลี่ยงการเผชิญหน้าในการแก้ปัญหา
1(0)	ระดับตอบคำถามไม่สอดคล้องกับคำถามหรือตอบคำถามในเชิงลบ	ผู้เรียนไม่ตอบคำถามหรือ ผู้เรียนเพิกเฉยต่อการหาวิธีในการแก้ปัญหา รวมถึงการใช้วิธีการทำให้เกิดผลเสียต่อตนเองและผู้อื่น	ผู้เรียนไม่ตอบคำถามหรือคำตอบไม่สอดคล้องกับคำถาม หรือแสดงการกระทำที่เป็นผลทางลบต่อสิ่งนั้น	ผู้เรียนไม่ตอบคำถาม หรือตอบคำถามไม่เกี่ยวข้อง หรือผู้เรียนเอาชนะด้วยการแข่งขันหรือกดดันผู้อื่นด้วยความคิดเห็นของตนเองแสดงการกระทำที่เป็นผลทางลบต่อสิ่งนั้น

3.2 ขั้นตอนการสร้างข้อคำถาม (task design) ในการพัฒนาข้อคำถามเชิงสถานการณ์แบบปรนัย 5 ตัวเลือก จากกรอบการประเมินที่เกิดจากการกำหนดระดับสมรรถนะเพื่อการเปลี่ยนแปลงในการวิจัยระยะที่ 1 จากการถอดบทเรียนจากอัตชีวประวัติของผู้ที่ประสบความสำเร็จ เพื่อใช้เป็นแนวทางในการกำหนดทิศทางระดับของสมรรถนะเพื่อการเปลี่ยนแปลง จำนวน 3 คน และการสัมภาษณ์โดยใช้เทคนิคเหตุการณ์วิกฤติ

(critical incident interview) ซึ่งเป็นการสร้างสถานการณ์ให้ผู้ถูกสัมภาษณ์แก้ปัญหาจากเหตุการณ์ต่าง ๆ จากคัดเลือกบุคคลต้นแบบที่เป็นที่ยอมรับว่ามีสมรรถนะเพื่อการเปลี่ยนแปลงในโลกยุคใหม่ตามที่กำหนดเบื้องต้นในเอกสารและงานวิจัยที่เกี่ยวข้อง โดยมีเกณฑ์ในการคัดเลือก คือ เป็นบุคคลที่เป็นที่ยอมรับในโรงเรียนที่มีผลงานเชิงประจักษ์ เช่น ประธานนักเรียน ผู้เรียนที่ได้รับทุนความเป็นเลิศทางคณิตศาสตร์และวิทยาศาสตร์ ความสามารถทางศิลปะ ดนตรี และกีฬา จำนวน 5 คน เพื่อกำหนดระดับสมรรถนะในระดับสูงในแต่ละมิติ และยังขาดคำอธิบายสมรรถนะในระดับที่ต่ำลงมา จึงปรับปรุงคำถามจากการสัมภาษณ์ไปสู่การสร้างเป็นข้อคำถามเชิงสถานการณ์ปลายเปิด 15 ข้อ เพื่อนำไปใช้กับผู้เรียนระดับชั้นมัธยมศึกษาตอนต้น จำนวน 120 คน เพื่อตรวจสอบคุณภาพของกรอบการประเมินสมรรถนะเพื่อการเปลี่ยนแปลง โดยใช้โมเดลรากลัส พบว่ากรอบการประเมินที่เหมาะสมกับสมรรถนะเพื่อการเปลี่ยนแปลง ประกอบด้วย 3 มิติแต่ละมิติ ประกอบด้วย 4 จุดตัด ทำให้สามารถกำหนดกรอบการประเมินได้ 5 ระดับ ดังภาพ 1 และมีคำอธิบายในแต่ละมิติดังตาราง 1

จากกรอบการประเมินที่ใช้ที่สร้างขึ้นดังกล่าวนำไปสู่การพัฒนาเครื่องมือในระยะที่ 2 ที่ผู้วิจัยนำเสนอผลการวิจัยในบทความนี้ ในการนำเข้าแพลตฟอร์มดิจิทัลอัตโนมัติแบบเรียลไทม์ ผู้วิจัยจึงพัฒนาข้อคำถามให้มีการตรวจให้คะแนนที่ง่ายต่อการรายงานผลแบบทันทีหลังจากทำการประเมิน โดยการปรับข้อคำถามให้เป็นปรนัยแบบเลือกตอบ โดยมีขั้นตอนการพัฒนาข้อคำถามเชิงสถานการณ์แบบเลือกตอบดังนี้

(1) สร้างข้อคำถามให้ตรงกับนิยามของแต่ละมิติ และกำหนดตัวเลือกให้สอดคล้องกับแผนผังที่โครงสร้างในแต่ละมิติเพื่อให้ได้ข้อมูลตามที่บรรยายไว้ในแผนผังที่โครงสร้าง

(2) กำหนดขอบเขตเนื้อหาและพฤติกรรม (มิติ) ที่ต้องการวัด เป็นการกำหนดน้ำหนักของพฤติกรรมที่ต้องการวัด ในการที่จะกำหนดว่าในแต่ละเนื้อหาต้องการที่จะให้เกิดพฤติกรรมอะไรบ้างของสมรรถนะเพื่อการเปลี่ยนแปลง โดยผู้วิจัยสร้างสถานการณ์ และกำหนดตัวเลือกให้มีระดับสูงสุดและต่ำสุดทุกข้อ โดยแต่ละข้อเป็นข้อคำถามวัดสถานการณ์ จำนวน 5 ตัวเลือก รวมถึงการกำหนดตัวเลือก “อื่น ๆ” หากความคิดเห็นหรือระดับการปฏิบัติของผู้สอบไม่ตรงตามตัวเลือก เพื่อพิจารณาคำตอบที่มีความเป็นไปได้เพิ่มเติม

(3) กำหนดตารางโครงสร้างและสร้างเครื่องมือประเมินสมรรถนะเพื่อการเปลี่ยนแปลงสถานการณ์ให้ครอบคลุมทุกมิติ แต่ละข้อจะวัดเพียงมิติใดมิติหนึ่งเท่านั้น ตัวอย่างข้อคำถามปรากฏดังภาพ 2 เป็นการวัดมิติความรับผิดชอบ โดยเป็นการวัดได้สูงสุดถึงระดับที่ 5 มีลักษณะเป็นสถานการณ์ที่เน้นการแก้ปัญหาและการปฏิบัติตนเมื่ออยู่ร่วมกับผู้อื่น

