

การพัฒนาคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ
ของนักเรียนชั้นมัธยมศึกษาตอนปลาย

The Development of an Item Bank for a Management and Responsibility
Competency Scale for Upper Secondary Students

โสธญา แสนเมือง^{1*} ประกฤติยา ทักซิโน² และ กรชวัล ชายผา³

Soraya Sanmuang^{1*} Prakittiya Tuksino² and Kornchawal Chaipah

^{1,2}สาขาวิชาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

(Department of Educational Measurement and Evaluation, Faculty of Education, Khon Kaen University)

³สาขาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น

(Department of Computer Engineering, Faculty of Engineering, Khon Kaen University)

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อตรวจสอบความตรงเชิงโครงสร้างของโมเดลการวัดสมรรถนะและหาคุณภาพคลังข้อคำถามสำหรับประเมินสมรรถนะด้านการจัดการและความรับผิดชอบของนักเรียนชั้นมัธยมศึกษาตอนปลาย รวมถึงการกำหนดจุดตัดสมรรถนะเพื่อการจำแนกระดับสมรรถนะอย่างแม่นยำ คลังข้อคำถามมีลักษณะเป็นที่ใช้เป็นแบบวัดเชิงสถานการณ์ ที่มีตัวเลือกคำตอบ 4 ระดับ ตรวจให้คะแนนมากกว่า 2 ค่า พัฒนามาตามกรอบแนวคิดด้านจิตพิสัยของ Krathwohl เพื่อสะท้อนพฤติกรรมในสถานการณ์จริง วิเคราะห์ข้อมูลด้วยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) โมเดล Graded Response Model (GRM) กลุ่มตัวอย่างคือ นักเรียนชั้นมัธยมศึกษาตอนปลาย จำนวน 832 คน ผลการวิจัย พบว่า 1) โมเดลการวัดประกอบด้วยสมรรถนะย่อย 4 ด้าน รวม 14 พฤติกรรมบ่งชี้ ได้แก่ (1) ความเชื่อมั่นในตนเอง (5 พฤติกรรมบ่งชี้) (2) การจัดการความเครียดและความไม่แน่นอน (3 พฤติกรรมบ่งชี้) (3) การวางแผนและการจัดการองค์กร (3 พฤติกรรมบ่งชี้) และ (4) ความรับผิดชอบ (3 พฤติกรรมบ่งชี้) รวม 84 ข้อ ผลการตรวจสอบความตรงเชิงโครงสร้างพบว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ ($\chi^2 = 73.874$, $df = 57$, $p\text{-value} = 0.0661$, $RMSEA = 0.019$, $SRMR = 0.016$, $CFI = 0.997$, $TLI = 0.994$) ข้อคำถามทุกข้อผ่านเกณฑ์พารามิเตอร์ความจำแนก (α) มีค่าเฉลี่ย 0.904 (0.516–1.891) และพารามิเตอร์เกณฑ์ (β) มีค่าเฉลี่ยที่ $\beta_1 = -0.281$, $\beta_2 = 0.906$, $\beta_3 = 2.122$ ค่าความเชื่อมั่นรวมของแบบวัดทั้งฉบับ (Cronbach's alpha) เท่ากับ 0.925 กราฟฟังก์ชันสารสนเทศมีความแม่นยำสูงในช่วง $\theta \approx -1.0$ ถึง $+3$ เหมาะสำหรับนักเรียนที่มีสมรรถนะระดับปานกลางถึงสูง กำหนดจุดตัดสมรรถนะโดยใช้ค่าเฉลี่ยพารามิเตอร์ threshold และ Wright Map แบ่งเป็น 4 ระดับ ได้แก่ ระดับ 1 $\theta \leq -0.28$ (มีสมรรถนะไม่เพียงพอ) ระดับ 2 $-0.28 < \theta \leq 0.91$ (เริ่มมีสมรรถนะ) ระดับ 3 $0.91 < \theta \leq 2.12$ (มีสมรรถนะ) และ ระดับ 4 $\theta > 2.12$ (มีสมรรถนะสูง) คลังข้อคำถามนี้เหมาะสมต่อการเข้าสู่ระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) ในอนาคต

คำสำคัญ: คลังข้อคำถาม สมรรถนะการจัดการและความรับผิดชอบ ทฤษฎีการตอบสนองข้อสอบ

โมเดล Graded Response Model (GRM) จุดตัด

ABSTRACT

This study aimed to develop and examine the quality of an item bank for assessing management and responsibility competency among upper secondary school students, as well as to establish cut scores for accurate competency classification. The instrument was a situational test with four response options, polytomously scored, developed based on the affective domain framework of Krathwohl to reflect authentic behaviors in real-life situations. Data were analyzed using Item Response Theory (IRT) with the Graded Response Model (GRM). The sample comprised 832 upper secondary school students. The results showed that 1) The measurement model consisted of four sub-competencies with a total of 14 behavioral indicators: (1) self-efficacy (5 indicators), (2) stress and uncertainty management (3 indicators), (3) planning and organizational management (3 indicators), and (4) responsibility (3 indicators), totaling 84 items. Confirmatory factor analysis indicated that the model fit the empirical data well ($\chi^2 = 73.874$, $df = 57$, $p\text{-value} = 0.0661$, $RMSEA = 0.019$, $SRMR = 0.016$, $CFI = 0.997$, $TLI = 0.994$). All items met the selection criteria for discrimination (α) with a mean of 0.904 (range: 0.516–1.891) and threshold (β) parameters with means of $\beta_1 = -0.281$, $\beta_2 = 0.906$, $\beta_3 = 2.122$. The overall reliability (Cronbach's alpha) was 0.925, and the test information function demonstrated high measurement precision for ability levels in the range $\theta \approx -1.0$ to $+3$, this is suitable for students with intermediate to high ability. Competency cut scores were established using the mean threshold parameters and Wright Map, dividing competency into four levels: $\theta \leq -0.28$ (insufficient), $-0.28 < \theta \leq 0.91$ (emerging), $0.91 < \theta \leq 2.12$ (competent), and $\theta > 2.12$ (highly competent). The validated item bank is appropriate for accurate competency assessment and is ready for integration into a computerized adaptive testing (CAT) system in the future.

KEYWORDS: Item bank, Management and responsibility competency, Item Response Theory (IRT), Graded Response Model (GRM), Cut scores

*Corresponding author, E-mail: S.soraya@kkumail.com Tel. 0933194954

Received: 14 August 2025 /Revised: 9 October 2025 /Accepted: 31 October 2025 /Published online: 26 December 2025

บทนำ

การเปลี่ยนแปลงอย่างรวดเร็วของเศรษฐกิจ สังคม และเทคโนโลยีในศตวรรษที่ 21 ได้ส่งผลให้ระบบการศึกษาและตลาดแรงงานทั่วโลกต้องเผชิญกับความท้าทายใหม่ที่ซับซ้อนและยืดหยุ่นมากขึ้น รายงาน *Education at a Glance 2024* ขององค์การเพื่อความร่วมมือและการพัฒนาทางเศรษฐกิจ (OECD, 2024) ชี้ให้เห็นว่า แม้อัตราการสำเร็จการศึกษาระดับมัธยมศึกษาตอนปลายของประชากรอายุ 25–34 ปีจะเพิ่มขึ้นเป็นร้อยละ 86 แต่อย่างไรก็ตาม ปัญหาความไม่สอดคล้องระหว่างทักษะของผู้สำเร็จการศึกษาและความต้องการของตลาดแรงงาน (skills mismatch) ยังคงเป็นประเด็นสำคัญที่ส่งผลต่อประสิทธิภาพของระบบการศึกษาและความสามารถในการแข่งขันของประเทศในระยะยาว ในทำนองเดียวกัน รายงาน *The State of Global Teenage Career Preparation 2025* ของ OECD (2025) พบว่า นักเรียนอายุ 15 ปีทั่วโลกกว่าครึ่งหนึ่งไม่เคยมีโอกาสร่วมกิจกรรมแนะแนวอาชีพหรือการพบปะกับผู้ประกอบการอาชีพจริง ซึ่งลดโอกาสในการเตรียมความ

พร้อมสำหรับการทำงานในอนาคตอย่างมีนัยสำคัญ ขณะที่รายงาน *Future of Jobs Report 2025* ของ World Economic Forum (WEF, 2025) ระบุว่า โลกของการทำงานกำลังเผชิญกับภาวะ “ถูกรบกวนด้านทักษะ” (skills disruption) อย่างรุนแรง ซึ่งเป็นผลจากเทคโนโลยีปัญญาประดิษฐ์ ระบบอัตโนมัติ และความเปลี่ยนแปลงทางเศรษฐกิจในระดับโลก โดยทักษะที่มีความต้องการสูงในอนาคตได้แก่ ความสามารถในการจัดการตนเอง (self-management) ความยืดหยุ่นและการปรับตัว (resilience & adaptability) การแก้ปัญหาที่ซับซ้อน (complex problem solving) ความคิดสร้างสรรค์ (creativity) การสื่อสารและการทำงานร่วมกับผู้อื่นอย่างมีประสิทธิภาพ (Effective Communication and Collaboration with Others) ซึ่งล้วนเป็นสมรรถนะสำคัญของนักเรียนในศตวรรษที่ 21

พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 และแผนการศึกษาแห่งชาติ พ.ศ. 2560–2579 ตลอดจนสำนักงานเลขาธิการสภาการศึกษา ปี พ.ศ. 2564 ได้กำหนดให้การพัฒนาสมรรถนะนักเรียนเป็นเป้าหมายสำคัญ โดยเน้นการบูรณาการร่วมกันระหว่างความรู้ ทักษะ เจตคติ และคุณลักษณะ เพื่อสามารถประยุกต์ใช้ในชีวิตและการทำงานอย่างมีประสิทธิภาพ (Ministry of Education, 1999; Office of the Education Council, 2017; Office of the Education Council, 2021) แนวคิดนี้สอดคล้องกับกรอบสมรรถนะสากลขององค์การเพื่อความร่วมมือและการพัฒนาทางเศรษฐกิจ (OECD, 2018; 2019) ที่เน้นสมรรถนะการจัดการตนเอง (self-management) และความรับผิดชอบ (responsibility-taking) เป็นหัวใจของการเตรียมพลโลก รวมทั้งกรอบสมรรถนะเพื่อการเรียนรู้ตลอดชีวิตของคณะกรรมาธิการยุโรป (European Commission, 2018) และ Competency Framework ของ UNESCO (2015) ซึ่งระบุชัดเจนว่าทักษะการวางแผน การจัดการตนเอง และการรับผิดชอบเป็นปัจจัยสำคัญของความสำเร็จในชีวิตและการทำงาน

ในบริบทของประเทศไทย การปฏิรูปการศึกษาในปัจจุบันให้ความสำคัญกับการพัฒนาสมรรถนะนักเรียนตามแนวคิดของ Competency-Based Education (CBE) โดยกรอบสมรรถนะหลักของประเทศไทย (CBE Thailand, 2021) ได้ระบุ “สมรรถนะการจัดการและความรับผิดชอบ” เป็นหนึ่งในสมรรถนะพื้นฐานที่นักเรียนทุกคนต้องได้รับการพัฒนา เพื่อเตรียมความพร้อมในการดำรงชีวิตในสังคมที่ซับซ้อนและเชื่อมโยงถึงกันอย่างลึกซึ้ง ขณะเดียวกัน องค์การระหว่างประเทศ เช่น UNESCO (2015) และ European Commission (2018) ต่างยืนยันว่า ความสามารถในการวางแผน จัดระบบ จัดการตนเอง และรับผิดชอบต่อผลการกระทำ เป็นหัวใจของความสำเร็จในศตวรรษที่ 21 ขณะเดียวกัน World Economic Forum (2020) ระบุทักษะการจัดการตนเอง ความยืดหยุ่นต่อความเครียด การปรับตัว และการทำงานร่วมกับผู้อื่นเป็นทักษะสำคัญที่ตลาดแรงงานอนาคตต้องการสูง

