

## การสร้างและการตรวจสอบคุณภาพแบบทดสอบการคิดอภิमानแบบอัตรันยพหุมิติ Creating and Quality Testing of Metacognition Multidimensional Subjective Test

พรวิมล ระวันประโคน<sup>1\*</sup> และ สมบัติ ท้ายเรือคำ<sup>2</sup>

Pornwimon Rawanprakhon<sup>1\*</sup>, and Sombat Tayraukham<sup>2</sup>

(Received: April 26, 2020 ; Revised: May 23, 2020 ; Accepted: May 29, 2020)

### บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์ 1) เพื่อสร้างแบบทดสอบการคิดอภิमानแบบอัตรันยพหุมิติ 2) เพื่อสร้างเกณฑ์การให้คะแนนแบบทดสอบการคิดอภิमानแบบอัตรันยพหุมิติ และ 3) เพื่อหาคุณภาพแบบทดสอบการคิดอภิमानแบบอัตรันยพหุมิติ ประชากรเป็นนักเรียนชั้นมัธยมศึกษาตอนปลาย จากโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาในภาคตะวันออกเฉียงเหนือ และกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาตอนปลาย จากโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาในภาคตะวันออกเฉียงเหนือ จำนวน 1,222 คน จาก 29 โรงเรียน ได้มาโดยการสุ่มอย่างง่ายโดยใช้โรงเรียนเป็นหน่วยการสุ่ม เครื่องมือที่ใช้เป็นแบบทดสอบวัดความสามารถการคิดอภิमानด้วยข้อสอบอัตรันยพหุมิติ จำนวน 9 ข้อ สถิติที่ใช้ คือ ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน การวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor analysis) การคำนวณค่าความเชื่อมั่นด้วยสูตรสัมประสิทธิ์แอลฟาและวิธีการวิเคราะห์พหุมิติ ใช้ค่าความเชื่อมั่น EAP reliability โมเดลพหุมิติ Multidimensional Random Coefficients Multinomial Logit Model (MRCMLM) ผลการวิจัยพบว่า 1) แบบทดสอบการคิดอภิमानแบบอัตรันยพหุมิติมีจำนวนทั้งสิ้น 9 ข้อ มีองค์ประกอบทั้งสิ้น 6 องค์ประกอบ 2) เกณฑ์การให้คะแนนแบบทดสอบการคิดอภิमानแบบอัตรันยพหุมิติ มีความเหมาะสม อยู่ในระดับมากที่สุด ( $\bar{x} = 2.67$ , S.D. = .02) และค่าความเชื่อมั่น (Inter-Rater Reliability: IRR) เท่ากับ .994 3) แบบทดสอบวัดความสามารถการคิดอภิमानแบบอัตรันยพหุมิติ มีคุณภาพด้าน 3.1) ความเที่ยงตรงเชิงโครงสร้าง จากการวิเคราะห์พหุมิติ พบว่าโมเดลการคิดอภิमानแบบพหุมิติมีความเหมาะสมกว่าโมเดลเอกมิติ และการวิเคราะห์องค์ประกอบเชิงยืนยันโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์มีค่าสถิติไคสแควร์ เท่ากับ 3.972 (df = 3, p = .265) GFI = .995, AGFI = .985, RMR = .0048, RMSEA = .0163 3.2) ค่าชั้นความยาก

<sup>1</sup> นิสิตปริญญาเอก สาขาวิชาวิจัยและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

<sup>2</sup> รองศาสตราจารย์ ภาควิชาวิจัยและพัฒนาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

<sup>1</sup> Doctoral degree student in Educational Research and Evaluation Program, Faculty of Education, Mahasarakham University

<sup>2</sup> Associate Professor, Department of Research and Educational Development, Faculty of Education, Mahasarakham University

\*Corresponding Author E-mail: Jiapjee@gmail.com

(Threshold) ใช้เป็นจุดตัดเกณฑ์ความสามารถของนักเรียน ซึ่งในแต่ละมิติมีชั้นความยาก 4 ชั้น จึงมีจุดตัดความสามารถ 4 จุดตัด แบ่งระดับความสามารถของนักเรียนได้เป็น 5 ระดับ 3.3) ค่าความเชื่อมั่นโดยการวิเคราะห์พหุมิติ มิติที่ 1-6 มีค่าระหว่าง .714 - .853 ค่าความเชื่อมั่นของแบบทดสอบที่วิเคราะห์ด้วยสูตรสัมประสิทธิ์แอลฟา มิติที่ 1-6 มีค่าระหว่าง .707 - .752

**คำสำคัญ** แบบทดสอบอัตนัย แบบทดสอบพหุมิติ การวิเคราะห์พหุมิติ

## Abstract

The purposes of this research were to 1) develop a metacognition multidimensional subjective test, 2) create scoring rubrics for the metacognition multidimensional subjective test, and 3) find the quality of the metacognition multidimensional subjective test. The population consisted of students in the upper secondary school level in schools under the Offices of Secondary Educational Service Area in the Northeast Region. The sample consisted of 1,222 students in the upper secondary school level in 29 schools under the Offices of Secondary Educational Service Area in the Northeast Region, obtained through simple random sampling technique, using the school as the sampling unit. The tool was a multidimensional subjective metacognitive-ability test of 9 items. The statistics used were the mean, standard deviation, confirmatory factor analysis, alpha coefficient, the multidimensional analysis employing EAP reliability, and multidimensional random coefficients multinomial logit model (MRCMLM).

The findings were as follows:

1) The multidimensional subjective metacognitive-ability test consisted of 9 questions, 6 elements.

2) The expert's opinion about the scoring rubrics of the multidimensional subjective metacognitive-ability test was at the highest level ( $\bar{x} = 2.67$ ) and the Inter-Rater Reliability (IRR) was .994.

3) The quality of the multidimensional subjective metacognitive-ability test was as follows:

3.1) Regarding the structural validity, the multidimensional model was better fitting than the unidimensional model. In the analysis using a confirmatory factor analysis, the model was fit to the empirical data. The CFA = 3.972 (df = 3, p = .265), GFI = .995, AGFI = .985, RMR = .0048, RMSEA = .0163.

3.2) Thresholds were used as cut scores. In each dimension, there were 4 difficulty levels, so there were 4 cutting points. The students' ability was divided into 5 levels. It was used to interpret and diagnose students.

3.3) The EAP reliability was from .714 to .853. The reliability (alpha coefficient) was from .707 to .752.

**Keywords:** subjective test, multidimensional test, multidimensional analysis

## บทนำ

การเรียนรู้ในศตวรรษที่ 21 ต้องก้าวข้ามสาระวิชาไปสู่การเรียนรู้ทักษะเพื่อการดำรงชีวิตที่ครูต้องลดการสอนให้น้อยลงแต่ต้องออกแบบการเรียนรู้และอำนวยความสะดวกในการเรียนรู้ให้นักเรียนเรียนรู้จากการลงมือทำและเรียนรู้จากการค้นคว้าด้วยตนเอง สังคมทุกวันนี้เป็นสังคมยุคข่าวสารไร้พรมแดน เนื่องจากอินเทอร์เน็ตทำให้โลกแคบลง ข้อมูลความรู้ ข่าวสารต่าง ๆ เปลี่ยนแปลงอยู่ตลอดเวลา บุคคลจึงต้องอาศัยทักษะการคิดวิเคราะห์และการตัดสินใจในการรับข้อมูลข่าวสาร การคิดอภิमान (metacognition) จึงมีบทบาทสำคัญในการเรียนรู้ในฐานะที่เป็นสิ่งนำบุคคลในการเรียนรู้ด้านปัญญา สามารถทำให้ผู้เรียนเรียนรู้อย่างลึกซึ้งและทำให้การเรียนรู้ดีขึ้น เนื่องจากผู้เรียนมีการจัดระบบความคิดอย่างมีประสิทธิภาพ สามารถกำกับควบคุมตนเอง รวมถึงการวางแผน การควบคุม ดำเนินงาน วางระเบียบ การทำงานและการเรียนรู้ของตนเอง ทั้งนี้การคิดอภิमानเป็นการคิดเกี่ยวกับความคิด หรือการสะท้อนการรู้คิดเป็นความสามารถของคนที่จะตระหนักถึงกระบวนการเรียนรู้ของตน รู้ว่าอะไรที่เหมาะสมกับตนเองที่สุดในการเรียนรู้ในเรื่องต่าง ๆ ตลอดจนสามารถเลือกกลวิธีในการวางแผน กำกับ ติดตาม การควบคุมกระบวนการ การประเมินและการตรวจสอบ ประเมินการเรียนรู้ของตนเองได้ ดังนั้น การคิดอภิमानจึงมีความสำคัญยิ่งในการเรียนรู้ เพราะเป็นสิ่งสำคัญในการแก้ปัญหา การตัดสินใจ การสร้างความหมายในสถานการณ์ที่เกิดขึ้น ซึ่งถ้าหากผู้เรียนที่เรียนโดยอาศัยการคิดอภิमान จะเป็นนักเรียนที่เรียนแบบมีทิศทาง เป็นผู้มีโอกาสในการตรวจดูความก้าวหน้าความสำเร็จและทิศทางของตนเองในอนาคต