3.3 การกำหนดแนวทางการให้คะแนน (outcome space) ในแต่ละระดับ โดยมีขั้นตอน ดังนี้

(1) กำหนดเกณฑ์การให้คะแนนเพื่อเป็นแนวทางในการตรวจคะแนนตามแผนผังที่โครงสร้าง โดยมีการให้คะแนนในแต่ละในแต่ละมิติอยู่ระหว่าง 0-4 คะแนน ซึ่งเป็นการให้คะแนนแบบความรู้บางส่วน

มิติที่วัด:	ด้านความรับผิดชอบ	ระดับการวัด:	ระดับ 5
คำชี้แจง:	ให้ผู้เรียนพิจารณาสถานการณ์ในแต่ละข้อและตอบคำถาม โดยเลือกข้อที่ตรงกับความรู้สึกและปฏิบัติมากที่สุดเพียง 1 ข้อ แต่ถ้าตัวเลือกไม่ตรงกับความรู้สึกและจะปฏิบัติ ให้ผู้เรียนเลือกตัวเลือก "อื่นๆ" และเขียนคำอธิบายในช่องที่เตรียมไว้		
คำถาม:	หากเราเจอเพื่อนที่ซาร์จแบตเตอรี่โทรศัพท์มือถือของเพื่อนไปใช้ และลืมนำไปคืนโดยไม่ได้ตั้งใจ เมื่อผ่านมาเดือนหนึ่งเพิ่งสังเกตเห็นว่าที่ซาร์จแบตยังอยู่ที่ตนเอง ผู้เรียนจะอย่างไร		
	<ol style="list-style-type: none"> 1. ไม่กล้าเอาไปคืน เพราะมันนานมาแล้วกลัวเพื่อนด่า 2. เอาไปคืนไว้ในกระเป๋าเพื่อนแบบเนียนๆ 3. เอาไปคืนเพื่อนตรงๆ เพราะคิดว่าเพื่อนก็จำเป็นต้องใช้ 4. เอาไปคืนเพื่อนและขอโทษที่หยิบไปโดยไม่ได้ออก 5. เอาไปคืนเพื่อนและปรับเปลี่ยนตนเองใหม่ว่าต้องเตรียมของจำเป็นของตนเองอยู่เสมอจะได้ไม่ทำให้คนอื่นเดือดร้อนอีก 		
	อื่นๆ.....		

ภาพ 2 ตัวอย่างคำถามมิติด้านความรับผิดชอบ

(2) เครื่องมือที่ใช้กับผู้สอบจำนวน 487 คน โดยการเก็บข้อมูลในทุก ๆ ชั้นเรียน ครูผู้สอนจะชี้แจงกับผู้เรียนว่าผลการตอบไม่มีผลต่อคะแนนหรือเกรดเฉลี่ย จึงให้ตอบตามความคิดเห็นหรือการปฏิบัติให้ตรงกับตนเองมากที่สุด และดำเนินการตรวจให้คะแนน ตามเกณฑ์การให้คะแนนที่กำหนดข้างต้นในแต่มิติ ดังตัวอย่างการให้คะแนนในมิติด้านการรับผิดชอบต่อข้อคำถามในภาพที่ประกอบ 2 ซึ่งมีการแบ่งระดับตั้งแต่ระดับที่ 1-5 โดยมีการให้คะแนนตั้งแต่ 0-4 โดยข้อคำถามดังกล่าว ผู้สอบจะได้คะแนนสูงสุด 4 คะแนนและอยู่ในระดับที่ 5 (ขั้นขยายการคิด) คือ ผู้สอบเลือกตอบตัวเลือกที่ 5 “เอาไปคืนเพื่อนและปรับเปลี่ยนตนเองใหม่ว่าต้องเตรียมของจำเป็นของตนเองอยู่เสมอจะได้ไม่ทำให้คนอื่นเดือดร้อนอีก” เป็นตัวเลือกที่สะท้อนให้เห็นว่าการแก้ปัญหาจะใช้วิธีการคำนึงถึงผลกระทบต่อตนเองและส่วนรวมก่อนจะกระทำการใดเสมอ ผู้เรียนรับผิดชอบต่อสิ่งที่ตนเองกระทำ โดยยึดหลักคุณธรรมจริยธรรมและเคารพสิทธิของตนเองและผู้อื่นอย่างเต็มใจและพร้อมจะขอโทษเมื่อตนเองทำผิด และแก้ไขอย่างถูกต้อง

3.4 การออกแบบโมเดลการวัดสมรรถนะเพื่อการเปลี่ยนแปลง ที่แสดงถึงความสัมพันธ์ของตัวแปรแฝงในแต่ละมิติ กับข้อคำถามโดยมีขั้นตอนดังนี้

(1) ตรวจให้คะแนนผู้สอบรายบุคคล ในแต่ละข้อของแต่ละมิติตามโมเดลพหุมิติระหว่างข้อคำถาม (between-items multidimensional model)

(2) นำคำตอบในแต่ละมิติไปสู่การวิเคราะห์ร่วมกับการพิจารณา wright map ที่แสดงถึงลักษณะการกระจายความสามารถที่แท้จริงของผู้สอบที่มีความสัมพันธ์กับความยากของข้อสอบแต่ละข้อตามโมเดลการวัดที่กำหนดพร้อมทั้งเป็นดัชนีบ่งชี้คุณภาพของกรอบการประเมินที่พัฒนาขึ้น

4. การวิเคราะห์ข้อมูล

สำหรับการวิจัยในครั้งนี้ ผู้วิจัยทำการวิเคราะห์ข้อมูล เพื่อตรวจสอบคุณภาพกรอบการประเมิน โดยใช้โปรแกรม ConQuest 2.0 (Wu, Adams, Wilson, & Haldane, 2007) ในการหาหลักฐานด้านความตรงและหลักฐานความเที่ยง รวมถึงการตรวจสอบความเหมาะสมรายข้อคำถาม โดยมีขั้นตอนการวิเคราะห์ข้อมูล ดังนี้

4.1 หลักฐานความตรง (AERA, APA, & NCME, 2014) ประกอบด้วย (1) หลักฐานความตรงด้านเนื้อหาของรายการประเมิน (validity evidence based on test content) โดยพิจารณาจากการกระจาย