จากบริบทดังกล่าว การพัฒนาสมรรถนะการจัดการและความรับผิดชอบจึงเป็นประเด็นสำคัญที่ได้รับการกล่าวถึงในนโยบายการศึกษาในระดับชาติและนานาชาติ เนื่องจากสมรรถนะด้านนี้ถือเป็นพื้นฐานของการเรียนรู้ตลอดชีวิตและการเป็นพลเมืองที่มีคุณภาพ อย่างไรก็ตาม ในบริบทของประเทศไทยยังคงขาดเครื่องมือวัดสมรรถนะดังกล่าวที่ได้รับการพัฒนาและตรวจสอบคุณภาพอย่างเป็นระบบและมีหลักฐานเชิงประจักษ์ โดยเฉพาะในระดับมัธยมศึกษาตอนปลาย ซึ่งเป็นช่วงวัยที่นักเรียนกำลังเตรียมความพร้อมสำหรับการศึกษาต่อและการเข้าสู่ตลาดแรงงานในอนาคต

จากการศึกษางานวิจัยที่ผ่านมาพบว่า มีการพัฒนาเครื่องมือประเมินสมรรถนะของนักเรียน แต่เครื่องมือเหล่านี้ส่วนใหญ่เน้นระดับประถมศึกษาหรืออาชีวศึกษา ใช้โมเดลการวัดสมรรถนะอ้างอิงตามกรอบสมรรถนะของหลักสูตรเดิม (Srisakda, 2015; Tuksino et al., 2016; Phrommabun et al., 2020; Langka et al., 2019) โดยยังไม่ปรากฏงานวิจัยที่มุ่งพัฒนาเครื่องมือเฉพาะสมรรถนะร่วมด้านการจัดการและความรับผิดชอบโดยมีการตรวจสอบคุณภาพครบถ้วนทั้งด้านความตรง (validity) ความเที่ยง (reliability) และค่าพารามิเตอร์ของข้อสอบ จึงทำให้ขาดแคลนแบบวัดสมรรถนะที่มีคุณภาพ เนื่องจากเครื่องมือวัดส่วนใหญ่ในประเทศไทยยังอาศัยทฤษฎีการวัดแบบดั้งเดิม หรือ Classical Test Theory (CTT) ซึ่งมีข้อจำกัดในหลายด้าน เช่น ค่าความเที่ยงและค่าความยากง่ายของข้อสอบขึ้นอยู่กับกลุ่มตัวอย่าง ทำให้ไม่สามารถเปรียบเทียบผลระหว่างกลุ่มนักเรียนต่าง ๆ ได้อย่างเที่ยงตรงและยุติธรรม (Embretson & Reise, 2000) ดังนั้น การนำทฤษฎีการตอบสนองข้อสอบ

(Item Response Theory: IRT) มาใช้ในการพัฒนาเครื่องมือวัดจึงเป็นแนวทางที่มีความเหมาะสมกว่า เนื่องจากสามารถประมาณค่าความสามารถของนักเรียนได้อย่างแม่นยำในทุกระดับ (Lord, 1980) และมีคุณสมบัติที่สำคัญคือ item invariance ซึ่งหมายถึงพารามิเตอร์ของข้อสอบไม่ขึ้นกับกลุ่มตัวอย่าง อีกทั้งยังเหมาะต่อการพัฒนาระบบทดสอบแบบปรับเหมาะ (Computerized Adaptive Testing: CAT) ที่ใช้คอมพิวเตอร์ในการเลือกข้อสอบให้เหมาะกับระดับความสามารถของผู้เข้าสอบแต่ละคน (Hambleton & Swaminathan, 1985)

นอกจากนี้ ทฤษฎี IRT ยังสามารถใช้วิเคราะห์ข้อสอบที่มีระดับคะแนนหลายระดับ (polytomous items) ได้อย่างมีประสิทธิภาพ โดยเฉพาะการใช้โมเดล Graded Response Model (GRM) ซึ่ง Samejima (1969) ได้พัฒนาเพื่อใช้กับข้อสอบที่มีการให้คะแนนมากกว่า 2 ค่า แทนที่จะเป็นเพียง 2 ค่า ซึ่งเหมาะสมอย่างยิ่งสำหรับการวัดสมรรถนะด้านการจัดการและความรับผิดชอบที่มีพฤติกรรมคำตอบแบบหลายระดับ เพราะสามารถวิเคราะห์พฤติกรรมคำตอบในแต่ละระดับของสมรรถนะได้ละเอียดกว่าที่ใช้ในแบบสอบทั่วไป สอดคล้องกับงานวิจัยของ Chanchusakul (2021) ที่พบว่าเครื่องมือวัดที่มีการให้คะแนนมากกว่า 2 ค่า มีค่าความเที่ยงและความสามารถในการจำแนกนักเรียนได้ดีกว่าแบบวัดที่มีการตรวจให้คะแนน 2 ค่า

อีกปัญหาหนึ่งที่สำคัญคือ การขาดการกำหนดจุดตัดสมรรถนะที่มีหลักฐานเชิงประจักษ์ ปัจจุบันการแบ่งระดับสมรรถนะมักอาศัยการพิจารณาจากความคิดเห็นของผู้เชี่ยวชาญเพียงอย่างเดียว โดยไม่ได้นำข้อมูลเชิงสถิติจากการวิเคราะห์ข้อสอบมาสนับสนุน ส่งผลให้การแปลผลคะแนนอาจขาดความเที่ยงตรง (Cizek & Bunch, 2007) การใช้ข้อมูลเชิงประจักษ์จากการวิเคราะห์ข้อสอบตามแนวทางของ IRT จึงเป็นสิ่งจำเป็นในการกำหนดระดับสมรรถนะอย่างมีหลักฐานรองรับ โดยการกำหนดจุดตัดระดับสมรรถนะ (competency cut scores) ถือเป็นขั้นตอนสำคัญที่ทำให้การประเมินมีความหมายเชิงตีความสามารถจัดจำแนกนักเรียนออกเป็นกลุ่มตามระดับความสามารถได้อย่างชัดเจน โดยในการวิจัยครั้งนี้ จุดตัดถูกกำหนดโดยใช้ค่าพารามิเตอร์ threshold จากโมเดล Graded Response Model (GRM) ซึ่งบ่งชี้ระดับความสามารถที่นักเรียนต้องมีเพื่อให้ได้คะแนนในแต่ละระดับของข้อคำถาม เมื่อนำค่าพารามิเตอร์เหล่านี้มาวิเคราะห์และเฉลี่ย จะสามารถกำหนดช่วงคะแนนความสามารถ (theta scores) ที่สอดคล้องกับแต่ละระดับสมรรถนะได้ โดยการกำหนดจุดตัดระดับสมรรถนะเช่นนี้ ช่วยให้ผลการประเมินไม่เพียงบอกคะแนนความสามารถของนักเรียนในหน่วยเชิงสถิติ แต่ยังสามารถสื่อสารเป็นระดับสมรรถนะเชิงพฤติกรรม (performance level descriptors) ที่ผู้บริหาร ครู และผู้เกี่ยวข้องสามารถนำไปใช้กำหนดแผนพัฒนา การจัดการกิจกรรมการเรียนรู้ หรือการให้คำปรึกษาได้อย่างตรงจุดและมีประสิทธิภาพ

แนวคิดการพัฒนาคคลังข้อสอบ (Item Bank) ตามหลักทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ได้รับการยอมรับอย่างกว้างขวางในวงการการวัดและประเมินผล เนื่องจากเป็นกลไกสำคัญที่ช่วยให้ระบบการทดสอบมีความยืดหยุ่น มีประสิทธิภาพ และสามารถนำไปใช้ซ้ำได้อย่างคุ้มค่าในระยะยาว (Wainer et al., 2000) โดยเป็นการรวบรวมข้อคำถามที่ผ่านการวิเคราะห์เชิงสถิติและเชิงจิตพิสัยอย่างละเอียด โดยเฉพาะการตรวจสอบค่าพารามิเตอร์ของข้อสอบแต่ละข้อ เช่น ค่า a (discrimination) ซึ่งบ่งชี้ถึงความสามารถของข้อในการจำแนกผู้ที่มีระดับสมรรถนะต่างกันได้ดีเพียงใด, ค่า b (difficulty/threshold) ซึ่งแสดงระดับความสามารถที่ผู้ตอบมีโอกาสตอบถูกประมาณ 50%, และในกรณีของแบบวัดที่ให้คะแนนมากกว่า 2 ค่า (polytomous items) ยังมีค่าพารามิเตอร์เพิ่มเติมที่สะท้อนระดับการยากของแต่ละระดับคะแนน (Samejima, 1969; Embretson & Reise, 2000) การมีคลังข้อสอบคุณภาพสูงยังเอื้อให้สามารถ ติดตามพัฒนาการของนักเรียน (progress monitoring) ได้อย่างต่อเนื่อง โดยเมื่อข้อสอบทุกข้อในคลังมีพารามิเตอร์ที่ถูกต้องและอยู่ในสเกลเดียวกัน การนำข้อจากคลังเดียวกันมาทดสอบในช่วงเวลาต่าง ๆ จะทำให้สามารถเปรียบเทียบผลคะแนนในระยะยาวได้อย่างแม่นยำ ซึ่งมีประโยชน์ต่อการประเมินผลสัมฤทธิ์เชิงสมรรถนะและการวางแผนพัฒนานักเรียนรายบุคคลในระดับสถานศึกษาและระดับนโยบาย (Lord, 1980; van der Linden, 2018)

การศึกษาค้นคว้าครั้งนี้มุ่งเน้นการพัฒนาเครื่องมือประเมินสมรรถนะการจัดการและความรับผิดชอบของนักเรียนชั้นมัธยมศึกษาตอนปลาย โดยใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) ผ่านโมเดล Graded Response Model (GRM) เพื่อให้สามารถวัดสมรรถนะได้อย่างแม่นยำในทุกระดับ รวมทั้งตรวจสอบคุณภาพของข้อคำถามและคลังข้อคำถามอย่างเข้มงวดตามหลักจิตพิสัยเชิงประจักษ์ ตลอดจนกำหนดจุดตัดสมรรถนะที่อิงข้อมูลจริง เพื่อเพิ่มความถูกต้องและโปร่งใสในการแปลผลคะแนน ทั้งนี้ ผลการวิจัยจะเป็นพื้นฐานสำคัญในการพัฒนาระบบการประเมินสมรรถนะในรูปแบบ Computerized Adaptive Testing (CAT) ที่สามารถใช้ได้จริงในบริบทของการศึกษาขั้นพื้นฐานของไทย อันจะนำไปสู่การยกระดับคุณภาพการวัดและประเมินผลให้เทียบเท่ามาตรฐานสากลต่อไป

วัตถุประสงค์

1. เพื่อพัฒนา ตรวจสอบความตรงเชิงโครงสร้างของโมเดลการวัดสมรรถนะ และหาคุณภาพของคลังข้อคำถาม สมรรถนะการจัดการและความรับผิดชอบ ของนักเรียนชั้นมัธยมศึกษาตอนปลาย
2. เพื่อพัฒนาเกณฑ์จุดระดับสมรรถนะคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ ของนักเรียนชั้นมัธยมศึกษาตอนปลาย