การคิดอภิमानเป็นทักษะการคิดขั้นสูง (high order thinking skills) มีโครงสร้างของพฤติกรรมที่ซับซ้อนและไม่สามารถสังเกต อธิบายหรือกำหนดความหมายได้โดยตรงเพราะความคิดของบุคคลนั้นเปลี่ยนแปลงได้ตลอดเวลาและโครงสร้างของพฤติกรรมมีความสัมพันธ์กันจนไม่สามารถแยกจากกันได้ การวัดการคิดอภิमानจึงต้องอาศัยการวัดทางอ้อม โดยใช้สิ่งเร้าหรือสถานการณ์เป็นตัวกระตุ้นให้บุคคลแสดงออกถึงกระบวนการคิดของตนในลักษณะต่าง ๆ แบบทดสอบเป็นเครื่องมือหนึ่งที่สามารถไปกระตุ้นให้แสดงพฤติกรรมที่ต้องการวัดออกมาด้วยการตอบสนองต่อข้อสอบ และแบบทดสอบที่นิยมใช้วัดทักษะการคิดขั้นสูงนั้นคือแบบทดสอบอัตนัย (Subjective test) (ศิริชัย กาญจนวาสี, 2558) เนื่องจากแบบทดสอบอัตนัยมีคุณสมบัติที่ดีหลายประการ คือ สามารถวัดพฤติกรรมได้ทุกระดับ โดยเฉพาะสามารถวัดด้านการคิดวิเคราะห์สังเคราะห์และประเมินค่า และผู้ตอบได้ใช้ความรู้ความสามารถในการแสดงความคิดเห็นได้อย่างอิสระและแสดงความสามารถ

ในการใช้ภาษาได้เต็มที่ ซึ่งถ้าผู้ตอบที่ไม่มีความรู้ในเรื่องนั้นจะไม่สามารถเดาคำตอบได้ เป็นการช่วยลดความคลาดเคลื่อนในการวัดได้เป็นอย่างดี

ในการวัดคุณลักษณะแบบดั้งเดิมนั้นอยู่ภายใต้โครงสร้างแบบวัดที่เป็นเอกมิติ (unidimensionality UIRT) ซึ่งมีข้อตกลงเบื้องต้นที่สำคัญคือ ข้อคำถามแต่ละข้อมุ่งวัดเพียงคุณลักษณะแฝงเพียงมิติเดียวและแต่ละมิติต้องไม่มีความสัมพันธ์กัน (ชัยวิชิต เขียรชนะ, 2552) ดังนั้นเพื่อให้ข้อคำถามครอบคลุมคุณลักษณะของบุคคล จึงต้องสร้างข้อคำถามหลาย ๆ ข้อ ซึ่งการพยายามที่จะวัดแยกคุณลักษณะเดี่ยว ๆ และจำนวนข้อคำถามที่มากทำให้ผู้ตอบอาจเกิดความเบื่อหน่ายหรือเหนื่อยล้า ส่งผลให้เกิดความคลาดเคลื่อนในการวัด (Boyd, 1999) แต่ในความเป็นจริงแล้วธรรมชาติของคุณลักษณะทางจิตวิทยา มีโครงสร้างของพฤติกรรมที่ซับซ้อน ความเป็นพหุมิติ คือ ในการตอบคำถามหนึ่งข้อต้องอาศัยคุณลักษณะแฝงมากกว่า 1 องค์ประกอบ เพื่อที่จะตอบข้อสอบข้อนั้นได้ ซึ่งโมเดลการตอบสนองข้อสอบพหุมิติ (multidimensional item response theory : MIRT) สอดคล้องกับสภาพความเป็นจริงของการตอบข้อคำถามที่ผู้ตอบต้องใช้คุณลักษณะที่หลากหลาย เพื่อที่จะตอบข้อสอบข้อนั้นได้และยังให้ผลการสรุปอ้างอิงที่ใกล้เคียงกับคุณลักษณะภายในที่แท้จริงมากขึ้น และยังช่วยลดข้อคำถามในการวัดคุณลักษณะให้น้อยลงกว่าการทดสอบด้วยโมเดลโครงสร้างแบบเอกมิติ แต่ยังให้ผลการวัดที่มีประสิทธิภาพสูง (Frey and Seitz, 2009) ทำให้สามารถสร้างข้อคำถามได้ครอบคลุมเนื้อหาทั้งหมดได้ การตรวจข้อสอบก็ทำได้รวดเร็วขึ้น

ด้วยเหตุผลดังกล่าว เพื่อเป็นการแก้ไขข้อจำกัดของข้อสอบอัตนัยที่เป็นการวัดแบบเอกมิติ ผู้วิจัยจึงได้นำแนวคิดของโมเดลการตอบสนองข้อสอบแบบพหุมิติ (multidimensional item response model : MIRT) มาสร้างแบบทดสอบอัตนัยการคิดอภิमानแบบพหุมิติ สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย ที่สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา ในภาคตะวันออกเฉียงเหนือ

### ความมุ่งหมายของการวิจัย

1. เพื่อสร้างแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ
2. เพื่อสร้างเกณฑ์การให้คะแนนแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ
3. เพื่อหาคุณภาพแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ

### วิธีดำเนินการวิจัย

การวิจัยนี้ ผู้วิจัยดำเนินการเป็น 3 ระยะ ตามวัตถุประสงค์ของการวิจัย มีรายละเอียดดังนี้

ระยะที่ 1 การสร้างแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ

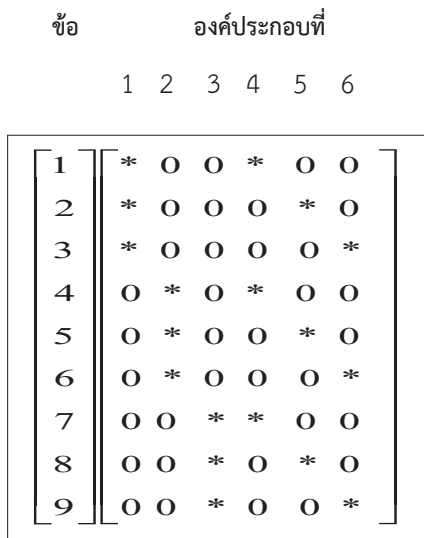
วิธีดำเนินการสร้างแบบทดสอบวัดความสามารถการคิดอภิमानแบบอัตนัยพหุมิติสำหรับนักเรียนมัธยมศึกษาตอนปลาย โดยผู้วิจัยศึกษาแนวคิดทฤษฎีที่เกี่ยวข้องกับการคิดอภิमान ทฤษฎีการตอบสนองข้อสอบแบบ พหุมิติที่ตรวจให้คะแนนมากกว่า 2 ค่า การสร้างข้อสอบแบบอัตนัยที่เป็นพหุมิติวิธีการวัดการคิดการให้คะแนนและการแปลความหมายคะแนนแล้ววิเคราะห์และสังเคราะห์กรอบแนวคิดและนิยามเชิงปฏิบัติการเกี่ยวกับการคิดอภิमान โดยศึกษาแนวคิดของ Flavell (1979), Baker and Brown (1984), Brown (1987),

Blakery and Spence (1990), Cross and Paris (1988), Pintrich (2000), Lee and Bayer (2006), Veenman (2006), McCormick (2009), Yang (2009), Negel (2014) ได้ทั้งหมด 6 องค์ประกอบย่อย จากนั้นกำหนดโมเดลการตอบสนองข้อสอบแบบพหุมิติที่มีลักษณะเป็นโมเดลการตอบสนองข้อสอบพหุมิติแบบพหุมิติภายในข้อสอบ (within – items multidimensional IRT model) และสร้างผังข้อสอบความสามารถการคิดอภิमान ดังตาราง 1 พร้อมทั้งเขียนเมตริกซ์การสร้างข้อสอบ ดังภาพ 1 และเขียนโมเดลการวัดความสามารถการคิดอภิमान ดังภาพ 2