ความสามารถ และรายการประเมินบน Wright Map (2) หลักฐานความตรงของกระบวนการตอบ (validity evidence based on response processes) จากการสัมภาษณ์ผู้สอบหลังจากใช้เครื่องมือประเมินว่ามีความเข้าใจในข้อคำถามได้ตรงตามที่ผู้วิจัยต้องการถามหรือไม่ (3) หลักฐานความตรงเชิงโครงสร้างภายใน (validity evidence based on internal structure) เพื่อเปรียบเทียบว่าโมเดลการวัดที่ผู้วิจัยพัฒนาขึ้นนี้เหมาะสมกับโมเดลแบบพหุมิติหรือไม่ พิจารณาจากการประเมินระหว่างโมเดลการประเมินระดับสมรรถนะเพื่อการเปลี่ยนแปลงแบบพหุมิตินั้นกับโมเดลแบบเอกมิติรวม โดยเปรียบเทียบจากค่าสถิติดีวีเยนซ์ (Deviance Statistic; G^2) ค่า AIC และ BIC ที่มีค่าน้อยกว่าจะมีโครงสร้างเหมาะสมกว่า (Yao & Schwarz, 2006) และ (4) คำนวณจุดตัดร่วมกับการพิจารณาการกำหนดเกณฑ์พื้นที่บน wright map แต่ละระดับของสมรรถนะเพื่อการเปลี่ยนแปลง ซึ่งพิจารณาจากค่าเฉลี่ยของความยากหรือระดับ threshold เดียวกัน หากด้วยจำนวนข้อคำถามที่นำมาใช้คำนวณ แยกตามมิติที่ทำการศึกษาคำนวณจุดตัดในแต่ละมิติเพื่อนำไปใช้กำหนดมาตรฐานการประเมินระดับสมรรถนะเพื่อการเปลี่ยนแปลง

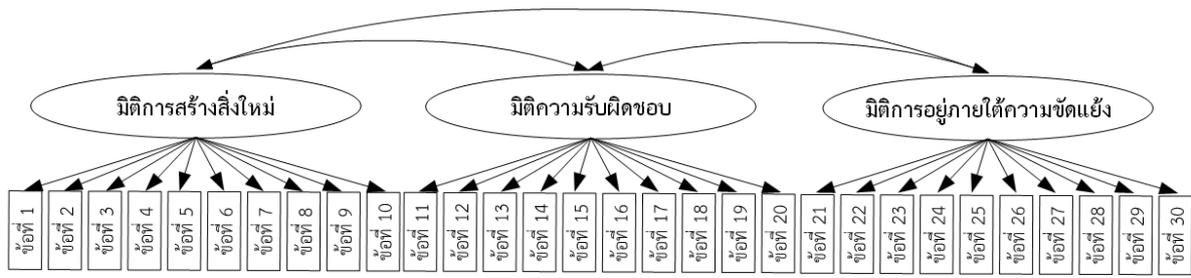
4.2 หลักฐานความเที่ยง พิจารณาจากค่าความเที่ยงแบบ EAP/PV และความเที่ยงแบบสอดคล้องภายในจากสัมประสิทธิ์แอลฟาของครอนบาค และการพิจารณาความคลาดเคลื่อนมาตรฐานของการวัดและขอบเขตเกณฑ์เมื่อสมรรถนะเพื่อการเปลี่ยนแปลงแยกเป็นมิติน้อยที่สัมพันธ์กัน

4.3 การตรวจสอบความเหมาะสมรายข้อ (Item Fit) เป็นการวิเคราะห์ค่าสถิติความเหมาะสมรายข้อผ่าน Wright Map บนพื้นฐานการศึกษาโมเดลพหุมิติ โดยพิจารณาจากดัชนีวัดความเหมาะสมของข้อคำถามรายข้อที่ยังไม่ได้แปลงค่า คือ ค่าสถิติ OUTFIT MNSQ (OUTFIT Mean Square หรือ Unweighted Mean Square) ซึ่งเป็นค่ากำลังสองของความคลาดเคลื่อนที่ไม่มีการถ่วงน้ำหนัก (Unweighted Fit) และค่าสถิติ INFIT MNSQ (INFIT Mean Square หรือ Weighted Fit) ซึ่งเกณฑ์ในการพิจารณาความเหมาะสมของค่า OUTFIT MNSQ และ INFIT MNSQ ควรมีค่าอยู่ระหว่าง 0.75 – 1.33 (Adam & Khoo, 1996; Wilson, Allen & Li, 2006)

ผลการวิจัย

1. ผลการพัฒนาเครื่องมือประเมินระดับสมรรถนะเพื่อการเปลี่ยนแปลงสำหรับผู้เรียนระดับมัธยมศึกษาตอนต้นแบบพหุมิติ

จากการพัฒนาเครื่องมือประเมินสมรรถนะเพื่อการเปลี่ยนแปลงสำหรับผู้เรียนระดับชั้นมัธยมศึกษาตอนต้นจากกรอบการประเมินสมรรถนะของ Podjana et al. (in press) สามารถพัฒนาเป็นข้อคำถามเชิงสถานการณ์แบบเลือกตอบจำนวน 5 ตัวเลือก จำนวน 30 ข้อ ประกอบด้วย 3 มิติ มิติละ 10 ข้อ คือ มิติการสร้างสิ่งใหม่ ข้อ 1-10 มิติความรับผิดชอบ ข้อที่ 11-20 และมิติการอยู่ภายใต้ความขัดแย้ง ข้อที่ 21-30 สมรรถนะเพื่อการเปลี่ยนแปลงนั้นมีความซับซ้อนของโครงสร้าง ซึ่งเมื่อพิจารณาคุณลักษณะที่ปรากฏในแต่ละระดับต้องอาศัยความรู้ทักษะ และเจตคติ เพื่อที่จะอยู่ในระดับสมรรถนะที่สูงได้ ผู้วิจัยจึงกำหนดโมเดลการวัดเป็นโมเดลพหุมิติระหว่างข้อสอบ (Between-item Multidimensional Model) ดังปรากฏในภาพประกอบ 3 ซึ่งการสร้างข้อคำถามในการวิจัยนี้ได้พัฒนาตามแผนที่โครงสร้างประกอบไปด้วย 5 ระดับ คือระดับสูงสุดคือ ระดับ 5 ขึ้นขยายการคิด ไปจนถึงระดับต่ำสุด คือ ชั้นระดับตอบคำถามไม่สอดคล้องกับคำถามหรือตอบคำถามในเชิงลบ