นิยามศัพท์เฉพาะ

1. สมรรถนะนักเรียน หมายถึง การใช้ ความรู้ ทักษะ และคุณลักษณะต่าง ๆ ที่ประกอบขึ้นร่วมกันภายในตัวบุคคล ผ่านสถานการณ์หรืองาน ที่เราให้นักเรียนแสดงออกมาซึ่งสมรรถนะ และมีเกณฑ์ความสำเร็จที่เป็นเป้าหมายในการวัดระดับสมรรถนะว่าอยู่ในระดับใด และหากจะทำได้ดียิ่งขึ้น หรือมีประสิทธิภาพมากขึ้น ก็ขึ้นอยู่กับการสั่งสมสมรรถนะ
2. สมรรถนะการจัดการและความรับผิดชอบ (Management and taking responsibility Competency) หมายถึง ความสามารถในการตัดสินใจความสามารถและการกำกับตนเอง เพื่อให้สามารถบรรลุเป้าหมายที่กำหนดไว้ สามารถจัดการกับความไม่แน่นอน พร้อมทั้งรับมือกับความซับซ้อน จัดการปัญหาและภาวะวิกฤตได้อย่างครอบคลุมทุกบริบท สามารถฟื้นคืนสู่สภาวะสมดุล (Resilience) เพื่อไปสู่ความสำเร็จของเป้าหมายในชีวิต มีสุขภาวะที่ดีและมีสัมพันธภาพกับผู้อื่นได้ดีสามารถวางแผนและจัดระเบียบการดำเนินงานให้สำเร็จตามเป้าหมายได้อย่างมีประสิทธิภาพ และสามารถพิจารณาผลที่ตามมาจากการกระทำและพร้อมที่จะรับผิดชอบต่อผลที่ตามานั้น และมีส่วนร่วมในการปรับปรุงพัฒนาความเป็นอยู่ของคนในสังคม ประกอบด้วย 4 องค์ประกอบย่อย 14 พฤติกรรมบ่งชี้

วิธีดำเนินการวิจัย

ขอบเขตเนื้อหา

จากการสังเคราะห์ข้อมูลและแนวคิดจากงานวิจัยและกรอบสมรรถนะระดับสากล พบว่า สมรรถนะการจัดการและความรับผิดชอบ หมายถึง ความสามารถในการตัดสินใจและกำกับตนเองเพื่อให้บรรลุเป้าหมายที่ตั้งไว้ สามารถจัดการกับความไม่แน่นอนและรับมือกับความซับซ้อน จัดการปัญหาและภาวะวิกฤตในทุกบริบทได้อย่างมีประสิทธิภาพ พร้อมฟื้นคืนสู่สภาวะสมดุล (resilience) เพื่อมุ่งสู่ความสำเร็จของเป้าหมาย มีสุขภาวะที่ดี และสร้างสัมพันธภาพเชิงบวกกับผู้อื่นได้ สามารถวางแผนและจัดระเบียบการดำเนินงานให้สำเร็จตามเป้าหมาย รวมถึงพิจารณาผลที่ตามมาจากการกระทำและยอมรับความรับผิดชอบต่อผลลัพธ์นั้น อีกทั้งยังมีส่วนร่วมในการพัฒนาความเป็นอยู่ของสังคม ในการวิจัยครั้งนี้ ผู้วิจัยได้กำหนดโครงสร้างสมรรถนะการจัดการและความรับผิดชอบให้ครอบคลุม 4 องค์ประกอบ 14 พฤติกรรมบ่งชี้ ได้แก่ 1) การรับรู้ความสามารถตนเอง (5 พฤติกรรมบ่งชี้) 2)การจัดการกับความตึงเครียดและความไม่แน่นอน (3 พฤติกรรมบ่งชี้) 3)การวางแผนและการจัดการ

ระเบียบ (3 พฤติกรรมบ่งชี้) และ 4) ความรับผิดชอบ (3 พฤติกรรมบ่งชี้) องค์ประกอบและพฤติกรรมบ่งชี้ดังกล่าวถูกนำมาเป็นพื้นฐานในการสร้างคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ โดยกำหนดข้อคำถามจำนวน 6 ข้อต่อ 1 พฤติกรรมบ่งชี้ รวมเป็นข้อคำถามทั้งหมด 84 ข้อ เพื่อให้สามารถประเมินสมรรถนะได้อย่างครอบคลุมและแม่นยำ และเพียงพอในแต่ละพฤติกรรมบ่งชี้เพื่อรองรับการทดสอบแบบปรับเหมาะ (Computerized Adaptive Testing: CAT) นำเสนอ ดัง Figure 1

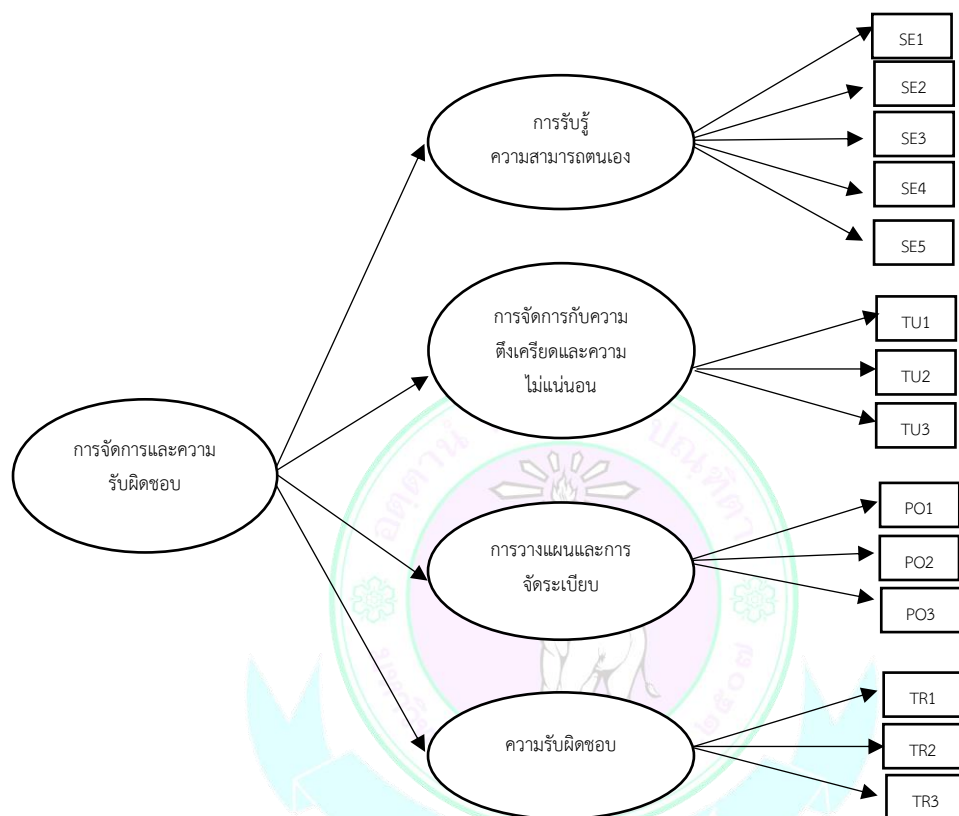


Figure 1 โมเดลการวัดของคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ

ประชากร และกลุ่มตัวอย่างวิจัย

1) ประชากร

นักเรียนชั้นมัธยมศึกษาตอนปลาย ระดับชั้นมัธยมศึกษาปีที่ 4-6 สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ ในภาคตะวันออกเฉียงเหนือ ในภาคเรียนที่ 2 ปีการศึกษา 2566

2) กลุ่มตัวอย่าง

การวิจัยครั้งนี้ใช้วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันในการตรวจสอบความตรงของโมเดลการวัด และการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ ในการตรวจสอบคุณภาพของเครื่องมือ กลุ่มตัวอย่างจึงควรมีขนาดใหญ่และการแจกแจงประชากรต้องเป็นโค้งปกติ ตามแนวคิดของ Hair, Anderson, Black, Babin, Aderson & Tatham (2006) กำหนดขนาดกลุ่มตัวอย่าง 5-20 เท่าของของจำนวนพารามิเตอร์ De Ayala (2009, pp. 223-224) ได้เสนอว่า ขนาดกลุ่มตัวอย่างที่ใช้การวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ (IRT) โมเดล Graded Response Model (GRM) ในการประมาณค่าพารามิเตอร์ในโมเดล GRM ควรมีไม่ต่ำกว่า 500 คน ดังนั้น ผู้วิจัยจึงกำหนดกลุ่มตัวอย่างเพื่อให้เพียงพอต่อการวิเคราะห์ข้อมูลและการเป็นตัวแทนของประชากร เป็นนักเรียนชั้นมัธยมศึกษาตอนปลาย จำนวน 900 คน โดยกลุ่มตัวอย่างได้มาจากการสุ่มแบบหลายขั้นตอน

ขั้นที่ 1 สุ่มอย่างง่ายโดยให้จังหวัดเป็นหน่วยสุ่ม ได้ 4 จังหวัดที่เป็นกลุ่มตัวอย่าง ขั้นที่ 2 จากนั้นสุ่มอย่างง่ายโดยใช้สถานศึกษาในแต่ละจังหวัดเป็นหน่วยสุ่ม โดยเป็นสถานศึกษาในเมืองและนอกเขตอำเภอเมือง ได้ 8 สถานศึกษา และขั้นที่ 3 สุ่มแบบยกกลุ่มโดยใช้ห้องเรียนของสถานศึกษาเป็นหน่วยสุ่ม ได้จำนวน 20 ห้องเรียน โดยจำนวนกลุ่มตัวอย่างที่เก็บข้อมูลได้จริงจำนวน 832 คน (คิดเป็นร้อยละ 92.44) ซึ่งมากกว่าจำนวนตัวอย่างขั้นต่ำที่ต้องการตามเกณฑ์ โดยมีข้อมูลขาดหายไปไม่เกินร้อยละ 30

เครื่องมือวิจัย

การวิจัยในครั้งนี้ใช้เครื่องมือ คือ คลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ ชั้นมัธยมศึกษาตอนปลาย โดยมีเกณฑ์การคัดเลือก ตัวแปรที่มีความใหม่จากหน่วยงานและการศึกษาวิจัยในไทยและนานาชาติ เพื่อรับการเปลี่ยนแปลงและการพัฒนาสมรรถนะนักเรียนให้เท่าทันโลก เพื่อให้สอดคล้องกับการประเมินสมรรถนะที่เป็น content free ไม่ขึ้นอยู่กับเนื้อหาของการเรียน โดยเป็นการให้นักเรียนรายงานการปฏิบัติของตนเอง ตามระดับการปฏิบัติสมรรถนะของตนเอง ผ่านข้อคำถามเชิงสถานการณ์เป็นแบบวัดเชิงสถานการณ์ 4 ตัวเลือก ที่มีการตรวจให้คะแนนมากกว่า 2 ค่า ตามระดับพฤติกรรมตามทฤษฎีทางด้านจิตพิสัยของ Krathwohl และคณะ (Krathwohl et al., 1964) เพื่อให้ได้ข้อมูลที่ตรงกับพฤติกรรมการปฏิบัติของนักเรียนมากที่สุด โดยมีระดับการให้คะแนนตัวเลือก 1-4 คะแนน ตามระดับการปฏิบัติสมรรถนะของนักเรียน ดังนี้ ระดับ 1: ความตระหนักรู้ (คล้ายกับการรับ) ระดับ 2: ความสนใจ (คล้ายกับการตอบสนอง) ระดับ 3: ความสำคัญ (คล้ายกับการให้คุณค่า) และระดับ 4: การกระทำ (ผสมผสานการจัดระเบียบและการกำหนดลักษณะ) โดยผู้วิจัยพัฒนาข้อคำถาม 6 ข้อ ต่อ 1 พฤติกรรมบ่งชี้ เพื่อพัฒนาเป็นคลังข้อคำถามสมรรถนะ รวมจำนวน 84 ข้อ ใช้เวลาทำแบบวัด 30 - 40 นาที