ตาราง 1 ผังข้อสอบวัดความสามารถทางการคิดอภิमान

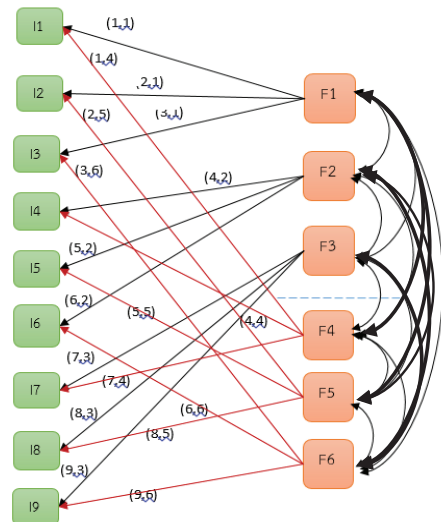
องค์ประกอบที่ต้องการวัด	ข้อ/ข้อย่อยที่																	
	1		2		3		4		5		6		7		8		9	
	1.1	1.2	2.1	2.2	3.1	3.2	4.1	4.2	5.1	5.2	6.1	6.2	7.1	7.2	8.1	8.2	9.1	9.2
1. ความรู้ในการคิดอภิमान																		
1.1 ความรู้เกี่ยวกับการให้นิยามความหมาย (K1)	✓		✓		✓													
1.2 ความรู้ที่เป็นลำดับขั้นตอนหรือวิธีการ (K2)							✓		✓		✓			✓		✓		✓
1.3 ความรู้ที่เป็นเงื่อนไข(K3)													✓		✓			✓
2. ประสบการณ์ด้านการคิดอภิमान																		
2.1 การวางแผน (K4)		✓						✓						✓				
2.2 การตรวจสอบ (K5)				✓						✓						✓		
2.3 การประเมินผล (K6)						✓					✓		✓					✓

เขียนเมตริกซ์การสร้างข้อสอบ ได้ดังนี้



ภาพ 1 เมตริกซ์การสร้างข้อสอบ

ข้อสอบ องค์ประกอบ



ภาพ 2 โมเดลการวัดความสามารถในการคิดอภิमान

โดยอธิบายได้ว่า แบบทดสอบวัดการคิดอภิमानแบบอัตนัยพหุมิติมีองค์ประกอบการคิดอภิमान 6 องค์ประกอบ ทั้งฉบับมีทั้งสิ้นจำนวน 9 ข้อ โดยข้อสอบแต่ละข้อสามารถมุ่งวัดได้มากกว่า 1 องค์ประกอบ เช่น ข้อ 1 มุ่งวัดองค์ประกอบที่ 1 และองค์ประกอบที่ 4 ข้อ 2 มุ่งวัดองค์ประกอบที่ 1 และองค์ประกอบที่ 5 ซึ่งเมื่อเขียนเป็นเมตริกซ์ข้อสอบ ใช้สัญลักษณ์ \* แทนการมีองค์ประกอบในข้อนั้น ๆ และ 0 แทน การไม่มีองค์ประกอบในข้อนั้น ๆ หลังจากนั้นเขียนนิยามความหมายขององค์ประกอบการคิดอภิमान และสร้างสถานการณ์ และข้อคำถาม โดยเขียนข้อสอบการคิดอภิमानตามนิยามเชิงปฏิบัติการ กำหนดรูปแบบข้อสอบเป็นแบบอัตนัย ข้อคำถามมีลักษณะเป็นพหุมิติ ตามแผนผังและโมเดลที่กำหนด แล้วเสนอผู้เชี่ยวชาญจำนวน 19 คน ตรวจสอบคุณภาพด้านการใช้ภาษาและความตรงเชิงโครงสร้าง โดยตรวจสอบความเที่ยงตรงเชิงโครงสร้าง จากการพิจารณาจากค่าดัชนีความสอดคล้องของผู้เชี่ยวชาญ (index of consistency : IOC) มีค่าดัชนีความสอดคล้องอยู่ระหว่าง .75 – 1.00

ระยะที่ 2 การสร้างเกณฑ์การให้คะแนนแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ

กลุ่มเป้าหมายที่ใช้ในการวิจัยครั้งนี้ แบ่งเป็น 2 กลุ่ม คือ กลุ่มเป้าหมายที่ใช้ในการสร้างเกณฑ์การให้คะแนน ได้แก่ นักเรียนระดับชั้นมัธยมศึกษาตอนปลาย จำนวน 18 คน จากโรงเรียนผดุงนารี จำนวน 6 คน โรงเรียนสาธิตมหาวิทยาลัทยมหาสารคาม จำนวน 6 คน และโรงเรียนเชียงขวัญพิทยาคม จำนวน 6 คน ได้มาโดยการเลือกแบบเจาะจง และกลุ่มเป้าหมายที่ใช้ในการตรวจสอบเกณฑ์ ได้แก่ นักเรียนระดับชั้นมัธยมศึกษาตอนปลาย โรงเรียนนรย์บุรี จำนวน 30 คน ได้มาโดยการเลือกแบบเจาะจง เครื่องมือที่ใช้ในการวิจัยในระยะนี้ ได้แก่ แบบทดสอบความสามารถการคิดอภิमानแบบอัตนัยพหุมิติที่สร้างในขั้นตอนที่ 1 และเครื่องบันทึกเสียง วิเคราะห์ข้อมูล โดยการวิเคราะห์เนื้อหา (content analysis) จากคำตอบที่ได้จากการใช้เทคนิค Think Aloud Protocol หาค่าเฉลี่ยของคะแนนความคิดเห็นของผู้เชี่ยวชาญที่มีต่อเกณฑ์การให้คะแนน และวิเคราะห์หาค่าความเชื่อมั่นของเกณฑ์การให้คะแนน (inter-rater reliability: IRR) โดยสหสัมพันธ์เพียร์สัน (Pearson product moment correlation coefficient)

ขั้นตอนการวิจัย ผู้วิจัยศึกษาการสร้างเกณฑ์การให้คะแนนตามแนวทางของการสร้างรูบริกส์เกณฑ์การให้คะแนนคำตอบของนักเรียนตามมาตรวัดเมตาคognition แบบรายบุคคล (IAMS) ของ Paris and Jacob (1987) รวมทั้งเทคนิคการคิดออกเสียง (think aloud protocol) ซึ่งเป็นวิธีการให้นักเรียนรายงานความคิดขณะที่กำลังทำอยู่ในกระบวนการคิด หลังจากนั้นนำแบบทดสอบไปใช้กับกลุ่มเป้าหมายในการสร้างเกณฑ์การให้คะแนน (rubrics score) โดยให้นักเรียนเขียนตอบลงในกระดาษและสัมภาษณ์เพิ่มเติม แล้วสังเคราะห์คำตอบของนักเรียนเพื่อหาแนวคำตอบในแต่ละระดับเสนอเกณฑ์การให้คะแนนการวัดความสามารถการคิดอภิमान และแนวคำตอบให้ผู้เชี่ยวชาญตรวจสอบคุณภาพด้านการใช้ภาษา และความเหมาะสมของเกณฑ์การให้คะแนน โดยแบบประเมินสำหรับผู้เชี่ยวชาญ มีลักษณะเป็นมาตราประมาณค่า 3 ระดับ ซึ่ง คะแนน 3 หมายถึง ดีมาก คะแนน 2 หมายถึง พอใช้ คะแนน 1 หมายถึง ควรปรับปรุง จากนั้นนำแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ ไปทดสอบกับนักเรียนชั้นมัธยมศึกษาตอนปลายจากโรงเรียนนรย์บุรีพิทยาคม รัชชังคลาพิเชก จ.บุรีรัมย์ จำนวน 30 คน แล้วนำคำตอบชุดเดียวกันไปให้ผู้เชี่ยวชาญ 2 คนตรวจสอบเพื่อหาค่าความเชื่อมั่นระหว่างผู้ประเมิน (inter-rater reliability: IRR)

### ระยะที่ 3 การหาคุณภาพแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ

ผู้วิจัยเก็บข้อมูลกับนักเรียนระดับชั้นมัธยมศึกษาตอนปลายในภาคตะวันออกเฉียงเหนือ จำนวน 1,222 คน ซึ่งได้มาโดยการสุ่มอย่างง่ายโดยใช้โรงเรียนเป็นหน่วยการสุ่มจากโรงเรียนทั้งหมด 933 โรงเรียน กำหนดขนาดกลุ่มตัวอย่างจากการคำนวณ 20 เท่าของพารามิเตอร์ (Hair et al., 2010) นำแบบทดสอบวัดความสามารถการคิดอภิमानแบบอัตนัยพหุมิติ ที่สร้างในระยะที่ 1 ส่งไปยังโรงเรียนต่าง ๆ เพื่อให้ครูผู้สอนช่วยดำเนินการทดสอบ หลังจากนั้นรวบรวมแบบทดสอบจากโรงเรียนมาตรวจให้คะแนน โดยใช้เกณฑ์การให้คะแนนที่สร้างขึ้นในระยะที่ 2 จากนั้นวิเคราะห์หาคุณภาพความเที่ยงตรงเชิงโครงสร้าง (construct validity) ด้วยวิธีการวิเคราะห์พหุมิติ และการวิเคราะห์องค์ประกอบเชิงยืนยัน (confirm factor analysis) หาค่าความยาก (threshold) ของชั้นการตอบและหาค่าความเชื่อมั่นของแบบทดสอบความสามารถการคิดอภิमानแบบอัตนัยพหุมิติ ด้วยวิธีการคำนวณด้วยสูตรสัมประสิทธิ์แอลฟาและวิธีการวิเคราะห์พหุมิติ โดยใช้ค่าความเชื่อมั่น EAP reliability ด้วยโปรแกรมสำเร็จรูป SPSS for window statistical package for social science), โปรแกรม ConQuest และโปรแกรม LISREL (linear structure relationship)