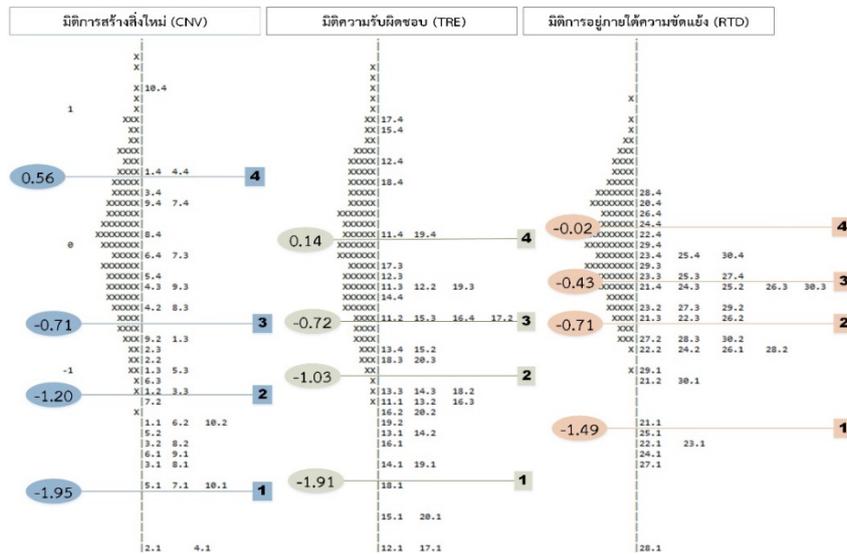


ภาพ 3 โมเดลการวัดระดับสมรรถนะเพื่อการเปลี่ยนแปลง

2. ผลการตรวจสอบคุณภาพเครื่องมือประเมินระดับสมรรถนะเพื่อการเปลี่ยนแปลงสำหรับผู้เรียนระดับชั้นมัธยมศึกษาตอนต้นแบบพหุมิติ

2.1 หลักฐานความตรง

โมเดล MRCML สามารถนำมาใช้พิจารณาหลักฐานความตรงด้านเนื้อหาของรายการประเมินพบว่ากรอบการประเมินมีความเป็นตัวแทนในรายการประเมินในแต่ละมิติเพียงพอสำหรับการประเมิน โดยมีการกระจายของค่าพารามิเตอร์ระดับสมรรถนะของผู้สอบ (θ) ครอบคลุมรายการประเมินจากระดับสูงสุดมาสู่ระดับต่ำสุด อย่างไรก็ตาม จะเห็นได้ว่าขั้นการตอบใน Threshold 1 ในแต่ละมิตียังไม่มีผู้สอบที่เป็นตัวแทนในระดับดังกล่าว และมีการกระจายของค่าพารามิเตอร์สมรรถนะของผู้สอบ (θ) ระหว่าง logit -2 ถึง +2 ดังภาพ 4



ภาพ 4 การกระจายความสามารถของผู้สอบโดยพิจารณาจาก Wright Map

หลักฐานความตรงด้านกระบวนการตอบ ผู้วิจัยได้สัมภาษณ์ผู้สอบภายหลังจากทำเครื่องมือประเมินสมรรถนะเพื่อการเปลี่ยนแปลง จำนวน 62 คน พบว่าผู้สอบเข้าใจโจทย์หรือคำถามตรงตามที่ผู้วิจัยคาดหวัง

หลักฐานความตรงด้านโครงสร้างภายในจากความสอดคล้องกลมกลืนของโมเดลการวัดประกอบกับการพิจารณา Wright Map ด้วยวิธีการวิเคราะห์พหุมิติ พิจารณาเปรียบเทียบระหว่างโมเดลแบบเอกมิติและโมเดลพหุมิติจากค่าสถิติดีเวียนซ์ (Deviance Statistic; G^2) ค่า AIC และ BIC (Yao & Schwarz, 2006) ผลปรากฏดังตาราง 2

ตาราง 2 ผลการเปรียบเทียบความสอดคล้องกลมกลืนระหว่างโมเดลพหุมิติและเอกมิติ

โมเดลสมรรถนะเพื่อการเปลี่ยนแปลง	สถิติดีเวียนซ์	จำนวนพารามิเตอร์	ค่าเกณฑ์สารสนเทศเอไคคี (AIC)	ค่าเกณฑ์สารสนเทศเบเซียน (BIC)
โมเดลสมรรถนะเพื่อการเปลี่ยนแปลงแบบเอกมิติรวม	33814.12	121	34056.12	34139.31
โมเดลสมรรถนะเพื่อการเปลี่ยนแปลงแบบพหุมิติ	33741.91	126	33993.91	34080.54

พหุมิติ (Multidimensional Approach) เปรียบเทียบกับ เอกมิติรวม (Composite Approach)

สถิติไคสแควร์อัตราส่วนไลค์ลิฮูด (Likelihood Ratio Chi-Squared Statistic; G^2): $\chi^2 = 72.21$, $df = 5$, $p = .01$

ค่า Akaike Information Criterion (AIC): $33993.91 < 34056.12$ ค่า Bayesian Information Criterion (BIC): $34080.54 < 34139.31$

เมื่อเปรียบเทียบความเหมาะสมของโมเดลระหว่างโมเดลเอกมิติและพหุมิติ จากตาราง 2 พบว่า โมเดลพหุมิติเป็นโมเดลที่เหมาะสมในการประเมินระดับสมรรถนะเพื่อการเปลี่ยนแปลง เนื่องจากมีค่าสถิติดีเวียนซ์น้อยกว่าโมเดลเอกมิติรวม อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ผลการวิเคราะห์เปรียบเทียบประสิทธิภาพโมเดลเมื่อพิจารณาเปรียบเทียบ AIC และ BIC ระหว่างสองโมเดล พบว่า โมเดลพหุมิติมีค่าต่ำกว่าโมเดลเอกมิติ ซึ่งเป็นตัวบ่งชี้ถึงโครงสร้างเพื่อประเมินระดับสมรรถนะเพื่อการเปลี่ยนแปลงว่ามีความเหมาะสมกับการวัดลักษณะพหุมิติ (Lin & Dayton, 1997)