การพัฒนาเครื่องมือวิจัย

1) พัฒนาชุดเครื่องมือที่เป็นร่างข้อคำถามตามพฤติกรรมบ่งชี้ขององค์ประกอบของสมรรถนะ จากโมเดลการวัดที่ผู้วิจัยได้ศึกษาจากหน่วยงาน องค์กรทางการศึกษา รวมถึงงานวิจัยทั้งในประเทศไทยและต่างประเทศ โดยสังเคราะห์จากความหมายและพัฒนาพฤติกรรมบ่งชี้ตามความหมายของพฤติกรรมการปฏิบัติหรือคำที่มีความหมายคล้ายกัน (OECD, 2019; OECD Future of Education and Skills 2030, 2019; Partnership for 21st Century Skills, 2019; European Commission, 2018; World Economic Forum, 2020; OECD, 2018; UNESCO, 2016; Office of the Basic Education Commission, 2021; Phrommaboon et al., 2020; CBE Thailand, 2021; Tuksino et al., 2016) จากนั้นเขียนข้อคำถามตามพฤติกรรมบ่งชี้ โดยพัฒนาข้อคำถามเชิงสถานการณ์ที่สอดคล้องกับสมรรถนะและบริบทของนักเรียน พฤติกรรมบ่งชี้ละ 6 ข้อ เพื่อให้ครอบคลุมทุกระดับความสามารถ โดยคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบมีทั้ง 84 ข้อ ข้อคำถามมี 4 ตัวเลือก ให้คะแนน 1-4 คะแนน ตามระดับความสามารถที่แสดงออกในแต่ละสถานการณ์ ตามหลักพฤติกรรม Krathwohl's Affective Domain (1964)

2) นำร่างชุดเครื่องมือวัดสมรรถนะไปให้ผู้เชี่ยวชาญจำนวน 5 ท่าน ประกอบไปด้วยผู้เชี่ยวชาญทางการวัดและประเมินผลจำนวน 3 ท่าน ผู้เชี่ยวชาญทางด้านการพัฒนาสมรรถนะจำนวน 2 ท่าน โดยมีเกณฑ์การคัดเลือกคือต้องจบการศึกษาตั้งแต่ระดับปริญญาเอกขึ้นไป และมีประสบการณ์ทางด้านการวัดและประเมินผลและด้านการพัฒนาสมรรถนะ

3) วิเคราะห์ความตรงเชิงเนื้อหา โดยนำผลการลงความเห็นของผู้ทรงคุณวุฒิว่าข้อคำถามแต่ละข้อวัดได้ตามนิยามที่กำหนดหรือไม่ แล้วนำมาคุณภาพด้านความตรงเชิงเนื้อหาด้วยดัชนี CVR (Content validity ratio) แล้วคัดเลือกข้อคำถามที่มีค่าดัชนี CVR มากกว่า 0.50 (Lawshe, 1975) พบว่าทุกข้อคำถามสอดคล้องกับพฤติกรรมบ่งชี้ มีค่าดัชนี CVR 1.00 หลังจากนั้นผู้วิจัยปรับค่าข้อความและการตรวจให้คะแนนของตัวเลือกเพื่อให้เกิดความชัดเจน จากข้อเสนอแนะของผู้ทรงคุณวุฒิ แล้วจึงจัดทำคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ

4) ดำเนินการขอจริยธรรมการวิจัยในมนุษย์ จากศูนย์จริยธรรมการวิจัยในมนุษย์ มหาวิทยาลัยขอนแก่น โดยงานวิจัยนี้ได้ผ่านการรับรองจากคณะกรรมการจริยธรรมการวิจัยในมนุษย์ มหาวิทยาลัยขอนแก่น โดยยึดหลักเกณฑ์ตามรายงานเบลมอนต์ และ GCP in Social and Behavioral Research เลขที่โครงการ HE663399

5) นำคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ จำนวน 84 ข้อ ที่มีช่องของการเขียนข้อเสนอแนะ ไปทดลองใช้แบบออนไลน์ผ่าน Google form กับกลุ่มวัยเดียวกันกับกลุ่มตัวอย่างจริง โดย Javali et al. (2011) กล่าวว่าขนาดตัวอย่างที่เพียงพอต่อการทดสอบความเชื่อมั่น (Reliability) คือ มีจำนวน 50 คนขึ้นไป ผู้วิจัยจึงต้องกำหนดขนาดของกลุ่มทดลองใช้เครื่องมือ จำนวน 90 คน (Midren L Patten, 2017) พบว่า คลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ มีค่าความเชื่อมั่นโดยใช้สูตรสัมประสิทธิ์แอลฟา (α -Coefficient) ของครอนบาช (Cronbach) มีค่าเท่ากับ 0.91 มีค่าอำนาจจำแนกตามทฤษฎีการทดสอบแบบดั้งเดิมตั้งแต่ 0.2 ขึ้นไป ทุกข้อ

6) นำรายละเอียดของการทดลองใช้กับใช้กับกลุ่มวัยเดียวกันกับกลุ่มตัวอย่างจริง มาพัฒนา ปรับปรุงข้อคำถามด้านการใช้ภาษา สถานการณ์คำถาม และความชัดเจนในการเขียนตัวเลือกให้สามารถแบ่งระดับของพฤติกรรมได้อย่างชัดเจน เวลาที่ใช้ในการตอบข้อคำถาม เพื่อให้เกิดความชัดเจนและสมบูรณ์ จากนั้นนำไปให้อาจารย์ที่ปรึกษาตรวจสอบความสมบูรณ์อีกครั้ง แล้วจัดทำคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบฉบับสมบูรณ์ จำนวน 84 ข้อ เพื่อนำไปใช้เก็บข้อมูลต่อไป

การวิเคราะห์ข้อมูล

1) วิเคราะห์คุณภาพของเครื่องมือทั้งฉบับ ด้านความตรงเชิงโครงสร้างโดยการตรวจสอบความตรงตามโมเดลจากการวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis: CFA) ด้วยโปรแกรมการวิเคราะห์ Mplus 8.0

2) วิเคราะห์คุณภาพของข้อคำถามรายข้อ เพื่อคัดเลือกข้อคำถามจัดทำคลังข้อคำถาม (Item Bank) ด้วยการหาค่าอำนาจจำแนกด้วยค่าพารามิเตอร์ความชันร่วมของข้อคำถาม (α) ค่าความยากด้วยค่าพารามิเตอร์เทรชโฮลด์ ที่เป็นไปตามเกณฑ์การคัดเลือก ตามแนวคิดของทฤษฎีการตอบสนองข้อคำถาม (IRT) โมเดล Graded Response Model (GRM) ด้วยโปรแกรม R Studio version 4.4.2

3) วิเคราะห์ความเชื่อมั่นของเครื่องมือทั้งฉบับและรายองค์ประกอบ ด้วยวิธีการวิเคราะห์สัมประสิทธิ์แอลฟา (α -Coefficient) ของครอนบาช (Cronbach) โดยใช้ทฤษฎีแบบดั้งเดิม ด้วยโปรแกรม SPSS version 28 และวิเคราะห์ความเชื่อมั่นของเครื่องมือทั้งฉบับด้วยทฤษฎีการตอบสนองข้อสอบ ด้วยการวิเคราะห์ ฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Function: TIF) ด้วยโปรแกรม R Studio version 4.4.2

4) การวิเคราะห์เกณฑ์จุดตัดระดับสมรรถนะการจัดการและความรับผิดชอบ ด้วยเทคนิคการกำหนดคะแนนจุดตัดโดยใช้เกณฑ์พื้นที่จากพารามิเตอร์ Threshold บน Wright Map เพื่อจำแนกระดับสมรรถนะของนักเรียนให้อยู่ในกลุ่มที่มีลักษณะแตกต่างกันตามระดับที่กำหนดไว้ โดยใช้ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ดำเนินการวิเคราะห์ข้อมูลการตอบของนักเรียนด้วยโปรแกรม R แพ็กเกจ mirt เพื่อหาค่าพารามิเตอร์ threshold ของแต่ละข้อ จากนั้นสร้าง Wright Map แผนภาพแสดงการกระจายระดับความสามารถของนักเรียนร่วมกับพารามิเตอร์ threshold บนสเกลเดียวกัน ทำให้เห็นว่าแต่ละพฤติกรรมของผู้ตอบสัมพันธ์กับระดับความยากของตัวเลือกใด แล้วคำนวณค่าเฉลี่ยของ threshold แต่ละระดับนำค่าพารามิเตอร์ threshold ที่ 1 (b1), ที่ 2 (b2) ฯลฯ ของทุกข้อ มาหาค่าเฉลี่ยเพื่อกำหนดจุดตัดระหว่างระดับสมรรถนะ ด้วยโปรแกรม R Studio version 4.4.2

ผลการวิจัย

1. ผลการพัฒนาคำถามสัมภาษณ์สมรรถนะการจัดการและความรับผิดชอบ ของนักเรียนชั้นมัธยมศึกษาตอนปลาย

1.1) คลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ ของนักเรียนชั้นมัธยมศึกษาตอนปลาย

คลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ ของนักเรียนชั้นมัธยมศึกษาตอนปลาย ประกอบด้วย 4 องค์ประกอบ 14 พฤติกรรมบ่งชี้ ได้แก่ 1) การรับรู้ความสามารถตนเอง จำนวน 5 พฤติกรรมบ่งชี้ ได้แก่ SE1) การประเมินตนเองตามสภาพจริง และกำกับตนเองเพื่อให้บรรลุผลตามเป้าหมายที่ได้กำหนดไว้ SE2) สามารถฟื้นคืนสู่สภาวะสมดุล (Resilience) เพื่อไปสู่ความสำเร็จของเป้าหมาย SE3) มีแรงจูงใจ กระตือรือร้นและเพียรพยายามที่จะพัฒนาความสามารถของตนเอง SE4) สามารถดูแลตนเองทั้งร่างกายและจิตใจให้แข็งแรงและมีความสุข และ SE5) สร้างมนุษยสัมพันธ์ ทำงานร่วมกับผู้อื่น ได้เหมาะสมกับบทบาทและหน้าที่ 2) การจัดการกับความตึงเครียดและความไม่แน่นอน จำนวน 3 พฤติกรรมบ่งชี้ ได้แก่ TU1) คาดการณ์ความเสี่ยงและอุปสรรค และยอมรับเหตุการณ์ฉุกเฉินที่อาจเกิดขึ้นได้นอกเหนือจากวางแผน TU2) สามารถจัดการ ภาวะที่กลืนไม่เข้าคายไม่ออก และ TU3) จัดการปัญหาและภาวะวิกฤตได้อย่างครอบคลุมทุกบริบท 3) องค์ประกอบด้านการวางแผนและจัดการระเบียบ จำนวน 3 พฤติกรรมบ่งชี้ ได้แก่ PO1) การวางแผนจัดระบบการทำงาน PO2) กำหนดแนวทางปฏิบัติที่เป็นจริงและเป็นระบบเพื่อให้บรรลุวัตถุประสงค์ และ PO3) สามารถบริหารจัดการทรัพยากรได้อย่างมีประสิทธิภาพ และ 4) องค์ประกอบด้านความรับผิดชอบ จำนวน 3 พฤติกรรมบ่งชี้ ได้แก่ TR1) รับผิดชอบต่อการส่งมอบผลลัพธ์ที่เป็นเป้าหมาย TR2) สามารถพิจารณาผลที่จะตามมาในอนาคตของการกระทำ เพื่อประเมินความเสี่ยงผลตอบแทน และเพื่อยอมรับความรับผิดชอบในผลงานของตน และ TR3) รับผิดชอบต่อประเด็นหรือสถานการณ์ของท้องถิ่น โลก หรือวัฒนธรรมที่แตกต่าง และมีส่วนร่วมในการปรับปรุงและพัฒนา ซึ่งผู้วิจัยได้นำองค์ประกอบและพฤติกรรมบ่งชี้ดังกล่าวมาเป็นข้อมูลในการสร้างคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ โดยสร้างข้อคำถาม 6 ข้อ ต่อ 1 พฤติกรรมบ่งชี้ เพื่อสามารถพัฒนาเป็นคลังข้อคำถามรวมทั้งหมด 84 ข้อ