### ผลการวิจัย

1. กระบวนการสร้างแบบทดสอบวัดความสามารถการคิดอภิमानแบบอัตนัยพหุมิติ สำหรับนักเรียนมัธยมศึกษาตอนปลาย มีดังนี้

1.1 วางแผนการสร้างข้อสอบ โดยกำหนดจุดมุ่งหมายการสร้างข้อสอบแบบอัตนัยพหุมิติ ศึกษาแนวคิดทฤษฎีที่เกี่ยวข้องกับการคิดอภิमान ทฤษฎีการตอบสนองข้อสอบพหุมิติที่ตรวจให้คะแนนมากกว่า 2 ค่า วิธีการวัดการคิด การให้คะแนนและการแปลความหมายคะแนน แล้วสังเคราะห์กรอบแนวคิดของนักการศึกษาเกี่ยวกับการคิดอภิमान ได้องค์ประกอบใหญ่ 2 องค์ประกอบ ได้แก่ องค์ประกอบด้านความรู้ในการคิดอภิमान และองค์ประกอบด้านประสบการณ์ด้านการคิดอภิमान และองค์ประกอบย่อย 6 องค์ประกอบ ได้แก่ องค์ประกอบด้านความรู้เกี่ยวกับการให้นิยามความหมาย (Declarative knowledge: K1) ความรู้ที่เป็นลำดับขั้นตอนหรือวิธีการ (Procedural knowledge: K2) ความรู้ที่เป็นเงื่อนไข (Conditional knowledge: K3) ประสบการณ์การวางแผน (Planning: K4) ประสบการณ์การตรวจสอบ (Monitoring: K5) และประสบการณ์การประเมินผล (Evaluating: K6) จากนั้นกำหนดนิยามความหมาย และสร้างผังข้อสอบวัดความสามารถทางการคิดอภิमान แล้วนำมาเขียนเมตริกซ์การสร้างข้อสอบ และเขียนโมเดลการวัดการคิดอภิमान

1.2 สร้างและพัฒนาแบบทดสอบ ตามองค์ประกอบและนิยามของข้อสอบ ได้ข้อสอบทั้งฉบับจำนวน 9 ข้อ ประกอบด้วย 3 สถานการณ์ แต่ละสถานการณ์ถามคำถาม 3 ข้อ แต่ละข้อมุ่งวัด 2 องค์ประกอบย่อยข้อสอบทั้ง 9 ข้อ ครอบคลุม 6 องค์ประกอบย่อย ตรวจสอบความถูกต้องเหมาะสมของภาษาจากผู้เชี่ยวชาญจำนวน 19 คน

1.3 ตรวจสอบความเที่ยงตรงเชิงโครงสร้าง โดยหาค่าดัชนีความสอดคล้องของผู้เชี่ยวชาญ (Index of Consistency: IOC) จากผู้เชี่ยวชาญชุดเดิม พบว่า ข้อคำถามมีค่า IOC ตั้งแต่ .75 - 1.00 ซึ่งสูงกว่าเกณฑ์การคัดเลือก .50 ขึ้นไป



2. ผลการสร้างเกณฑ์การให้คะแนนแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ มีรายละเอียดดังนี้

เกณฑ์การให้คะแนนแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ สามารถสรุปแนวทางการตอบได้ 3 ระดับคะแนน คือ 0, 1 และ 2 ซึ่ง 0 คะแนน คือ นักเรียนไม่ตอบหรือตอบไม่ตรงประเด็น 1 คะแนน คือ นักเรียนตอบได้แต่ไม่ตรงประเด็นหรือให้เหตุผลไม่สมเหตุสมผล และ 2 คะแนน คือ นักเรียนตอบได้ตรงประเด็น พร้อมให้เหตุผลที่สมเหตุสมผล โดยมีความเหมาะสมของเกณฑ์การให้คะแนนจากความเห็นของผู้เชี่ยวชาญ พบว่า มีความเหมาะสมอยู่ในระดับมากที่สุด ( $\bar{x} = 2.67$ , S.D. = 0.02) (วิเชียร เกตุสิงห์, 2538) และค่าความเชื่อมั่นของเกณฑ์การให้คะแนนจากการให้คะแนนความสอดคล้องของผู้ประเมิน (Inter-Rater Reliability: IRR) จำนวน 2 คน พบว่า มีความเชื่อมั่นเท่ากับ .994

3. ผลการหาคุณภาพแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ

3.1 ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบวัดความสามารถการคิดอภิमानแบบอัตนัยพหุมิติ จำแนกเป็นความเที่ยงตรงที่วิเคราะห์ด้วยวิธีพหุมิติ และวิธีวิเคราะห์องค์ประกอบเชิงยืนยัน (CFA) ดังนี้

1) ผลการวิเคราะห์ความเที่ยงตรงเชิงโครงสร้างด้วยวิธีพหุมิติ พบว่า โมเดลการคิดอภิमानแบบพหุมิติ มีค่าสถิติดีเวียนซ์ (Deviance statistic:  $G^2$ ) เท่ากับ 299870.16805 (จำนวนพารามิเตอร์เท่ากับ 57) ซึ่งน้อยกว่า โมเดลการคิดอภิमानแบบเอกมิติ ที่มีค่าสถิติดีเวียนซ์ (Deviance statistic) เท่ากับ 30179.99250 (จำนวนพารามิเตอร์เท่ากับ 37) และผลการเปรียบเทียบตามค่า เกณฑ์สารสนเทศไคคิ (Akaike information criterion; AIC) ระหว่างโมเดลการคิดอภิमानแบบพหุมิติกับโมเดลการคิดอภิमानแบบเอกมิติ พบว่า โมเดลการคิดอภิमानแบบพหุมิติ มีค่าเกณฑ์สารสนเทศไคคิ (AIC) เท่ากับ 29984.16805 โมเดลการคิดอภิमानแบบเอกมิติ มีค่าเกณฑ์สารสนเทศไคคิ (AIC) เท่ากับ 30253.99250 ซึ่งโมเดลการคิดอภิमानแบบพหุมิติมีค่าเกณฑ์สารสนเทศไคคิ (AIC) น้อยกว่าโมเดลการคิดอภิमानแบบเอกมิติแสดงว่า โมเดลความสามารถการคิดอภิमानแบบพหุมิติมีความเหมาะสมของโมเดลที่สอดคล้องกับข้อมูลมากกว่าโมเดลความสามารถการคิดอภิमानแบบเอกมิติ ซึ่งเป็นตัวบ่งชี้ถึงโครงสร้างของความสามารถการคิดอภิमानมีความเหมาะสมกับการวัดพหุมิติ อันเป็นหลักฐานแสดงถึงความตรงตามโครงสร้างทฤษฎีของแบบทดสอบวัดความสามารถการคิดอภิमानแบบอัตนัยพหุมิติ