เมื่อพิจารณาถึงการคำนวณจุดตัดโดยพิจารณาค่าเฉลี่ย Threshold ของข้อคำถามในแต่ละข้อปรากฏดังภาพประกอบ 4 ที่แสดงการกระจายของระดับสมรรถนะของผู้เรียนและการประมาณค่าพารามิเตอร์ความยากในแต่ละ Threshold ของข้อคำถามในแต่ละมิติด้วยการวิเคราะห์โมเดลพหุมิติ พบว่า โดยภาพรวม Threshold ในแต่ละขั้นของการตอบแบ่งเป็น 4 กลุ่ม คือ กลุ่ม Threshold ที่ 1 ถึงกลุ่ม Threshold ที่ 4 โดยมีการลดหลั่นระดับความยากได้อย่างเหมาะสม และเป็นเหตุเป็นผลโดย Threshold 1 < Threshold 2 < Threshold 3 < Threshold 4 สะท้อนถึงความเหมาะสมสอดคล้องกับกรอบการประเมินเพื่อวินิจฉัยระดับสมรรถนะเพื่อการเปลี่ยนแปลง รายละเอียดของแต่ละข้อปรากฏดังตาราง 3

ตาราง 3 ค่าความยากและ ค่า Threshold ของข้อคำถามในเครื่องมือประเมินสมรรถนะเพื่อการเปลี่ยนแปลง

มิติการสร้างสิ่งใหม่						มิติความรับผิดชอบ						มิติการอยู่ภายใต้ความขัดแย้ง					
ข้อ	ความยาก	Threshold				ข้อ	ความยาก	Threshold				ข้อ	ความยาก	Threshold			
		1	2	3	4			1	2	3	4			1	2	3	4
1	-0.74	-1.44	-1.15	-0.99	0.55	11	-0.55	-1.25	-0.61	-0.39	0.08	21	-0.86	-1.37	-1.06	-0.59	-0.38
2	-0.59	-3.32	-0.95	-0.80	2.77	12	-0.59	-2.53	-0.28	-0.24	0.61	22	-0.74	-1.56	-0.85	-0.62	0.06
3	-0.97	-1.69	-1.57	-1.14	0.38	13	-1.17	-1.48	-1.27	-1.16	-0.82	23	-0.64	-1.55	-0.51	-0.25	-0.10
4	-0.67	-2.43	-0.51	-0.32	0.56	14	-1.19	-1.79	-1.49	-1.18	-0.40	24	-0.69	-1.66	-0.87	-0.32	0.13
5	-1.15	-1.88	-1.49	-1.04	-0.25	15	-0.66	-2.11	-0.82	-0.56	0.84	25	-0.60	-1.51	-0.36	-0.30	-0.12
6	-1.01	-1.61	-1.44	-1.05	-0.09	16	-1.15	-1.54	-1.36	-1.27	-0.57	26	-0.38	-0.88	-0.57	-0.37	0.22
7	-0.74	-1.86	-1.21	-0.14	0.29	17	-0.62	-2.61	-0.63	-0.20	0.95	27	-0.84	-1.72	-0.75	-0.55	-0.25
8	-0.92	-1.75	-1.53	-0.48	0.01	18	-0.88	-1.87	-1.13	-0.94	0.44	28	-0.95	-2.66	-0.86	-0.72	0.40
9	-0.61	-1.62	-0.72	-0.36	0.27	19	-0.88	-1.84	-1.38	-0.32	0.03	29	-0.46	-0.97	-0.50	-0.22	-0.04
10	-0.71	-1.87	-1.43	-0.76	1.15	20	-1.01	-2.11	-1.29	-0.93	0.27	30	-0.57	-1.05	-0.75	-0.36	-0.09
Mean		-1.95	-1.20	-0.71	0.56	Mean		-1.91	-1.03	-0.72	0.14	Mean		-1.49	-0.71	-0.43	-0.02

2.2 หลักฐานความเที่ยง

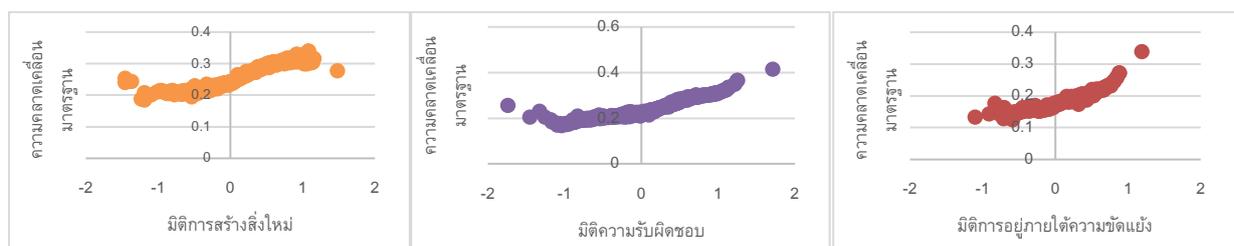
เครื่องมือที่สร้างขึ้นมีค่าความเที่ยงแบบ EAP/PV ในมิติการสร้างสิ่งใหม่ มิติความรับผิดชอบ และ มิติการอยู่ภายใต้ความขัดแย้ง เท่ากับ 0.80 0.83 และ 0.82 ตามลำดับ ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้ (Adams and Khoo, 1996) และค่าความเที่ยงแบบสอดคล้องภายใน เท่ากับ 0.87 ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้เช่นกัน

เมื่อพิจารณาข้อมูลพื้นฐานของการประมาณค่าระดับสมรรถนะในแต่ละมิติ ดังตาราง 4 พบว่า ทั้ง 3 มิติ มีค่าเฉลี่ยใกล้เคียง 0 สะท้อนให้เห็นถึงการกระจายระดับความสามารถของผู้สอบมีลักษณะการแจกแจงแบบปกติ โดยมีระดับการกระจายระดับสมรรถนะของมิติความรับผิดชอบมีช่วงกว้างมากกว่ามิติอื่นอย่างเห็นได้ชัด โดยมี θ_{TRE} อยู่ในช่วง 1.71 ถึง 1.74 ในขณะที่มิติการสร้างสิ่งใหม่ และ มิติการอยู่ภายใต้ความขัดแย้ง มีช่วงแคบกว่า คือ θ_{CNV} อยู่ในช่วง 1.48 ถึง -1.46 และ θ_{RTD} อยู่ในช่วง 1.19 ถึง -1.10 ตามลำดับ ซึ่งการประมาณค่า θ ทั้ง 3 มิติ มีช่วงของการเกิดความคลาดเคลื่อนมาตรฐานในระดับที่ใกล้เคียงกัน ดังปรากฏในตาราง 4