1.2) ผลการวิเคราะห์ความตรงเชิงโครงสร้าง (construct validity) โดยการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง (confirmatory factor analysis) ของโมเดลการวัดสมรรถนะการจัดการและความรับผิดชอบ ของนักเรียนชั้นมัธยมศึกษาตอนปลาย และการวิเคราะห์นำเสนอตามรายละเอียด ดังนี้

1.2.1) ผลการค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรสังเกตได้ทั้งหมดมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยตัวแปรสังเกตได้มีความสัมพันธ์ทางบวกทุกคู่ มีค่าดัชนี Kaiser - Meyer - Olkin (KMO) และค่าสถิติ Bartlett's test of Sphericity พบว่า มีค่าเท่ากับ .955 และ 4867.196 ($p < 0.01$) ตามลำดับ แสดงให้เห็นว่าจำนวนกลุ่มตัวอย่างที่ใช้มีความเหมาะสมในการวิเคราะห์องค์ประกอบของข้อมูลชุดนี้ ตัวแปรแต่ละตัวมีความสัมพันธ์กัน และเหมาะสมที่จะนำไปใช้ในการพัฒนาและตรวจสอบความตรงเชิงโครงสร้างของคลังข้อคำถามต่อไป

1.2.2) ผลการวิเคราะห์ความตรงเชิงโครงสร้าง (construct validity) โดยการวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor analysis) ของโมเดลการวัดสมรรถนะการจัดการและความรับผิดชอบ ของนักเรียนชั้นมัธยมศึกษาตอนปลาย พบว่า มี 4 องค์ประกอบ มีค่าสัมประสิทธิ์น้ำหนักองค์ประกอบอยู่ระหว่าง 0.927 ถึง 0.972 และสามารถอธิบายสมรรถนะการจัดการและความรับผิดชอบ ได้ร้อยละ 85.9 ถึง 94.4 ดัง Table1

Table1 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลการวัดสมรรถนะการจัดการและความรับผิดชอบ

องค์ประกอบย่อย	เมทริกซ์น้ำหนักองค์ประกอบ			R ²
	β	SE	t	
การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับ 1				
การรับรู้ความสามารถตนเอง (SE)				
SE1	0.698	0.021	32.021	0.488
SE2	0.639	0.024	26.811	0.409
SE3	0.653	0.023	28.627	0.426
SE4	0.624	0.024	25.801	0.389
SE5	0.680	0.022	31.146	0.462
การจัดการกับความตึงเครียดและความไม่แน่นอน (TU)				
TU1	0.583	0.026	22.556	0.340
TU2	0.770	0.018	42.185	0.593
TU3	0.728	0.020	36.662	0.530
การวางแผนและจัดการระเบียบ (PO)				
PO1	0.435	0.032	13.693	0.190
PO2	0.603	0.026	22.838	0.364
PO3	0.739	0.023	32.441	0.547
ความรับผิดชอบ (TR)				
TR1	0.775	0.018	43.622	0.601
TR2	0.739	0.019	38.256	0.546
TR3	0.735	0.019	37.912	0.459
การวิเคราะห์องค์ประกอบเชิงยืนยันอันดับ 2				
การรับรู้ความสามารถตนเอง (SE)	0.958	0.014	69.512	0.918
การจัดการกับความตึงเครียดและความไม่แน่นอน (TU)	0.969	0.015	65.788	0.939
การวางแผนและจัดการระเบียบ (PO)	0.972	0.022	44.132	0.944
ความรับผิดชอบ (TR)	0.927	0.014	65.378	0.859

เมื่อพิจารณาค่าดัชนีความสอดคล้องของโมเดลการวัดสมรรถนะการจัดการและความรับผิดชอบ ของนักเรียนชั้นมัธยมศึกษาตอนปลาย พบว่าโมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ ($\chi^2=73.874$, df = 57, p-value = 0.0661, RMSEA=0.0019, SRMR= 0.016, CFI= 0.997, TLI = 0.994) หมายความว่าคลังสมรรถนะการจัดการและความรับผิดชอบ มีความตรงเชิงโครงสร้าง

1.3) ผลการวิเคราะห์คุณภาพของข้อคำถามรายข้อ ด้วยทฤษฎีการตอบสนองข้อสอบ

ผลการวิเคราะห์คุณภาพของข้อคำถามรายข้อของคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ ด้วยการวิเคราะห์คุณภาพข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) โมเดล Graded Response Model (GRM) แบบตรวจให้คะแนนมากกว่า 2 ค่า (Polychomous IRT models) เป็นการประมาณค่าพารามิเตอร์ของรายการคำถามจากมาตรประมาณค่า 4 ระดับ ด้วยพิจารณาค่าอำนาจจำแนก (Discrimination Parameter: a) ด้วยค่าพารามิเตอร์ความชันร่วมของข้อคำถาม (Slope Parameter: α) และค่าความยาก (b) ด้วยค่าพารามิเตอร์เทรชโฮลด์ (β : Threshold parameters) ที่เป็นไปตามเกณฑ์การคัดเลือก พบว่า คลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ มีข้อคำถามที่ผ่านเกณฑ์ เพื่อคัดเลือกเข้าสู่คลังข้อคำถาม (Item bank) จำนวน 84 ข้อ คิดเป็นร้อยละ 100 โดยพบว่า ค่าพารามิเตอร์ความชันร่วมของข้อคำถาม (α) มีค่าตั้งแต่ 0.516–1.891 มีค่าเฉลี่ยอยู่ที่ 0.904 แสดงให้เห็นว่าแบบวัดมีข้อสอบที่สามารถจำแนกระดับความสามารถของผู้สอบได้ในระดับที่ดี โดยเฉพาะในช่วงระดับความสามารถปานกลางถึงสูงและมีค่าพารามิเตอร์เทรชโฮลด์

(Threshold parameters: $\beta_1, \beta_2, \beta_3$) เฉลี่ยอยู่ที่ -0.281, 0.906 และ 2.122 ตามลำดับ ครอบคลุมช่วงความสามารถของผู้สอบอย่างเหมาะสม (-3 ถึง +3) นำเสนอตั้ง Table 2

Table 2 แสดงค่าพารามิเตอร์ความชันร่วมของข้อคำถาม (α) และค่าพารามิเตอร์เทรชโฮลด์ ($\beta_1, \beta_2, \beta_3$)

จากการวิเคราะห์ทฤษฎีการตอบสนองข้อสอบ (IRT) โมเดล Graded Response Model (GRM) คลังข้อคำถาม
สมรรถนะการจัดการและความรับผิดชอบ

ข้อคำถาม	α	β_1	β_2	β_3	ความยาก	ผ่านเกณฑ์เข้าคลัง
Item_1	0.549	0.275	0.958	2.776	ปานกลางค่อนข้างไปทางยาก	ผ่าน
Item_2	1.143	0.983	1.868	2.745	ยาก	ผ่าน
Item_3	1.488	1.133	1.921	3.709	ยากมาก	ผ่าน
.	...	...	...	...	...	...
Item_82	1.606	0.580	0.977	2.167	ปานกลางค่อนข้างไปทางยาก	ผ่าน
Item_83	0.635	-0.398	1.132	2.108	ง่ายค่อนข้างไปทางปานกลาง	ผ่าน
Item_84	0.672	-0.360	1.368	2.134	ง่ายค่อนข้างไปทางปานกลาง	ผ่าน

1.4) ผลการวิเคราะห์คุณภาพของข้อคำถามทั้งฉบับด้วยทฤษฎีแบบดั้งเดิมและทฤษฎีการตอบสนอง

ข้อสอบ

1.4.1) การตรวจสอบความเชื่อมั่น (Reliability) ของคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบของนักเรียนชั้นมัธยมศึกษาตอนปลาย ด้วยทฤษฎีแบบดั้งเดิม โดยการวิเคราะห์สัมประสิทธิ์แอลฟา (α -Coefficient) ของครอนบาช (Cronbach) พบว่า คลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ จำนวน 84 ข้อของนักเรียนชั้นมัธยมศึกษาตอนปลาย มีค่าความเชื่อมั่นทั้งฉบับเป็น 0.925 เมื่อพิจารณารายองค์ประกอบพบว่ามีความเชื่อมั่นระหว่าง 0.674- 0.824 Table 3

Table 3 แสดงค่าความเชื่อมั่นของคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ ของนักเรียนชั้น

มัธยมศึกษาตอนปลายตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT)

สมรรถนะ	องค์ประกอบ	ค่าความเชื่อมั่น (α)
การจัดการและความ รับผิดชอบ	1) การรับรู้ความสามารถตนเอง (SE)	0.824
	2) การจัดการกับความตึงเครียดและความไม่แน่นอน(TU)	0.755
	3) การวางแผนและจัดการระเบียบ (PO)	0.674
	4) ความรับผิดชอบ (TR)	0.806
	รวม	0.925

1.4.2) การตรวจสอบความเชื่อมั่น (Reliability) ของคลังข้อคำถามสมรรถนะการจัดการนักเรียนชั้นมัธยมศึกษาตอนปลาย ด้วยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT)

เมื่อพิจารณาค่าความเชื่อมั่นของเครื่องมือทั้งฉบับจากกราฟฟังก์ชันสารสนเทศ (Test Information Function: TIF) ตามการวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) โมเดล Graded Response Model (GRM) ของคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ จากภาพที่ 4 พบว่า แบบวัดมีค่าฟังก์ชันสารสนเทศสูงสุดประมาณ 55 หน่วย โดยมีจุดยอดของกราฟอยู่ที่ค่า $\theta \approx 1.5$ แสดงว่าเครื่องมือนี้มีความแม่นยำมากที่สุดในการวัดนักเรียนที่มีสมรรถนะในระดับสูงกว่าค่าเฉลี่ย โดยสามารถให้ข้อมูลที่มีคุณภาพในช่วงความสามารถ

ตั้งแต่ $\theta \approx -1.0$ ถึง $+3$ ช่วงของค่า θ ที่ให้สารสนเทศสูงนี้สะท้อนว่า คลังข้อคำถามฉบับนี้มีความเหมาะสมอย่างยิ่งสำหรับการใช้งานกับกลุ่มนักเรียนที่มีแนวโน้มแสดงความรับผิดชอบสูง แสดงได้ดัง Figure 2

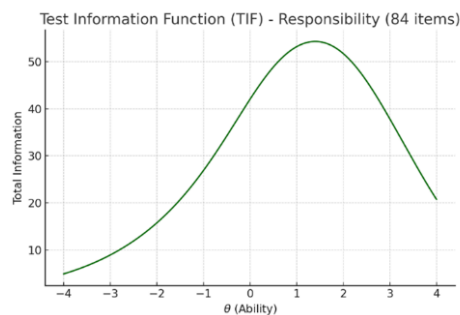


Figure 2 กราฟแสดงฟังก์ชันสารสนเทศของคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ

2) ผลการพัฒนาเกณฑ์จุดตัดระดับสมรรถนะการจัดการและความรับผิดชอบ

ผลการพัฒนาเกณฑ์จุดตัดระดับสมรรถนะการจัดการและความรับผิดชอบ ด้วยเทคนิคการกำหนดคะแนนจุดตัดที่ใช้เกณฑ์พื้นที่จากพารามิเตอร์ Threshold บนแผนภาพ Wright Map เพื่อให้จำแนกระดับสมรรถนะของนักเรียนให้อยู่ในกลุ่มที่มีลักษณะแตกต่างกันตามระดับที่กำหนดไว้ โดยนิยามระดับสมรรถนะ ไว้ 4 ระดับ ได้แก่ 1) มีสมรรถนะไม่เพียงพอ 2) เริ่มมีสมรรถนะ 3) มีสมรรถนะ และ 4) มีสมรรถนะสูง (Bates, 2019; Colby, 2021; Hess et al., 2020) ที่ และกำหนดคำอธิบายให้สามารถสะท้อนพฤติกรรมของนักเรียนได้ ใช้ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ด้วยโมเดล Graded Response Model (GRM) ซึ่งสามารถประมาณค่าความสามารถแฝง (θ) ของนักเรียนได้อย่างแม่นยำ โดยใช้แผนภาพ Wright Map ซึ่งเป็นแผนภาพที่แสดงการกระจายของระดับความสามารถของนักเรียนและพารามิเตอร์ของข้อสอบบนสเกลเดียวกัน โดยนำค่าพารามิเตอร์ threshold มาหาค่าเฉลี่ย และวิเคราะห์ในเชิงภาพบนแผนภาพ Wright Map ที่แสดงการกระจายของ threshold ทั้ง 3 ระดับ (β_1 - β_3) เพื่อแสดงระดับความสามารถที่จำเป็นสำหรับการตอบคำถามในระดับคะแนนที่สูงขึ้น ผลการวิเคราะห์ พบว่า ค่าพารามิเตอร์ threshold โดยเฉลี่ยมีค่า $\beta_1 = 0.96$, $\beta_2 = 1.59$ และ $\beta_3 = 2.73$ ซึ่งสามารถใช้เป็นจุดตัดในการจำแนกระดับสมรรถนะของนักเรียนได้เป็น 4 ระดับ ได้แก่ $\theta \leq -0.28$: ระดับที่ 1 – มีสมรรถนะไม่เพียงพอ, $-0.28 < \theta \leq 0.91$: ระดับที่ 2 – เริ่มมีสมรรถนะ, $0.91 < \theta \leq 2.12$: ระดับที่ 3 – มีสมรรถนะ, และ $\theta > 2.12$: ระดับที่ 4 – มีสมรรถนะสูง แสดง Figure 3 และ Table 4

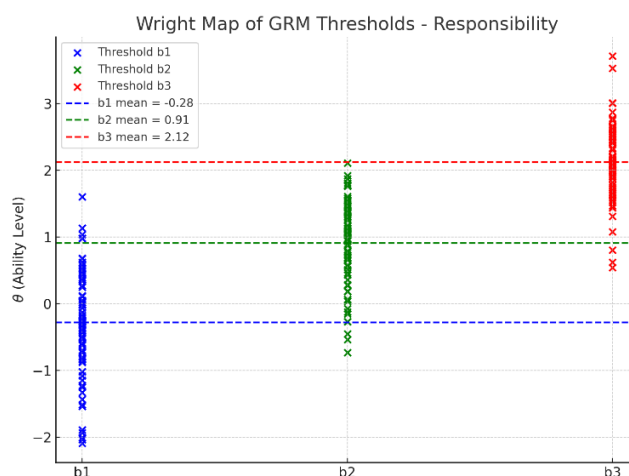


Figure 3 กราฟแสดงการกระจายของค่าพารามิเตอร์ threshold (β_1 , β_2 , β_3)

Table 4 การจำแนกระดับสมรรถนะการจัดการและความรับผิดชอบของนักเรียน ตามช่วงค่าความสามารถแฝง (θ) จาก Wright Map

ค่าความสามารถ (θ)	ระดับสมรรถนะ	ลักษณะพฤติกรรมของนักเรียน
$\theta \leq -0.28$	มีสมรรถนะไม่เพียงพอ	ไม่สามารถวางแผนหรือจัดการภาระงานได้ ขาดการควบคุมตนเอง
$-0.28 < \theta \leq 0.91$	เริ่มมีสมรรถนะ	เริ่มวางแผนและรับผิดชอบต่องานบางส่วน แต่ยังไม่สม่ำเสมอหรือชัดเจน
$0.91 < \theta \leq 2.12$	มีสมรรถนะ	สามารถจัดการงานอย่างเป็นระบบ ตัดสินใจและประเมินผลการทำงานได้ดี
$\theta > 2.12$	มีสมรรถนะสูง	วางแผนเชิงกลยุทธ์ มีวินัย และรับผิดชอบต่อบทบาททั้งตนเองและส่วนรวม

อภิปราย และข้อเสนอแนะ

การวิจัยครั้งนี้เป็นการพัฒนา ตรวจสอบความตรงเชิงโครงสร้าง (construct validity) ของโมเดลการวัด ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory factor analysis) และหาคุณภาพของคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ รวมถึงการวิเคราะห์เกณฑ์จุดตัดระดับสมรรถนะการจัดการและความรับผิดชอบของนักเรียนชั้นมัธยมศึกษาตอนปลาย โดยการประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) โมเดล Graded Response Model (GRM) อภิปรายผลการวิจัยตามผลการวิเคราะห์ข้อมูลคลัสเตอร์กับวัตถุประสงค์การวิจัย ได้ดังนี้

1. การพัฒนาและตรวจสอบความตรงเชิงโครงสร้าง (construct validity) ของโมเดลการวัด ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory factor analysis)

การวิจัยครั้งนี้ได้พัฒนาโมเดลการวัดสมรรถนะการจัดการและความรับผิดชอบ ของนักเรียนชั้นมัธยมศึกษาตอนปลาย ภายใต้ 4 องค์ประกอบย่อย ได้แก่ 1) การรับรู้ความสามารถตนเอง 2) การจัดการกับความตึงเครียดและความไม่แน่นอน 3) การวางแผนและจัดการระเบียบ และ 4) ความรับผิดชอบ (OECD, 2018; OECD, 2019; European Commission, 2018; UNESCO, 2015; World Economic Forum, 2020; Phrommaboon et al., 2020; CBE Thailand, 2021; Tuksino et al., 2016) จำนวน 14 พฤติกรรมบ่งชี้ นำพฤติกรรมบ่งชี้มาพัฒนาเป็นข้อคำถามในคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ ของนักเรียนชั้นมัธยมศึกษาตอนปลาย รวมทั้งหมด 84 ข้อ ใช้เวลาในการทำแบบวัดประมาณ 30-40 นาที

ผลการวิเคราะห์ความตรงเชิงโครงสร้าง (Construct Validity) โดยการวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis) ของคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบของนักเรียนชั้นมัธยมศึกษาตอนปลายที่พัฒนาขึ้น มีค่าดัชนีความสอดคล้องของโมเดลการวัดเป็นไปตามเกณฑ์มาตรฐาน เป็นหลักฐานสำคัญที่ชี้ให้เห็นว่าโครงสร้างสมรรถนะที่พัฒนาขึ้นมีความสอดคล้องกับข้อมูลเชิงประจักษ์ ผลนี้สอดคล้องกับเกณฑ์การประเมินที่ Hu และ Bentler (1999) เสนอไว้ โดยที่ CFI และ TLI ควรมีค่ามากกว่า 0.95, RMSEA ควรมีค่าต่ำกว่า 0.06, และ SRMR ควรมีค่าต่ำกว่า 0.08 และเมื่อพิจารณาค่าน้ำหนักองค์ประกอบ (Factor Loadings) ของทุกข้อคำถามมีค่า > 0.50 ซึ่งสูงกว่าเกณฑ์ที่ Hair et al. (2019) เสนอไว้ แสดงให้เห็นว่าข้อคำถามแต่ละข้อมีความสัมพันธ์กับองค์ประกอบที่กำหนดไว้อย่างชัดเจน และไม่มีข้อคำถามที่ไม่เกี่ยวข้องหรือเป็นสาเหตุของความคลาดเคลื่อนในการวัดซึ่งทำให้สามารถใช้คลังข้อคำถามนี้ในการศึกษาสมรรถนะการจัดการและความรับผิดชอบของนักเรียนในระดับมัธยมศึกษาได้อย่างเหมาะสม

โมเดลการวัดสมรรถนะการจัดการและความรับผิดชอบของนักเรียนชั้นมัธยมศึกษาตอนปลายมีค่าสัมประสิทธิ์น้ำหนักขององค์ประกอบทั้ง 4 องค์ประกอบในช่วง 0.927 ถึง 0.972 เป็นค่าที่สูงและแสดงให้เห็นว่าแต่ละองค์ประกอบมีความสัมพันธ์ที่แข็งแกร่งกับสมรรถนะที่ต้องการประเมิน ซึ่งองค์ประกอบที่มีค่าน้ำหนักสูงสุดคือ การวางแผนและจัดการระเบียบ (PO) ซึ่งมีค่าน้ำหนัก 0.972 ผลการวิเคราะห์พบว่าองค์ประกอบนี้มีความสำคัญสูงสุดในการทำนายสมรรถนะการจัดการและความรับผิดชอบ โดยสะท้อนถึงความสามารถของนักเรียนในการจัดระเบียบและบริหารงานต่าง ๆ อย่างมีประสิทธิภาพ การศึกษาของ Zhang & Liu (2019) พบว่า ความสามารถในการวางแผนและจัดการระเบียบเป็นทักษะสำคัญ

ในการพัฒนานักเรียนในด้านการจัดการเวลาหรือการทำงานในสถานการณ์ที่มีความซับซ้อน ซึ่งสอดคล้องกับผลการศึกษานี้ที่ใช้ให้เห็นว่าองค์ประกอบนี้มีบทบาทสำคัญในการพัฒนาสมรรถนะโดยรวม รองลงมาคือ การจัดการกับความตึงเครียดและความไม่แน่นอน (TU) ซึ่งมีค่าน้ำหนัก 0.969 การศึกษาของ Kim (2018) พบว่า การจัดการกับความตึงเครียดและความไม่แน่นอนช่วยให้นักเรียนมีทักษะในการรับมือกับความท้าทายต่าง ๆ ในชีวิตประจำวัน ซึ่งการศึกษาของ Johnson (2017) ก็ระบุว่า การจัดการกับความเครียดและความไม่แน่นอนมีความสำคัญต่อการตัดสินใจและการพัฒนาอารมณ์ที่สมดุลในช่วงวัยรุ่น ลำดับต่อมา คือ การรับรู้ความสามารถตนเอง (SE) ซึ่งมีค่าน้ำหนัก 0.958 การศึกษาของ Lee & Choi (2020) พบว่า การรับรู้ความสามารถตนเองมีบทบาทสำคัญในการพัฒนาความมั่นใจในตัวเองและทักษะการจัดการที่ดีขึ้น การศึกษานี้สะท้อนให้เห็นว่าเมื่อนักเรียนมีความเชื่อมั่นในความสามารถของตนเอง พวกเขามักจะสามารถทำงานได้อย่างมีประสิทธิภาพและจัดการปัญหาได้ดียิ่งขึ้น และสุดท้ายคือ ความรับผิดชอบ (TR) ซึ่งมีค่าน้ำหนัก 0.927 การศึกษาของ Brown & Miller (2016) พบว่า การรับผิดชอบเป็นคุณลักษณะที่ช่วยให้นักเรียนสามารถทำงานได้ดีขึ้นทั้งในด้านการเรียนและการทำงานกลุ่ม ความรับผิดชอบจึงเป็นองค์ประกอบสำคัญในการพัฒนาทักษะการจัดการโดยรวม โดยผลการวิเคราะห์ในครั้งนี้สามารถอธิบายสมรรถนะการจัดการและความรับผิดชอบได้ตั้งแต่ร้อยละ 85.9 ถึง 94.4 ซึ่งแสดงให้เห็นว่าคลังข้อคำถามนี้มีความแม่นยำสูงในการวัดสมรรถนะดังกล่าว และสามารถใช้งานได้อย่างมีประสิทธิภาพในการประเมินสมรรถนะของนักเรียนในระดับมัธยมศึกษาตอนปลาย