2) ผลการตรวจสอบความตรงเชิงโครงสร้างด้วยวิธีการวิเคราะห์องค์ประกอบเชิงยืนยัน พบว่า เมื่อทำการตรวจสอบความสอดคล้องของโครงสร้างทฤษฎีการคิดอภิमानแบบพหุมิติกับข้อมูลเชิงประจักษ์ โดยการทดสอบสถิติไคสแควร์ ( $\chi^2$ ) มีค่าเท่ากับ 3.972 (df = 3, p = .265) ไม่มีนัยสำคัญทางสถิติ ดังนั้นวัดระดับความกลมกลืน (GFI) เท่ากับ .995 ดังนั้นวัดระดับความกลมกลืนที่ปรับค่าแล้ว (AGFI) เท่ากับ .985 ค่ารากของค่าเฉลี่ยกำลังสองของส่วนเหลือ (RMR) เท่ากับ .0048 และค่ารากกำลังสองเฉลี่ยของค่าความแตกต่างโดยประมาณ (RMSEA) เท่ากับ .0163 จะเห็นได้ว่าค่าดัชนีต่าง ๆ อยู่ในเกณฑ์ที่ยอมรับได้ โดยค่าไคสแควร์มีค่าเข้าใกล้ศูนย์ และมีค่าใกล้เคียงกับค่าองศาอิสระ (df) แสดงว่าแบบจำลองมีความกลมกลืนกับข้อมูลที่นำมาวิเคราะห์ เมื่อพิจารณาค่า GFI และ AGFI ก็พบว่า มีค่าเข้าใกล้ 1.00 ซึ่งแสดงว่าแบบจำลองมีความกลมกลืนกับข้อมูลที่ใช้ในการวิเคราะห์ ส่วนค่า RMR เป็นค่าที่บอกขนาดของส่วนที่เหลือโดยเฉลี่ยจากการเปรียบเทียบระดับความกลมกลืนของแบบจำลองกับข้อมูลที่ใช้ในการวิเคราะห์ ซึ่งค่าดัชนี RMR ยิ่งเข้าใกล้ศูนย์



แสดงว่าแบบจำลองมีความกลมกลืนกับข้อมูลที่ใช้ในการวิเคราะห์ โดยทั่วไปดัชนี RMR ที่น้อยกว่า .10 ถือว่าอยู่ในเกณฑ์ที่ยอมรับได้ (Kline, 1998) และค่า RMSEA มีค่าน้อยกว่า 0.07 ซึ่งถือว่าอยู่ในเกณฑ์ที่ยอมรับได้ (Hu and Bentler, 1999) จึงสรุปได้ว่า จากค่าดัชนีความเหมาะสมสอดคล้องของโมเดลกับข้อมูลอยู่ในเกณฑ์ที่ยอมรับทุกค่า แสดงให้เห็นว่าโครงสร้างทฤษฎีความสามารถการคิดอภิमानมีความสอดคล้องกับข้อมูลเชิงประจักษ์อันเป็นหลักฐานแสดงถึงความตรงตามโครงสร้างทฤษฎีของความสามารถการคิดอภิमानแบบพหุมิติ

3.2 ค่าขั้นความยาก (threshold) ของแต่ละมิติ (ตาราง 2) พบว่า แต่ละมิติมีค่าขั้นความยาก (threshold) สูงขึ้นเรื่อย ๆ ตั้งแต่ขั้นที่ 1 ถึงขั้นที่ 4 โดยแต่ละขั้นเป็นจุดตัดความสามารถของนักเรียน ในขั้นนี้มี 4 จุดตัดทำให้แบ่งความสามารถของนักเรียนออกเป็น 5 ระดับ ซึ่งถ้านักเรียนจะตอบให้ได้คะแนนมากขึ้น นักเรียนจะต้องใช้ความสามารถในการตอบข้อสอบสูงขึ้น และเมื่อนักเรียนมีความสามารถตกอยู่ที่ขั้นใด จะแปลความหมายความสามารถของนักเรียนได้ว่าอยู่ระดับใด

ตาราง 2 ค่าขั้นความยาก (threshold) ของแต่ละมิติ

Threshold	มิติความสามารถการคิดอภิमान					
	K1	K2	K3	K4	K5	K6
1	-1.313	-1.068	-1.111	-1.165	-0.555	-0.914
2	-1.022	-1.076	-1.129	-0.822	-1.201	-0.941
3	0.975	0.707	0.811	0.712	0.831	0.685
4	1.359	1.437	1.429	1.275	0.925	1.170

3.3 ค่าความเชื่อมั่นของแบบทดสอบ ทั้งการวิเคราะห์ด้วยสัมประสิทธิ์แอลฟาซึ่งเป็นการพิจารณาความเชื่อมั่นตามทฤษฎีการทดสอบแบบดั้งเดิม และการวิเคราะห์พหุมิติซึ่งเป็นการพิจารณาความเชื่อมั่นตามทฤษฎีการตอบสนองข้อสอบที่ใช้โมเดลมาจิกแนล ประมาณค่าแบบมาร์จิ้นัลแม็กซ์ลิคฮูด (marginal maximum-likelihood: MML) พบว่า ทั้ง 6 องค์กรประกอบ มีค่าสัมประสิทธิ์แอลฟา อยู่ระหว่าง .707 - .752 และค่าความเชื่อมั่นแบบ EAP reliability มีค่าอยู่ระหว่าง .714 - .853 โดยการพิจารณาความเชื่อมั่นทั้งสองทฤษฎีนั้นใช้เกณฑ์เดียวกันในการพิจารณา ซึ่งยอมรับค่า .700 ขึ้นไป (Nunnally and Bernstein, 1994: Hair et al, 2010) จึงสรุปได้ว่า แบบทดสอบการคิดอภิमानแบบอัตรันัยพหุมิติเป็นแบบทดสอบที่มีความเชื่อมั่นอยู่ในระดับสูง รายละเอียดแสดงดังตาราง 3

**ตาราง 3** ค่าความเชื่อมั่นของแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ

องค์ประกอบ/มิติ	ค่าความเชื่อมั่น (reliability)	
	สัมประสิทธิ์แอลฟา	การวิเคราะห์พหุมิติ (EAP reliability)
ความรู้เกี่ยวกับการให้นิยามความหมาย (K1)	.707 (SEM = 1.496)	.846
ความรู้ที่เป็นลำดับขั้นตอนหรือวิธีการ (K2)	.746 (SEM = 1.494)	.853
ความรู้ที่เป็นเงื่อนไข (K3)	.711 (SEM = 1.519)	.845
การวางแผน (K4)	.714 (SEM = 1.365)	.717
การตรวจสอบ (K5)	.740 (SEM = 1.427)	.787
การประเมินผล (K6)	.752 (SEM = 1.375)	.714

### อภิปรายผล

1. ผลการสร้างแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ พบว่า แบบทดสอบมีองค์ประกอบ 6 องค์ประกอบ มีความเที่ยงตรงเชิงโครงสร้าง โดยมีค่า IOC อยู่ระหว่าง 0.75 - 1.00 ซึ่งสูงกว่าเกณฑ์การคัดเลือก 0.50 ขึ้นไป ทั้งนี้เนื่องจาก ผู้วิจัยได้สังเคราะห์การคิดอภิमान จากแนวคิดของนักการศึกษาหลายท่าน ที่มีแนวคิดที่คล้ายคลึงกัน (Cross and Paris, 1988; Flavell, 1979; Flavell, Miller and Miller, 1985; Lee and Baylor, 2006; Pintrich, 2000; Schraw and Dennison, 1994) จึงทำให้สามารถจัดหมวดหมู่ได้อย่างชัดเจน และมีองค์ประกอบที่เหมาะสมที่สุด นอกจากนี้ ผู้วิจัยได้กำหนด พฤติกรรมบ่งชี้ของแต่ละองค์ประกอบ ได้ชัดเจน และเขียนนิยามได้ตรงตามพฤติกรรมที่ต้องการวัด ทำให้สร้างข้อสอบได้สอดคล้องกับนิยาม เป็นผลให้ข้อสอบมีความเที่ยงตรงเชิงโครงสร้างสูง

2. ผลการสร้างเกณฑ์การให้คะแนนแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ มีรายละเอียดดังนี้

2.1 เกณฑ์การให้คะแนนแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ สามารถสรุปแนวทางการตอบได้ 3 ระดับคะแนน คือ 0, 1 และ 2 ซึ่ง 0 คะแนน คือ นักเรียนไม่ตอบหรือตอบไม่ตรงประเด็น 1 คะแนน คือ นักเรียนตอบได้แต่ไม่ตรงประเด็นหรือให้เหตุผลไม่สมเหตุสมผล และ 2 คะแนน คือ นักเรียนตอบได้ตรงประเด็น พร้อมให้เหตุผลที่สมเหตุสมผล ทั้งนี้เนื่องจาก ในการเก็บข้อมูลผู้วิจัยได้แบ่งกลุ่มตัวอย่างออกเป็นกลุ่มเก่ง กลาง อ่อน เพื่อตรวจสอบว่าเด็กกลุ่มเก่ง ปานกลาง อ่อน ให้คำตอบแตกต่างกันหรือไม่อย่างไร และพบว่า มีความแตกต่างกันอย่างเห็นได้ชัด โดยที่กลุ่มเก่ง สามารถให้คำตอบได้ตรงประเด็นครบถ้วน ให้เหตุผลที่สมเหตุสมผล ส่วนกลุ่มปานกลาง ให้คำตอบได้ตรงประเด็นแต่ไม่ครบถ้วนหรือให้เหตุผลได้ในบางประเด็นเท่านั้น และกลุ่มอ่อนตอบคำถามได้แต่ไม่ตรงประเด็น ดังนั้นผู้วิจัยจึงได้กำหนดระดับคะแนนเป็น 3 ระดับ คือ 0, 1 และ 2 ตามความสามารถของนักเรียน สอดคล้องกับ Moskal (2000) ที่อธิบายไว้ว่า หลังจากกำหนดเกณฑ์สำหรับการปฏิบัติที่ประสิทธิภาพระดับสูงสุดแล้ว ผู้ประเมินจะหันมากำหนดเกณฑ์สำหรับ ประสิทธิภาพต่ำสุด ความแตกต่างระหว่างเกณฑ์สำหรับประสิทธิภาพระดับบนสุดและประสิทธิภาพในระดับล่างมีแนวโน้มที่จะแนะนำเกณฑ์ที่เหมาะสมสำหรับระดับกลางของประสิทธิภาพ วิธีนี้จะส่งผลให้ได้คะแนนสามระดับ