ตาราง 4 ค่าสถิติพื้นฐานของการประมาณค่าพารามิเตอร์สมรรถนะของผู้สอบ (θ) และ SEM ในแต่ละมิติ

	θ_{CNV}	SEM _{CNV}	θ_{TRE}	SEM _{TRE}	θ_{RTD}	SEM _{RTD}
MEAN	0.41	0.03	0.45	0.03	0.32	0.02
S.D.	0.52	0.03	0.55	0.04	0.39	0.03
MAX	1.48	0.34	1.71	0.41	1.19	0.34
MIN	-1.46	0.18	-1.74	0.17	-1.10	0.12
RANGE	2.94	0.16	3.45	0.25	2.29	0.21

นอกจากนี้เมื่อพิจารณาหลักฐานความเที่ยงจากกราฟความคลาดเคลื่อนมาตรฐานของการวัด และขอบเขตเกณฑ์ เมื่อสมรรถนะเพื่อการเปลี่ยนแปลงแยกเป็นมิตีย่อยที่สัมพันธ์กัน คือ มิติการสร้างสิ่งใหม่ (θ_{CNV}) มิติความรับผิดชอบ (θ_{TRE}) และมิติการอยู่ภายใต้ความขัดแย้ง (θ_{RTD}) ค่าพารามิเตอร์ลักษณะแฝงของผู้สอบแต่ละคน จากโมเดลพหุมิติจะมีค่าความคลาดเคลื่อนมาตรฐานใกล้เคียงกัน แสดงได้ดังภาพประกอบ 5 เมื่อพิจารณาค่า SEM ในภาพรวมที่ได้จากโมเดลพหุมิติ พบว่า ผู้สอบมีค่าเฉลี่ยในแต่ละมิตีย่อยสูงใกล้เคียงกัน ดังกราฟความสัมพันธ์ระหว่างค่าความคลาดเคลื่อนมาตรฐานจากการวัดและค่าระดับสมรรถนะในแต่ละมิติ มีลักษณะโค้งขึ้นทั้งหมด ผลวิเคราะห์ดังกล่าวเป็นหลักฐานสนับสนุนความเที่ยงของการประเมินสมรรถนะเพื่อการเปลี่ยนแปลงที่สามารถให้ค่าประมาณมิตีย่อยแฝงรายด้านที่มีความคงเส้นคงวา



ภาพ 5 ความคลาดเคลื่อนมาตรฐาน (SEM) จากการวัดในแต่ละมิติ

2.3 ผลการวิเคราะห์สถิติความเหมาะสมรายข้อ (Item fit)

ผลการวิเคราะห์ค่าสถิติความเหมาะสมรายข้อ (item fit) โดยพิจารณาค่า OUTFIT MNSQ และ INFIT MNSQ ของข้อคำถามแต่ละข้อ ซึ่งควรมีค่าอยู่ระหว่าง 0.75 ถึง 1.33 (Adam & Khoo, 1996; Willson, Allen & Li, 2006) จากการวิเคราะห์ด้วยโมเดลพหุมิติพบว่าค่า OUTFIT MNSQ ของทั้งฉบับมีค่าอยู่ระหว่าง 0.77 ถึง 1.73 ส่วนค่า INFIT MNSQ มีค่าอยู่ระหว่าง 0.86 ถึง 1.30 ข้อคำถามทุกข้ออยู่ในเกณฑ์ที่ยอมรับได้

อภิปรายผล

จากผลการวิจัยในครั้งนี้พบว่าแผนที่โครงสร้างของสมรรถนะเพื่อการเปลี่ยนแปลงเป็นไปตามกรอบการประเมินสมรรถนะเพื่อการเปลี่ยนแปลงของ Podjana et al. (In press) ทั้งสามมิติ แต่อย่างไรก็ตาม การวิจัยครั้งนี้เป็นการพัฒนาข้อคำถามจากเดิมที่เป็นข้อคำถามสถานการณ์ปลายเปิดเพื่อสร้างกรอบการประเมินลักษณะของคำถามบางคำถามมีความเกี่ยวข้องกับเหตุการณ์ตั้งเครียดในปัจจุบัน ยกตัวอย่างคำถามในมิติการอยู่ภายใต้ความขัดแย้ง เช่น “หากผู้เรียนมีความคิดเห็นทางการเมืองไม่เหมือนกับเพื่อนสนิท ผู้เรียนจะทำอย่างไรเพื่ออธิบายจุดยืนของตนเอง” ข้อคำถามดังกล่าวมีผู้สอบที่ตอบคำถามในด้านบวกไปจนถึงด้านลบแบบกระจาย พบว่าในระดับที่ 1 หรือระดับต่ำสุดนั้นมักเป็นคำตอบที่แสดงถึงอารมณ์เชิงลบ ซึ่งไม่ก่อให้เกิดการจัดการปัญหาที่แท้จริง ในขณะที่ข้อคำถามในมิติเดียวกันแต่เป็นข้อคำถามที่เป็นสถานการณ์ทั่วไป เช่น “หากผู้เรียนเป็นหัวหน้าห้อง และรับผิดชอบการทำเล็อรุ่น ผู้เรียนจะทำอย่างไรหากความคิดเห็นของเพื่อนในห้องไม่ตรงกัน” คำตอบที่ได้จากข้อคำถามนั้นโดยส่วนใหญ่ผู้สอบจะแก้ไขปัญหาได้ในระดับที่สูงขึ้น ทำให้ในมิติเดียวกัน การตอบคำถามไม่สัมพันธ์กัน ดังนั้นจึงควรหลีกเลี่ยงข้อคำถามในลักษณะที่มีความตั้งเครียดมากเกินไป อันจะก่อให้เกิดความคลาดเคลื่อนจากการประเมินได้