2. การวิเคราะห์หาคุณภาพของคลังข้อคำถาม โดยการประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) โมเดล Graded Response Model (GRM) เป็นแนวทางที่ทันสมัยและให้ข้อมูลที่มีประโยชน์มากกว่าการวิเคราะห์แบบดั้งเดิม โดยผลการวิเคราะห์คลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ ของนักเรียนชั้นมัธยมศึกษาตอนปลาย มีลักษณะเป็นแบบวัดเชิงสถานการณ์ของการปฏิบัติกิจกรรมภายใต้สถานการณ์จำลองของสมรรถนะ มีการตรวจให้คะแนนมากกว่า 2 ค่า มีระดับการให้คะแนนตามระดับพฤติกรรมตามทฤษฎีทางจิตพิสัยของ Krathwohl และคณะ (Krathwohl et al., 1964) ได้แก่ ระดับ 1 ความตระหนักรู้ (คล้ายกับการรับ) ระดับ 2: ความสนใจ (คล้ายกับการตอบสนอง) ระดับ 3 ความสำคัญ (คล้ายกับการให้คุณค่า) และระดับ 4 การกระทำ (ผสมผสานการจัดระเบียบและการกำหนดลักษณะ) มีข้อคำถามที่ผ่านเกณฑ์ร้อยละ 100 เพื่อคัดเลือกเข้าสู่คลังข้อคำถาม (Item bank) โดยมีพารามิเตอร์ค่าความชันร่วมของข้อคำถาม (Slope Parameter: α) ที่เทียบเคียงได้กับพารามิเตอร์ค่าอำนาจจำแนก (Discrimination Parameter: a) อยู่ระหว่าง 0.822-2.681 สอดคล้องกับเกณฑ์ที่ Baker (2001) เสนอไว้ว่าข้อคำถามควรมีค่าความชันมากกว่า 0.7 เพื่อให้สามารถจำแนกความสามารถของผู้สอบได้อย่างเหมาะสม โดยเป็นค่าที่บ่งบอกถึงขนาดความสัมพันธ์ระหว่างข้อคำถามสามารถจำแนกผู้ทดสอบกลุ่มที่มีคุณลักษณะแฝงที่มุ่งวัด คือ สมรรถนะ อยู่ในระดับสูงกับกลุ่มที่มีคุณลักษณะแฝงที่มุ่งวัดอยู่ในระดับต่ำได้ดี เพื่อแสดงความสามารถในการจำแนกผู้สอบที่มีความสามารถต่างกันได้อย่างแม่นยำ (Baker, 2001; Hambleton et al., 1991) ข้อค้นพบที่สำคัญอีกประการหนึ่ง คือ ข้อคำถามส่วนใหญ่มีค่าความชันในช่วง 1.0-2.0 ซึ่งเป็นช่วงที่เหมาะสมที่สุดสำหรับการใช้ในการทดสอบแบบปรับเหมาะ (Computerized Adaptive Testing) ตามที่ Hambleton et al. (1991) ได้อธิบายไว้ว่าข้อคำถามที่มีค่าความชันในช่วงนี้จะให้ข้อมูลสูงสุดเกี่ยวกับความสามารถของผู้สอบ และช่วยให้การประมาณความสามารถมีความแม่นยำมากขึ้น และมีค่าพารามิเตอร์เทรชโฮลด์ (Threshold Parameter: β) ที่เทียบเคียงได้กับพารามิเตอร์ความยาก (Difficulty Parameter: b) โดยค่า β ต้องกระจายครอบคลุมช่วง θ ได้พอสมควร โดยแต่ละรายการคำถามมีค่าเทรชโฮลด์เรียงลำดับจากน้อยไปมาก ซึ่งเป็นไปตามเงื่อนไขของโมเดล Graded Response Model (GRM) ซึ่งบ่งบอกถึงค่าคุณลักษณะแฝงที่ตัวเลือกรายการคำตอบของข้อคำถามแต่ละข้อสามารถวัดได้จากช่วงค่าคุณลักษณะแฝงทั้งหมด มีการกระจายของค่าพารามิเตอร์เทรชโฮลด์ (β) ที่ครอบคลุมช่วงความสามารถที่กว้าง (-3.2 ถึง +3.5) เป็นลักษณะที่พึงประสงค์สำหรับคลังข้อคำถามที่จะนำไปใช้ในการทดสอบแบบปรับเหมาะ เนื่องจากแสดงให้เห็นว่าคลังข้อคำถามสามารถวัดนักเรียนที่มีระดับความสามารถแตกต่างกันได้อย่างเหมาะสม ตั้งแต่ระดับต่ำไปจนถึงระดับสูง ซึ่งสอดคล้องกับหลักการของ

Wainer (2000) ที่เน้นความสำคัญของการมีข้อคำถามที่หลากหลายในระดับความยากโดยอยู่ในช่วง -3 ถึง +3 เพื่อกระจายตัวครอบคลุมช่วงความสามารถของผู้สอบที่หลากหลาย (Samejima 1969; Embretson & Reise, 2000)

3) ความเชื่อมั่นของเครื่องมือทั้งฉบับ ด้วยวิธีการวิเคราะห์สัมประสิทธิ์แอลฟา (α -Coefficient) ของครอนบาช (Cronbach) โดยใช้ทฤษฎีแบบดั้งเดิมของคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ มีค่าความเชื่อมั่น 0.925 ซึ่งสูงกว่าเกณฑ์ 0.7 ที่ Nunnally และ Bernstein (1994) เสนอไว้ แสดงว่า คะแนนที่ได้จากการทำแบบวัดมีความคงเส้นคงวา และเหมาะสมต่อการนำไปใช้ สอดคล้องกับการศึกษาของ Streiner (2003) ที่พบว่าเครื่องมือวัดที่มีค่าความเชื่อมั่นสูงจะให้ผลการวัดที่คงที่และน่าเชื่อถือ เป็นผลมาจากกระบวนการศึกษาและพัฒนา ตั้งแต่การทบทวนทฤษฎี เอกสาร งานวิจัยที่น่าเชื่อถือ อย่างครอบคลุม การกำหนดนิยามปฏิบัติการที่ชัดเจน การสร้างข้อคำถามที่สอดคล้องกับทฤษฎี และการทดสอบความเข้าใจกับกลุ่มเป้าหมาย ซึ่งสอดคล้องกับแนวทางที่ DeVellis (2017) เสนอในการพัฒนาเครื่องมือวัดที่มีคุณภาพ ผู้วิจัยต้องตระหนักและให้ความสำคัญกับกระบวนการได้มาของข้อมูลจากหลายแหล่งที่น่าเชื่อถือและการตรวจสอบคุณภาพในทุกขั้นตอน โดยแบบวัดเชิงสถานการณ์ ได้รับความนิยมนมากขึ้น เนื่องจากสามารถวัดสมรรถนะเชิงพฤติกรรมที่ซับซ้อนได้อย่างมีประสิทธิภาพ เช่น ความสามารถในการตัดสินใจ การจัดการความขัดแย้ง และทักษะทางสังคม โดยมีจุดเด่นในด้านความสมจริง ความยุติธรรม และความสามารถในการทำนายผลลัพธ์ทางพฤติกรรม (Lievens & Sackett, 2012; McDaniel et al., 2007; Ployhart & Ehrhart, 2003; Weekley & Ployhart, 2013; Whetzel & McDaniel, 2009) มีจุดเด่นหลายประการที่เหนือกว่าแบบวัดทั่วไป ทั้งในด้านความสมจริง ความเป็นธรรม และความสามารถในการทำนายพฤติกรรมในบริบทจริง สอดคล้องกับ Chanchusakul (2021) ที่พบว่า เครื่องมือวัดที่มีการให้คะแนนมากกว่า 2 ค่า จะมีค่าเฉลี่ยดัชนีมาตรฐานความเชื่อมั่นสูงกว่า เครื่องมือวัดที่มีการให้คะแนนแบบ 2 ค่า และจากการวิเคราะห์ความเชื่อมั่นของคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบด้วยทฤษฎีการตอบสนองข้อสอบโมเดล Graded Response Model (GRM) พบว่าเครื่องมือ มีค่าฟังก์ชันสารสนเทศรวม (Test Information Function: TIF) สูงสุดประมาณ 55 หน่วย ที่บริเวณค่า $\theta \approx 1.5$ ซึ่งหมายความว่าเครื่องมือนี้มีระดับความแม่นยำ (measurement precision) สูงที่สุดในการประเมินสมรรถนะของนักเรียนที่มีความสามารถด้านนี้สูงกว่าค่าเฉลี่ย นอกจากนี้ยังสามารถให้ข้อมูลที่มีความเชื่อถือได้ในช่วงค่า θ ตั้งแต่ประมาณ -1.0 ถึง +3.0 ซึ่งครอบคลุมนักเรียนระดับปานกลางถึงสูงได้อย่างมีประสิทธิภาพ ค่า TIF ที่สูงแสดงถึง ค่าความเชื่อมั่น (reliability) ของแบบวัดในแต่ละระดับสมรรถนะ โดยตามแนวคิดของ Lord (1980) และ Embretson & Reise (2000) ความแม่นยำในการวัดจะเพิ่มขึ้นเมื่อค่า TIF สูงขึ้น ซึ่งสะท้อนถึงค่าความแปรปรวนของค่าความสามารถที่ต่ำลง กล่าวคือ คลังข้อคำถามสามารถจำแนกระดับสมรรถนะของนักเรียนได้ชัดเจนยิ่งขึ้นเมื่อค่า TIF อยู่ในระดับสูง จากบริบทของนักเรียนมัธยมศึกษาตอนปลาย การที่คลังข้อคำถามให้สารสนเทศสูงในช่วงค่า θ ตั้งแต่ -1 ถึง +3 ถือว่าเหมาะสมกับกลุ่มเป้าหมาย ซึ่งมีแนวโน้มแสดงความรับผิดชอบในระดับปานกลางถึงสูง อันสอดคล้องกับจุดประสงค์ของการประเมินเพื่อส่งเสริมศักยภาพด้านความรับผิดชอบตามกรอบแนวคิดในการพัฒนาสมรรถนะตามหลักสูตรการศึกษาประเทศไทยและกรอบสมรรถนะของนานาชาติ

4) การวิเคราะห์เกณฑ์จุดตัดระดับสมรรถนะการจัดการและความรับผิดชอบ ด้วยการวิเคราะห์ค่าความสามารถแฝง (θ) เพื่อกำหนดเกณฑ์จุดตัดของระดับสมรรถนะ ในงานวิจัยนี้ ได้ใช้วิธี คำนวณค่าเฉลี่ยของพารามิเตอร์ threshold ($\beta_1, \beta_2, \beta_3$) ที่ได้จากการประมาณค่าด้วยโมเดล Graded Response Model (Samejima, 1969) ของทฤษฎีการตอบสนองข้อสอบ ซึ่งสามารถใช้เป็นเกณฑ์ในการแบ่งระดับสมรรถนะของนักเรียนได้อย่างมีนัยสำคัญ โดยอาศัยการแปลความหมายตามช่วงของค่าความสามารถแฝง (theta: θ) บน Wright Map การแปลผลดังกล่าวอ้างอิงแนวคิดจาก Samejima (1997) และ Embretson & Reise (2000) ซึ่งเสนอว่า โมเดล GRM สามารถระบุพฤติกรรมตามระดับความสามารถได้อย่างแม่นยำผ่านพารามิเตอร์ threshold และค่า θ ซึ่งสามารถนำมาใช้ในการจำแนกกลุ่มนักเรียนเชิงพฤติกรรม เพื่อใช้เป็นจุดตัดในการจำแนกระดับความสามารถของนักเรียน แนวทางนี้ได้รับการสนับสนุนจากงานของ de Ayala (2009) และ Wilson (2005) ที่เสนอว่า การนำค่าเฉลี่ยของพารามิเตอร์ threshold ในแต่ละข้อคำถามมาใช้สามารถสร้างเกณฑ์จำแนกระดับสมรรถนะได้