2.2 ความเหมาะสมของเกณฑ์การให้คะแนน จากการตรวจสอบคุณภาพด้านการใช้ภาษา และความเหมาะสมของเกณฑ์การให้คะแนน พบว่า โดยภาพรวมของผู้เชี่ยวชาญ มีความคิดเห็นว่าเหมาะสมอยู่ในระดับมากที่สุด ( $\bar{x} = 2.67$ ) ทั้งนี้แบบประเมินสำหรับผู้เชี่ยวชาญ มีลักษณะเป็นมาตราประมาณค่า 3 ระดับ ซึ่งกำหนดคะแนนไว้ว่า คะแนน 3 หมายถึง ดีมาก คะแนน 2 หมายถึง พอใช้ คะแนน 1 หมายถึง ควรปรับปรุง และมีเกณฑ์การแปลความหมาย ดังนี้ คะแนนเฉลี่ยระหว่าง 2.34-3.00 หมายถึง มีความเหมาะสมมากที่สุด คะแนนเฉลี่ยระหว่าง 1.67-2.33 หมายถึง มีความเหมาะสมปานกลาง และคะแนนเฉลี่ยระหว่าง 1.00-1.66 หมายถึง มีความเหมาะสมน้อยที่สุด (วิเชียร เกตุสิงห์, 2538) ดังนั้น จากค่าเฉลี่ย 2.67 ซึ่งแปลความหมายได้ว่า มีความเหมาะสมมากที่สุด จึงอธิบายได้ว่า เกณฑ์การให้คะแนนสำหรับแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ ในแต่ละข้อระบุข้อความที่สามารถวัดพฤติกรรมได้ตรงตามตัวชี้วัดอยู่ในระดับมากที่สุด นั้นหมายความว่าเกณฑ์การให้คะแนนมีความเที่ยงตรงสูง ซึ่งความเที่ยงตรงเป็นคุณสมบัติที่สำคัญของเครื่องมือวัดทุกประเภท (ศิริชัย กาญจนวาสี, 2556)

2.3 ความเชื่อมั่นของเกณฑ์การให้คะแนน มีค่าเท่ากับ .994 แสดงให้เห็นว่าค่าความเชื่อมั่น และความสอดคล้องของผู้ตรวจให้คะแนน (Rater Agreement Index: RAI) อยู่ในระดับสูง ทั้งนี้เนื่องจาก เกณฑ์การให้คะแนนที่ผ่านการตรวจสอบโดยผู้เชี่ยวชาญแล้วมีความเที่ยงตรงอยู่ในระดับสูง จะทำให้มีความเชื่อมั่นของผู้ตรวจสูงเช่นกัน สอดคล้องกับ บุซวรรัช สแนปลีม (2556) ผลจากการสร้างและพัฒนาเกณฑ์การตรวจให้คะแนนเรียงความ พบว่า เกณฑ์การให้คะแนนแบบรวมองค์ประกอบ มีดัชนีความสอดคล้องของผู้ตรวจให้คะแนน เท่ากับ .833 และค่าความเชื่อมั่นเท่ากับ .921

### 3. ผลการหาคุณภาพแบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ

3.1 ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบวัดความสามารถการคิดอภิमानแบบอัตนัยพหุมิติ จำแนกเป็นความเที่ยงตรงที่วิเคราะห์ด้วยวิธีพหุมิติ และวิธีวิเคราะห์องค์ประกอบเชิงยืนยัน ดังนี้

1) ผลการวิเคราะห์ความเที่ยงตรงเชิงโครงสร้างด้วยวิธีพหุมิติ พบว่า โมเดลการคิดอภิमानแบบพหุมิติ มีค่าสถิติดีเวียนซ์ น้อยกว่า โมเดลการคิดอภิमानแบบเอกมิติ อย่างมีนัยสำคัญที่ระดับ .01 และผลการเปรียบเทียบตามค่าเกณฑ์สารสนเทศเอไคคี (Akaike information criterion; AIC) ระหว่างโมเดลการคิดอภิमानแบบพหุมิตินับกับโมเดลการคิดอภิमानแบบเอกมิติ พบว่าโมเดลการคิดอภิमानแบบพหุมิติ มีค่าเกณฑ์สารสนเทศเอไคคี (AIC) น้อยกว่าโมเดลการคิดอภิमानแบบเอกมิติ กล่าวได้ว่า โมเดลการคิดอภิमानแบบพหุมิติ มีความเหมาะสมกับข้อมูลมากกว่าโมเดลการคิดอภิमानแบบเอกมิติ โดยพิจารณาตามเกณฑ์ความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ ดังที่ Hoskens and Boeck (2001) กล่าวว่า สถิติดีเวียนซ์และเกณฑ์สารสนเทศเอไคคี เป็นสถิติที่ใช้ในการประเมินบ่งชี้ถึงความเหมาะสมของโมเดล โมเดลใดที่มีค่าสถิติดีเวียนซ์และเกณฑ์สารสนเทศเอไคคี แสดงว่าโมเดลนั้นมีความสอดคล้องกับข้อมูลเชิงประจักษ์มากกว่า สอดคล้องกับงานวิจัยของชัยวิชิต เขียรชนะ (2552) ที่พบว่า โมเดลโครงสร้างกลยุทธ์การเรียนรู้แบบพหุมิติมีความเหมาะสมมากกว่าแบบเอกมิติรวม และเหมาะสมมากกว่าแบบเอกมิติแยกตามมิติคล้ายคลึงกับ สุภัญญา ทองนาค (2555) ที่พบว่า แบบทดสอบมีความตรงเชิงโครงสร้างโดยโมเดลการวัดสมรรถนะนักศึกษาตามมาตรฐานวิชาชีพครูแบบพหุมิติภายในข้อมีความเหมาะสมมากกว่าแบบเอกมิติรวม และเหมาะสมมากกว่าแบบเอกมิติแยกตามมิติ