นอกจากนี้ในการพัฒนาข้อคำถามเชิงสถานการณ์แบบเลือกตอบ 5 ตัวเลือก ผลการวิจัยพบว่าการกระจายระดับสมรรถนะตาม wright map ของข้อคำถามแบบเลือกตอบ (logit +2 ถึง -2) มีช่วงที่แคบกว่าข้อคำถามสถานการณ์ปลายเปิดที่ได้จากการสร้างกรอบการประเมินในการวิจัยระยะการศึกษานำร่อง (logit -9 ถึง +8) อย่างไรก็ตาม การกระจายระดับสมรรถนะเพื่อการเปลี่ยนแปลงเป็นตัวแปรที่มีความซับซ้อนทั้งในการใช้ ความรู้ ทักษะและเจตคติ ซึ่งสะท้อนคุณลักษณะที่มีแนวโน้มส่วนใหญ่ออกมาทางด้านจิตใจหรือคุณธรรม (Non-cognitive) สอดคล้องกับผลการวิจัยของ Piyapimonsit and Suksiri (2017) ที่มีการพัฒนาเครื่องมือวัดบุคลิกลักษณะของนักเรียนตามทฤษฎีบุคลิกลักษณะห้าด้าน และ Wenukaset (2019) พัฒนาแบบวัดความเป็นพลเมืองดีด้วยเครื่องมือที่เป็นแบบวัดเชิงสถานการณ์แบบเลือกตอบ พบว่า การกระจายค่าความสามารถของผู้สอบอยู่ระหว่าง logit -1 ถึง +1 เท่านั้น ซึ่งจะแตกต่างจากการวัดคุณลักษณะที่เป็นทักษะทางปัญญา (cognitive) ที่มีการกระจายค่าความสามารถที่มีช่วงกว้าง และเมื่อพิจารณาจากผลของ wright map พบว่าในแต่ละมิติสามารถวัดได้จากระดับต่ำสุดไปสูงสุดเป็นระดับ แต่ยังคงขาดกลุ่มผู้สอบในระดับต่ำสุดในแต่ละมิติ สะท้อนให้เห็นในสองลักษณะของผลการศึกษาในครั้งนี้ โดยกรณีที่ 1 อาจเนื่องมาจากรายการประเมินที่สร้างขึ้นวัดสมรรถนะเพื่อการเปลี่ยนแปลงของผู้สอบได้จริงบางส่วน อาจไม่ครอบคลุมทั้งหมด ซึ่งอาจ

เกิดจากสถานการณ์ที่วัดเป็นคุณลักษณะด้านความคิด ผู้สอบอาจตอบไม่ตรงกับความจริงทำให้เกิดการตอบคำถามแบบโกหกในการประเมินระดับสมรรถนะระดับที่ 1 ซึ่งตัวเลือกอาจมีความสับสนทางลบอย่างชัดเจน ทำให้ผู้สอบไม่ตอบตัวเลือกดังกล่าว กรณีที่ 2 อาจจะเป็นเนื่องมาจาก การเก็บข้อมูลในครั้งนี้อาจยังไม่สามารถหากลุ่มที่มีสมรรถนะเพื่อการเปลี่ยนแปลงในแต่ละมิติได้เพียงพอที่จะเป็นตัวแทนของผู้เรียนที่มีระดับต่ำสุดของระดับสมรรถนะได้ ซึ่งผลดังกล่าวสอดคล้องกับผลการศึกษาของ Baker & Kim (2017); Damar (2010); Junpeng et al. (2019) ที่ให้ข้อเสนอแนะว่าในการออกแบบการประเมินในการพัฒนาเครื่องมือ ไม่ได้มุ่งเน้นการสุ่มกลุ่มตัวอย่างเพื่อสุ่มอ้างอิงไปสู่ประชากร แต่ควรมุ่งเน้นความเป็นตัวแทนที่มาจากกลุ่มที่เป็นตัวแทนของสมรรถนะเพื่อการเปลี่ยนแปลงทุกระดับ โดยข้อมูลเชิงประจักษ์ที่สามารถพิจารณาได้เบื้องต้น คือลักษณะการกระจายสมรรถนะควรอยู่ในช่วง logit -3 to +3 ดังนั้นในการวิจัยครั้งต่อไปจำเป็นต้องมีการให้ความสำคัญในการคัดเลือกกลุ่มผู้เรียนให้ครอบคลุมทุกระดับสมรรถนะ โดยเฉพาะในระดับที่ 1

นอกจากนี้ผลการวิจัยพบว่ากรอบการประเมินที่พัฒนาขึ้นจะมีความถูกต้องและคงเส้นคงวาเหมาะสมสำหรับผู้เรียนที่มีสมรรถนะปานกลางมาทางต่ำทั้ง 3 มิติ ซึ่งข้อค้นพบดังกล่าวมีประโยชน์อย่างยิ่งต่อการนำไปใช้ในการคัดกรองผู้เรียนที่มีสมรรถนะเพื่อการเปลี่ยนแปลงในผู้เรียนที่มีระดับต่ำ มีประโยชน์ต่อการนำเสนอสารสนเทศที่ได้ไปสู่การพัฒนาสมรรถนะของผู้เรียนให้อยู่ในระดับที่สูงขึ้น

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 จากการพัฒนาเครื่องมือประเมินสมรรถนะเพื่อการเปลี่ยนแปลง พบว่าเครื่องมือชุดนี้มีความเหมาะสมกับผู้สอบที่มีความสามารถกลางค่อนข้างต่ำ ซึ่งข้อค้นพบดังกล่าวมีประโยชน์อย่างยิ่งต่อการนำเครื่องมือไปใช้ประเมินในการคัดกรองผู้เรียนที่มีสมรรถนะเพื่อการเปลี่ยนแปลงที่อยู่ในระดับต่ำ เพื่อมีประโยชน์ต่อการนำเสนอสารสนเทศที่ได้ไปสู่การพัฒนาสมรรถนะของผู้เรียนให้อยู่ในระดับที่สูงขึ้น

1.2 ครูผู้สอนสามารถนำข้อมูลจากการกำหนดแผนที่โครงสร้างของระดับสมรรถนะเพื่อการเปลี่ยนแปลงที่มีคะแนนจุดตัดทั้ง 4 ระดับนี้ ไปวินิจฉัยผู้เรียนเพื่อพัฒนาผู้เรียนให้มีระดับที่สูงขึ้น และเป็นการเพิ่มศักยภาพในการให้ข้อมูลย้อนกลับแก่ผู้เรียนในระบบสารสนเทศ และนำไปสู่การปรับปรุงการเรียนการสอนให้มีประสิทธิภาพสูงขึ้น