อย่างมีความหมายในเชิงพฤติกรรม โดยเฉพาะเมื่อใช้ร่วมกับการวิเคราะห์ด้วย Wright Map ซึ่งแสดงตำแหน่งของนักเรียน และระดับความยากของแต่ละช่วงคะแนนอย่างเป็นระบบ โดยการใช้ค่าเฉลี่ย threshold เพื่อแบ่งระดับสมรรถนะมีความสำคัญต่อการแปลผลในระดับปฏิบัติการ ซึ่งการใช้ Wright Map ในการกำหนดจุดตัดระดับสมรรถนะเป็นแนวทางที่ได้รับการยอมรับในวงการวัดผลการศึกษา เพราะสามารถแปลผลข้อมูลเชิงสถิติให้เข้าใจง่ายและสอดคล้องกับพฤติกรรมนักเรียนในบริบทจริง (Wilson, 2005; Bond & Fox, 2015) นอกจากนี้ การวิเคราะห์ลักษณะนี้ยังแสดงถึงความโปร่งใสและจริยธรรมในการตีความผล เนื่องจากหลีกเลี่ยงการตีความเกินข้อมูล และใช้เกณฑ์เชิงประจักษ์ที่ตรวจสอบได้ การรายงานผลด้วยความเคารพต่อข้อมูลและผู้ถูกประเมินถือเป็นหลักจริยธรรมพื้นฐานในกระบวนการวิจัยด้านการวัดผลและประเมินผลการศึกษา (AERA, APA, & NCME, 2014) นอกจากนี้ กำหนดเกณฑ์จุดตัดของระดับสมรรถนะ วิธีนี้ยังช่วยให้สามารถออกแบบระบบการประเมินที่ยืดหยุ่นเพื่อนำไปสู่การพัฒนาระบบการประเมินสมรรถนะโดยการประยุกต์ใช้การทดสอบแบบปรับเหมาะด้วยความพิวเตอร์ (Computerized Adaptive Testing: CAT) ซึ่งจำเป็นต้องมีเกณฑ์จุดตัดระดับความสามารถที่แม่นยำและเชื่อมโยงกับข้อมูลเชิงพฤติกรรมของนักเรียนอย่างแท้จริง (Embretson & Reise, 2000; van der Linden & Glas, 2010)

ข้อเสนอแนะในการนำผลวิจัยไปใช้

1. ครูผู้สอนใช้คลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบกับนักเรียนมัธยมปลายทั้ง 3 ช่วง—ก่อนเรียน (diagnostic), ระหว่างเรียน (formative) และปลายภาค (summative) โดยอ้างอิง ระดับสมรรถนะ 4 ระดับจากจุดตัด และใช้ Wright Map ชี้ตำแหน่งจุดอ่อน-จุดแข็งเพื่อออกแบบกิจกรรม/แผนพัฒนารายบุคคล
2. ผู้บริหารสถานศึกษา สามารถใช้ผลการประเมินสมรรถนะที่สรุปผลระดับชั้น/รายห้อง ที่เป็นสัดส่วนแต่ละระดับสมรรถนะและองค์ประกอบ 4 ด้าน เพื่อนำเข้าในการออกแบบโครงการ/กิจกรรม ในแผนปฏิบัติการประจำปี และติดตามผลก่อน-หลังโครงการพัฒนา
3. หน่วยงานระดับนโยบายสามารถใช้ผลประเมินเพื่อตั้งเป้าหมายและออกแบบแนวทางสนับสนุนโรงเรียน โดยจัดทำรายงานเปรียบเทียบผลระหว่างโรงเรียน เพื่อให้เห็นจุดแข็งและจุดที่ควรพัฒนา ช่วยให้การแก้ปัญหาและยกระดับคุณภาพการศึกษาในพื้นที่เป็นไปอย่างตรงจุด
4. ระบบประเมินที่ประยุกต์ใช้การทดสอบแบบปรับเหมาะ (Computerized Adaptive Testing: CAT) สามารถนำคลังข้อคำถามสมรรถนะการจัดการและความรับผิดชอบ จำนวน 84 ข้อ ที่ผ่านการตรวจสอบตามเกณฑ์คุณภาพ เข้าระบบเพื่อเป็นส่วนหนึ่งของกระบวนการทดสอบแบบปรับเหมาะ (Computerized Adaptive Testing: CAT)

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

1. ควรตรวจสอบความเป็นธรรมของข้อสอบ (DIF/Fairness) ด้วยการตรวจการทำหน้าที่ต่างกันของข้อสอบ (DIF) ตามเพศ/ประเภทโรงเรียน/สายการเรียน เพื่อเพิ่มความน่าเชื่อถือของผลการวัดยิ่งขึ้น
2. ควรพัฒนาข้อสถานการณ์เพิ่มและปรับตัวเลือก โดยการทำ cognitive interview/think-aloud กับกลุ่มใกล้เคียงตัวอย่างเพื่อสร้างหรือปรับตัวเลือกให้สมจริง ให้ครอบคลุมช่วงความสามารถ (θ)
3. ควรศึกษาการวิเคราะห์จุดตัดระดับสมรรถนะด้วยวิธีการมาตรฐานที่หลากหลาย เพื่อเปรียบเทียบและประเมินความถูกต้องของการจัดกลุ่มระดับสมรรถนะ

กิตติกรรมประกาศ

งานวิจัยนี้เป็นส่วนหนึ่งของการศึกษาในระดับปริญญาเอก ที่ได้รับทุนวิจัยสำหรับคณาจารย์บัณฑิตศึกษา เพื่อให้สามารถรับนักศึกษาที่มีความสามารถและศักยภาพสูงเข้าศึกษาในหลักสูตรและทำวิจัยในสาขาที่อาจารย์มีความเชี่ยวชาญ (Research grants for graduate faculty to attract highly qualified students into their programs and conduct research in their areas of expertise) จากบัณฑิตวิทยาลัย มหาวิทยาลัยขอนแก่น

References

- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- Baker, F. B. (2001). *The Basics of Item Response Theory*. Washington, DC.: ERIC Clearinghouse on Assessment and Evaluation.
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). New York: Routledge.
- Brown, L., & Miller, K. (2016). Responsibility and Self-Management in Academic Success: A Structural Equation Modeling Approach. *Journal of Educational Research*, 80(3), 252-263.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ.: Lawrence Erlbaum Associates.
- Chanchusakul, S. (2021). A synthesis of research on the development of student competency measurement tools for 21st-century secondary education. *Journal of Educational Measurement, Mahasarakham University*, 27(1), 340-359. [in Thai]
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA.: SAGE Publications.
- De Ayala, R. J. (2009). *The theory and practice of item response theory*. New York: Guilford Press.
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Thousand Oaks, CA: Sage Publications.
- Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. New York: Psychology Press.
- European Commission. (2018). *Key competences for lifelong learning*. Luxembourg: Publications Office of the European Union.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis* (7th ed.). New York: Prentice Hall.
- Hair, J. F. Jr., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2014). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Thousand Oaks, CA.: SAGE Publications.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer- Nijhoff.
- Javali, S., Gudaganavar, N. V., & Raj, S. M. (2011). Effect of varying sample size in estimation of coefficients of internal consistency. *WebmedCentral Biostatistics*, 2(2), WMC001577.

- Johnson, R. (2017). The Importance of Responsibility in Academic Achievement. *International Journal of Educational Research*, 85(1), 48-61.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575.
- Langka, P., et al. (2019). *The development of a national standard toolkit for moral evaluation of students at all ages studying at academic institutions of the ministry of education*. *Journal of Education, Faculty of Education, Srinakharinwirot University*, 20(2), 1-15. [in Thai]
- Lievens, F., & Sackett, P. R. (2012). The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *Journal of Applied Psychology*, 97(2), 460-468.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L., III. (2007). Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology*, 60(1), 63-91.
- Ministry of Education. (1999). *National Education Act, B.E. 2542 (1999)*. Bangkok: Khuru Sapha Lad Phrao. [in Thai]
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- OECD. (2018). *The future of education and skills: Education 2030 – The future we want*. Retrieved from <https://www.oecd.org/education/2030-project/>
- OECD. (2019). *OECD learning compass 2030: A series of concept notes*. Retrieved from <https://www.oecd.org/education/2030-project/learning/learning-compass-2030/>
- OECD. (2024). *Education at a glance 2024: OECD indicators*. Paris: OECD Publishing.
- OECD. (2025). *The state of global teenage career preparation 2025*. Paris: OECD Publishing.
- Office of the Education Council. (2017). *National Education Plan B.E. 2560-2579 (2017-2036)*. Bangkok: Office of the Education Council, Ministry of Education. [in Thai]
- Office of the Education Council. (2021). *Guidelines for the Development of Basic Education Students' Competencies During the Transition to a Competency-Based Curriculum*. Bangkok: Office of the Education Council, Ministry of Education. [in Thai]
- Phommaboon, T., et al. (2020). *Development of a student competency testing system for upper primary level to promote student quality enhancement in the 21st century*. Bangkok: National Institute of Educational Testing Service (Public Organization) [in Thai]
- Ployhart, R. E., & Ehrhart, M. G. (2003). Be careful what you ask for: Effects of response instructions on the construct validity and reliability of situational judgment tests. *International Journal of Selection and Assessment*, 11(1), 1-16.
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, 34(4), 1-97.
- Srisakda, B. (2015). *Development of competency assessment tools for upper secondary*

- students according to the Basic Education Core Curriculum B.E. 2551(2008) (Doctoral dissertation). Faculty of Education, Chulalongkorn University. [in Thai]
- Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, 80(1), 99-103.
- Tuksino, P., et al. (2016). *Measurement and assessment model of desirable characteristics according to national education standards and the development of an attitude structure measurement tool for basic education students*. National Institute of Educational Testing Service (Public Organization). [in Thai]
- UNESCO. (2015). *UNESCO competency framework* [PDF]. Retrieved from https://en.unesco.org/sites/default/files/competency_framework_e.pdf
- UNESCO. (2015). *Education 2030: Incheon declaration and framework for action for the implementation of Sustainable Development Goal 4*. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000245656>
- van der Linden, W. J., & Glas, C. A. W. (2010). *Elements of adaptive testing*. New York: Springer.
- Wainer, H., Dorans, N. J., Flaugher, R., Green, B. F., Mislevy, R. J., Steinberg, L., & Thissen, D. (2000). *Computerized adaptive testing: A primer (2nd ed.)*. New York: Lawrence Erlbaum Associates.
- Wainer, H., & Mislevy, R. J. (2000). Item response theory, item calibration, and proficiency estimation. In H. Wainer (Ed.), *Computerized adaptive testing: A primer* (2nd ed., pp. 61–100). Mahwah, NJ: Lawrence Erlbaum Associates.
- Weekley, J. A., & Ployhart, R. E. (2013). Situational judgment: Antecedents and relationships with performance. *Human Performance*, 26(3), 213-234.
- Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review*, 19(3), 188-202.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- World Economic Forum. (2020). *The future of jobs report 2020*. World Economic Forum. Retrieved from <https://www.weforum.org/reports/the-future-of-jobs-report-2020>
- World Economic Forum. (2023). *The future of jobs report 2023*. World Economic Forum. Retrieved from <https://www.weforum.org/reports/the-future-of-jobs-report-2023>
- World Economic Forum. (2025). *The future of jobs report 2025*. World Economic Forum.
- Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago, IL: MESA Press.
- Zhang, Y., & Liu, S. (2019). Planning and Organization Skills as Predictors of Academic Success. *Journal of Educational Psychology*, 111(6), 1014-1025.