2) ผลการตรวจสอบความตรงเชิงโครงสร้างด้วยวิธีการวิเคราะห์องค์ประกอบเชิงยืนยันพบว่า ผลการตรวจสอบความสอดคล้องของโครงสร้างทฤษฎีการคิดอภิमानแบบพหุมิติมีความสอดคล้องกับข้อมูลเชิงประจักษ์ เมื่อพิจารณาตามเกณฑ์ของ Kelloway (1998) โดยค่าสถิติไคสแควร์ ( $\chi^2$ ) มีค่าเท่ากับ 3.972 (df = 3, p = 0.265) ไม่มีนัยสำคัญทางสถิติ แต่เนื่องจากค่าไคสแควร์มีความไวต่อขนาดกลุ่มตัวอย่าง จึงพิจารณาค่าสถิติตัวอื่น ๆ ร่วมด้วย คือดัชนีวัดระดับความกลมกลืน (GFI = .995) ซึ่งมีค่าเกิน .90 เป็นตัวแสดงประสิทธิภาพของโมเดลในภาพรวมทั้งหมด ดัชนีวัดระดับความกลมกลืนที่ปรับค่าแล้ว (AGFI = .985) มีค่าเกิน 0.90 มีคุณสมบัติเช่นเดียวกับค่าดัชนีวัดระดับความกลมกลืน (GFI) โดยนำค่าดัชนีวัดระดับความกลมกลืน (GFI) มาปรับแก้ ซึ่งคำนึงถึงขนาดขององศาอิสระ รวมถึงจำนวนตัวแปรและกลุ่มตัวอย่าง ค่ารากของค่าเฉลี่ยกำลังสองของส่วนเหลือ (RMR = .0048) มีค่าต่ำกว่า 0.05 เป็นค่าที่บ่งบอกขนาดของความคลาดเคลื่อนในการวัด ซึ่งเป็นความคลาดเคลื่อนที่แสดงความแตกต่างระหว่างความแปรปรวนที่ได้รับการพยากรณ์และความแปรปรวนที่แท้จริง และค่ารากกำลังสองเฉลี่ยของค่าความแตกต่างโดย ประมาณ (RMSEA = 0.0163) มีค่าต่ำกว่า 0.05 เป็นค่าที่แสดงถึงขนาดของความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์จากเกณฑ์การพิจารณา ค่าสถิติอยู่ในเกณฑ์การยอมรับทุกค่าจึงเป็นสิ่งบ่งชี้ว่าโมเดลความสามารถการคิดอภิमानมีความสอดคล้องกับข้อมูลเชิงประจักษ์อันเป็นหลักฐานความตรงเชิงโครงสร้างของแบบทดสอบ ซึ่งมีความสอดคล้องกับแนวคิดของ ศิริชัย กาญจนวาสิ (2558) และ McIntire and Miller (2007) ที่กล่าวไว้ว่าการวิเคราะห์องค์ประกอบเชิงยืนยันเป็นการแสดงหลักฐานความตรงเชิงโครงสร้างโดยหลักฐานที่แสดงนั้นเป็นการตรวจสอบความสอดคล้องของโมเดลกับข้อมูล หากโมเดลมีความสอดคล้องกับข้อมูลจะบ่งชี้ถึงโมเดลองค์ประกอบที่ศึกษาเป็นหลักฐานสำหรับยืนยันองค์ประกอบคุณลักษณะที่วัดจากหลักฐานดังกล่าวจึงแสดงถึงความเที่ยงตรงเชิงโครงสร้างทฤษฎีของแบบทดสอบวัดความสามารถการคิดอภิमानแบบพหุมิติด้วยวิธีการวิเคราะห์องค์ประกอบเชิงยืนยัน คล้ายคลึงกับงานวิจัยของ อัญชลี ศรีกลชาญ (2547) ได้ศึกษาเรื่องการพัฒนาแบบวัดพหุปัญญาสำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น พบว่า แบบวัดพหุปัญญามีความตรงเชิงโครงสร้างอยู่ในเกณฑ์ดี มีค่าไคสแควร์ = 120427 (p = .100), df = 1461, GFI = .95, AGFI = .94, CFI = 1.00, Standardized RMR = .03, RMSEA = .00 และองค์ประกอบทั้ง 8 องค์ประกอบ กับข้อสอบทั้ง 60 ข้อ มีนัยสำคัญทางสถิติที่ระดับ .05 และวนิดา ดีแป้น (2556) ได้ศึกษาการสร้างโมเดลการตอบสนองรายข้อแบบพหุมิติ พหุระดับของพฤติกรรมกรเป็นสมาชิก องค์กรของครู พบว่า ผลการวิเคราะห์องค์ประกอบเชิงยืนยัน พหุระดับบ่งชี้ความตรงเชิงโครงสร้างของแบบวัดพฤติกรรมกรเป็นสมาชิกองค์กรของครู โดยมีค่า ไคสแควร์ เท่ากับ 171.701 (df 161, p = 0.2178), CFI = 0.999, TLI = 0.998, RMSEA = 0.010, SRMRW = 0.012 และ SRMRb = 0.048

3.2 ค่าชั้นความยาก (threshold) ของแต่ละมิติ มีชั้นความยาก 4 ชั้น โดยแต่ละชั้นเป็นจุดตัดความสามารถของนักเรียน ในที่นี้มี 4 จุดตัดทำให้แบ่งความสามารถของนักเรียนออกเป็น 5 ระดับ ทำให้การแปลความหมายมีความเที่ยงตรง เนื่องจากเกิดจากการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติที่มีการวิเคราะห์ตามโมเดลการให้คะแนนแบบพหุภาคชนิดโมเดลการให้คะแนนความรู้บางส่วน (partial credit

model: PCM) ซึ่งเป็นการวิเคราะห์ที่สอดคล้องกับลักษณะหรือธรรมชาติของตัวแปรที่มีความเป็นพหุมิติ ทำให้การแปลความหมายตรงกับความเป็นจริงมากขึ้น

3.3 ค่าความเชื่อมั่นของแบบทดสอบความสามารถการคิดอภิमानแบบอัตนัยพหุมิติ วิเคราะห์ด้วยวิธีการวิเคราะห์พหุมิติ โดยหาค่าความเชื่อมั่นแบบ EAP reliability และวิธีการคำนวณด้วยสูตรสัมประสิทธิ์แอลฟา

ผลการวิเคราะห์ความเชื่อมั่นของแบบทดสอบความสามารถการคิดอภิमानแบบอัตนัยพหุมิติ จำนวน 9 ข้อ โดยวิเคราะห์ค่าความเชื่อมั่นตามทฤษฎีการทดสอบแบบดั้งเดิมด้วยการหาสัมประสิทธิ์แอลฟาของครอนบาค และการวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ (item response theory: IRT) ด้วยการหาค่าสัมประสิทธิ์ความเชื่อมั่นแบบ EAP reliability พบว่า การประมาณค่าความเชื่อมั่นด้วยสูตรสัมประสิทธิ์แอลฟามีค่าความเชื่อมั่นทั้ง 6 องค์ประกอบ เท่ากับ .707, .746, .711, .714, .740 และ .752 ตามลำดับ ค่าความคลาดเคลื่อนมาตรฐานของการวัด (SEM) เท่ากับ 1.496, 1.494, 1.519, 1.365, 1.427 และ 1.375 ตามลำดับ โดยการประมาณค่าแบบมาร์จิ้นัลแมกซ์ลิคูด (marginal maximum-likelihood: MML) มีค่าความเชื่อมั่นแบบ EAP reliability ทั้ง 6 องค์ประกอบ เท่ากับ 0.846, 0.853, 0.845, 0.717, 0.787 และ 0.714 ตามลำดับ จากผลการวิเคราะห์ค่าความเชื่อมั่นทั้งสองวิธีนั้นถือว่ามีความเหมาะสมสำหรับเป็นเครื่องมือในการวิจัย สอดคล้องกับ Nunnally and Bernstein (1994); Hair et al. (2010) ที่ได้เสนอเกณฑ์การพิจารณาความเชื่อมั่นที่ยอมรับได้ คือ ต้องมีค่าความเชื่อมั่นตั้งแต่ .700 ขึ้นไป ซึ่งเกณฑ์นี้สามารถใช้พิจารณาได้ทั้งการประมาณค่าความเชื่อมั่นโดยวิธีสัมประสิทธิ์แอลฟาและการประมาณค่าความเชื่อมั่นด้วยการวิเคราะห์พหุมิติ สอดคล้องกับ Adams (2005) ที่สรุปไว้ว่า ค่าความเชื่อมั่นทั้งตามทฤษฎีการทดสอบแบบดั้งเดิมเป็นค่าที่เหมือนกับทฤษฎีการตอบสนองข้อสอบที่ใช้โมเดลมาร์จิ้นัล (marginal model) โดยสามารถใช้ค่าความเชื่อมั่นในการแสดงถึงคุณภาพของการออกแบบการวัด (measurement design) จึงกล่าวได้ว่า แบบทดสอบวัดความสามารถการคิดอภิमानที่ผู้วิจัยสร้างขึ้นมีค่าความเชื่อมั่นเกินเกณฑ์พิจารณาทุกด้าน มีความคงเส้นคงวาของการวัด สอดคล้องกับงานวิจัยของ สุภัญญา ทองนาค (2555) ได้ศึกษาเรื่อง การพัฒนาแบบทดสอบสมรรถนะนักศึกษาตามมาตรฐานวิชาชีพครูแบบพหุมิติที่มีการตรวจให้คะแนนแบบพหุวิภาค แบบทดสอบสมรรถนะนักศึกษาตามมาตรฐานวิชาชีพครูแบบพหุมิติที่พัฒนาขึ้น มีจำนวน 125 ข้อ วัดสมรรถนะ 3 ด้าน คือ ด้านความรู้คิด ด้านทักษะและความสามารถ และด้านคุณลักษณะ พบว่า มีความเชื่อมั่นโดยการวิเคราะห์พหุมิติ (EAP reliability) เท่ากับ 0.8381 0.8303 และ 0.7875 ตามลำดับ นอกจากนี้ยังสอดคล้องกับงานวิจัยของ เมษา นวลศรี (2559) ที่ได้ศึกษาเรื่อง การพัฒนาและตรวจสอบโครงสร้างพหุมิติของความเป็นพลเมืองที่มีความรับผิดชอบของนักเรียนชั้นมัธยมศึกษาตอนต้น: การประยุกต์ใช้แนวคิดการสร้างแผนที่โครงสร้าง โดยสร้างคะแนนจุดตัดของแบบวัดตามโครงสร้างพหุมิติของความเป็นพลเมืองที่มีความรับผิดชอบของนักเรียนชั้นมัธยมศึกษาตอนต้น พบว่า ผลการตรวจสอบคุณภาพความเชื่อมั่นของแบบวัดความเป็นพลเมืองที่มีความรับผิดชอบของนักเรียนมัธยมศึกษาตอนต้น มีค่าสัมประสิทธิ์ความเชื่อมั่น แบบ EAP โดยมีมิติด้านความรู้ มิติด้านทักษะ และมีมิติด้านจิตพิสัย มีค่าความเชื่อมั่นเท่ากับ .744, .757 และ .746 ตามลำดับ