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 สำหรับการประเมินสมรรถนะเพื่อการเปลี่ยนแปลงจากผลการวิจัยที่แสดงให้เห็นว่าข้อคำถามสามารถวัดสมรรถนะอยู่ในระดับต่ำสุดได้ แต่กลุ่มผู้สอบในการวิจัยครั้งนี้ยังมีความสามารถในระดับนี้ อาจยังไม่มีความเป็นตัวแทนของกลุ่มระดับดังกล่าวที่เพียงพอ นั่นคือมีความจำเป็นต้องค้นหากลุ่มที่มีความเป็นตัวแทนระดับต่ำสุดของระดับสมรรถนะเพื่อการเปลี่ยนแปลง เนื่องจากสามารถวัดผู้สอบได้เหมาะสมในระดับที่ 2-5 แต่ยังไม่มีการวัดผู้สอบที่เป็นตัวแทนของระดับที่ 1 ซึ่งมีความจำเป็นต้องนำศึกษาต่อไปในอนาคต

2.2 การทำวิจัยในครั้งนี้อาจมีการตรวจสอบคุณภาพด้านความตรง คือ หลักฐานความตรงด้านเนื้อหา ของรายการประเมิน หลักฐานความตรงด้านกระบวนการตอบ และหลักฐานความตรงด้านโครงสร้างภายใน

ในการวิจัยครั้งต่อไปควรมีการตรวจสอบหลักฐานความตรงเพิ่มเติมอีก 2 ด้าน คือ (1) ความตรงเกี่ยวกับความสัมพันธ์กับเกณฑ์อื่น ซึ่งอาจจะใช้คะแนนผลงาน หรือ เกณฑ์เฉลี่ยสะสม เพื่อหาความสัมพันธ์ เพื่อใช้เป็นหลักฐานความตรงตามสภาพและความตรงเชิงทำนาย และ (2) หลักฐานความตรงที่เป็นผลจากการทดสอบ ซึ่งสะท้อนให้เห็นถึงผลกระทบทางบวกและทางลบจากผลการประเมินจากการสัมภาษณ์หรือการประเมินการนำผลจากการประเมินสมรรถนะเพื่อการเปลี่ยนแปลงไปใช้

References

- Adams, R. J., & Khoo, S. T. (1996). *Quest*. Australian Council for Educational Research.
- Adams, R. J., Wilson, M., & Wang, W. C. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement, 21*(1), 1-23
- AERA, APA, & NCME. (2014). *Standards for Educational and Psychological Testing* (6th ed.). American Educational Research Association.
- Baker, F. B., & Kim, S. H. (2017). *The Basics of Item Response Theory Using R*. Springer International Publishing.
- Canto-Sperber, M., & Depuy, J. P. (2001). *Competencies for good life and good society*. In *Defining and selecting competencies*, edited by D. Rychen and L. Salganik, pp. 67-92. Kirkland, WA: Hogrefe & Huber.
- DeMars, C. (2010). *Item Response Theory: Understanding Statistics Measurement*. Oxford University Press.
- European Commission. (2007). Key competences-Education and training-European Commission. http://ec.europa.eu/education/policy/school/competences_en.
- Grayling, A. (2017). Education 2030 - *Conceptual Learning Framework: Background papers: Observations on "Taking responsibility" and "Coping with tensions and dilemmas"*. <http://www.oecd.org/education/2030-project/about/documents>.
- Haste, H. (2001). "Ambiguity, autonomy, and agency: Psychological challenges to new competence", in Rychen, D. and L. Salganik (Eds.). *Defining and selecting key competencies*. (pp. 93-120). Hogrefe & Huber.
- Junpeng, P., Krotha, J., Chanayota, K., Tang, K. N. & Willson, M. (2019). Constructing Progress Maps of Digital Technology for Diagnostic Mathematical Proficiency. *Journal of Education and Learning, 8*(6), 90-102.
- Leadbeater, C. (2017). *Students as agents of change*. <http://www.oecd.org/education/2030-project/about/documents>.
- Lin, T.H., & Dayton, C. M. (1997). Model Selection Information Criteria for Non-Nested Latent Class Model. *Journal of Educational and Behavioral Statistics, 22*(3), 249-264.

- Nussbaum, M. (1997). *Cultivating Humanity: a Classical Defense of Reform in Liberal Education*. Harvard University Press.
- OECD. (2018). *The Future of Education and Skills: Education 2030*. [https://www.oecd.org/education/2030/E2030%20Position%20Paper%20\(05.04.2018\).pdf](https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf).
- Podjana, C., Junpeng, P., Luanganggoon, N., Nongna, C. & Tang, K. N. (in press). Developing a Framework for Assessing Transformative Competencies of Junior High School Students Through Construct Modeling Approach. *Humanities, Arts and Social Sciences Studies*.
- Reeves, Thomas, C. (2006). *Design research from a technology perspective*. In Van den Akker, Jan, Gravemeijer, Koeno, McKenney, Susan and Nieveen, Nienke. (Eds). *Educational Design Research*. Routledge, Taylor and Francis Group. (pp. 52-62).
- Sass D. A., Schmitt T. A., & Walker C. M. (2008). Estimating non-normal latent trait distributions with item response theory using true and estimated item parameters. *Applied Measurement in Education*, 21, 65-88
- Steinberg, L. (2017). *Transformative Competencies 2030: Taking Responsibility*. <http://www.oecd.org/education/2030-project/about/documents>
- Wilson, M. (2005). *Constructing measures: An item response modeling approach (Har/Cdr edition)*. Routledge.
- Wilson, M., Allen, D. D., & Li, J. C. (2006). Improving measurement in health education and health behavior research using item response modelling: Comparison with the classical test theory approach. *Health Education Research*, 2(1), 19-32.
- Yao, L., and Schwarz, R.D. (2006). A multidimensional partial credit model with associated item and test statistics: An application to mixed-format tests. *Applied Psychological Measurement*, 30(6), 469-492.

Translate Thai Reference

- Office of the Education Council. (2019). *A quick guide for people on the competency and a quick guide for Teachers, Administrators and Educational Personnel on Competency-Based Curriculum*. Bangkok, Thailand: 21st Century. (in Thai)
- Piyapimonsit, C. & Suksiri, W. (2017). *Developing Student's Personality Trait Inventory According to Big-Five Personality Trait Theory Using Construct Mapping and Rasch Measurement Approach*. National Institute of Educational Testing Service. (in Thai)
- Vongvanich, S. (2020). *Design research in education*. Bangkok, Thailand: Chulalongkorn University Printing House. (in Thai)
- Wenukaset, W. (2019). *A Development of Citizenship Measure: Applying Construct Modeling Approach* [Unpublished Master's thesis]. Khon Kean University. (in Thai)