## ข้อเสนอแนะ

### 1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 แบบทดสอบการคิดอภิमानแบบอัตนัยพหุมิติ สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลายนี้เป็นเพียงตัวอย่างหนึ่งเท่านั้น โดยแบบทดสอบมีลักษณะที่เป็นสถานการณ์ และมีคำถามที่มุ่งวัด 2 องค์ประกอบใน 1 ข้อ ผู้สนใจสามารถกำหนดสถานการณ์ขึ้นใหม่ที่ทันสมัยต่อเหตุการณ์ได้อย่างหลากหลายภายใต้โมเดลการคิดอภิमानแบบอัตนัยพหุมิตินี้ และสามารถนำไปพัฒนาแบบทดสอบในระดับชั้นอื่น ๆ โดยปรับเนื้อหาความยากง่ายให้เหมาะกับระดับชั้นนั้น ๆ

1.2 ครูผู้สอนสามารถนำ rubrics ไปใช้ในการพัฒนาหรือปรับปรุงความสามารถทางการคิดอภิमानของนักเรียนได้ โดยการตรวจให้คะแนนแบบรูบริคส์ทำให้ได้ทั้งข้อมูลเชิงคุณภาพที่ผนวกอยู่กับข้อมูลเชิงปริมาณมีประโยชน์ในการให้ข้อมูลป้อนกลับแก่ผู้ถูกประเมินซึ่งเป็นการตอบสนองหลักการของการประเมินผลเพื่อการปรับปรุง ช่วยให้ครูสามารถตั้งเป้าหมายทางการคิดของนักเรียนได้อย่างชัดเจน และสามารถให้นักเรียนเห็นได้อย่างชัดเจนว่าทำอะไรจึงสามารถคิดได้ตามความคาดหวังที่ตั้งไว้

### 2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

สำหรับการวิจัยครั้งนี้ ได้กำหนดเกณฑ์จุดตัดคะแนนความสามารถเพื่อเปรียบเทียบความสามารถของนักเรียนว่ามีความสามารถอยู่ในระดับใด ซึ่งจะต้องแปลคะแนนเป็นรายบุคคล ในกรณีที่นักเรียนจำนวนมากจะทำให้เสียเวลา ดังนั้น ในการวิจัยครั้งต่อไปจึงควรพัฒนาโปรแกรมเพื่อช่วยในการแปลความหมายได้อย่างรวดเร็วและสามารถวิเคราะห์ให้นักเรียนรายบุคคลได้ในคราวเดียวกัน

## เอกสารอ้างอิง

- ชัยวิชิต เขียรชนะ. (2552). *การพัฒนาแบบวัดกลยุทธ์การเรียนรู้แบบพหุมิติสำหรับนักเรียนมัธยมศึกษาตอนปลาย*. วิทยานิพนธ์ปริญญาครุศาสตรดุษฎีบัณฑิต สาขาการวัดและประเมินผลการจัดการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- บุษวรรษ์ แสนปลื้ม. (2556). *การศึกษาความสอดคล้องของการระบุความเข้มงวด/ใจดี และการทำหน้าที่ต่างกันของผู้ตรวจให้คะแนน*. *วารสารศิลปการศึกษาศาสตร์วิจัย*, 5(2), 335-347.
- เมษา นวลศรี. (2559). *การพัฒนาและตรวจสอบโครงสร้างพหุมิติของความเป็นพลเมืองที่มีความรับผิดชอบของนักเรียนชั้นมัธยมศึกษาตอนต้น: การประยุกต์ใช้แนวคิดการสร้างแผนที่โครงสร้าง*. วิทยานิพนธ์ปริญญาครุศาสตรดุษฎีบัณฑิต สาขาวิชาการวัดและประเมินผลการจัดการศึกษาภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์จุฬาลงกรณ์มหาวิทยาลัย.
- วนิดา ดีแป้น. (2556). *การสร้างโมเดลการตอบสนองรายข้อแบบพหุมิติ พหุระดับของพฤติกรรมความเป็นสมาชิกองค์กรของครู*. วิทยานิพนธ์ปริญญาครุศาสตรดุษฎีบัณฑิต สาขาการวัดและประเมินผลการจัดการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- วิเชียร เกตุสิงห์. (2538). *หลักการสร้างและวิเคราะห์เครื่องมือที่ใช้ในการวิจัย*. กรุงเทพฯ: ไทยวัฒนาพานิช.

- ศิริชัย กาญจนวาสี. (2556). *ทฤษฎีการทดสอบแบบดั้งเดิม*. พิมพ์ครั้งที่ 7. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2558). ข้อสอบอัตนัย: การสร้าง การวิเคราะห์ และการตรวจอย่างเป็นปรนัย (Constructed – Response Test Question: How to Construct, Analyze, and Score Item). *สารสมาคมวิจัยสังคมศาสตร์แห่งประเทศไทย*, 2(3), 16-27.
- สุกัญญา ทองนาค. (2555). *การพัฒนาแบบทดสอบสมรรถนะนักศึกษาตามมาตรฐานวิชาชีพครูแบบพหุมิติที่มีการตรวจให้คะแนนแบบพหุวิภาค*. วิทยานิพนธ์ปริญญาครุศาสตรดุษฎีบัณฑิต สาขาการวัดและประเมินผลการจัดการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- อัญชลี ศรีภักขาน (2547). การพัฒนาแบบวัดพหุปัญญาสำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น. *วารสารวิจัยและวัดผลการศึกษา*, 2(1), 112-129.
- Adams, R. J. (2005). Reliability as a measurement design effect. *Studies in Educational valuation*, 31, 162-172.
- Baker, L. and A. Brown. (1984). *Metacognitive Skills and Reading*. New York : Longman.
- Boyd, R. (1999). Homeostasis, Species and higher taxa. In R.A. Wilson(Ed), *Species : New interdisciplinary essays* (pp.141-185). Cambridge, MA : MIT Press.
- Brown, H.D. (1987). *Principles of Language Learning and Teaching*. New Jersey : Prentice Hall.
- Cross, D.R. and S.G. Paris. (1988). Developmental and Instructional Analyses of Children's Metacognition and Reading Comprehension. *Journal of Educational Psychology*, 80(2), 131.
- Flavell, J. H. (1979). Metacognition and Cognitive Monitoring : A new Are of Cognitive-development Inquiry. *American Psychologist*, 34(10), 906-911.
- Flavell, J. H., Miller, P. H., & Miller, S. A. (1985). *Cognitive development*. Englewood Cliffs, NJ: Prentice-Hall.
- Frey, A. and N.N., Seitz. (2009). Multidimensional Adaptive Testing in Educational and Psychological Measurement: Current State and Future Challenges. *Studies in Educational Evaluation*, 35, 89-94.
- Hoskens, M., and Boeck, P.D. (2001). Multidimensional componential item response theory models for polytomous items. *Applied Psychological Measurement*, 25(1), 19-37.
- Hair, J.F., Black, W.C., Babin, B.J., & Anderson, R.E. (2010). *Multivariate data analysis*. 7th ed. Prentice Hall.
- Hu L., and Bentler P.M. (1999). Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.



- Jacobs, J.E., and Paris, S.G. (1987). Children's metacognition about reading: Issue in definition, measurement and instruction. *Educational Psychologist*, 22, 255-278.
- Kelloway, E. (1998). *Using LISREL for structural equation modeling: A researcher's guide*. Thousand Oaks: SAGE Publications.
- Kline R.B. (1998). *Principles and practice of structural equation modeling*. Guilford Press.
- Lee, M. and A.L. Baylor. (2006). Designing Metacognitive Maps for Web-Based Learning. *Educational Technology and Society*, 9(1), 344-348.
- McCormick, R.J. (2009). *Applied Muscle Biology and Meat Science*. Boca Raton : CRC Press.
- McIntire, S.A. and L.A. Miller. (2007). *Foundations of Psychological Testing : A Practical Approach*. 2nd ed. California : SAGE.
- Moskal, Barbara M. (2000). Scoring rubrics: what, when and how?. *Practical Assessment, Research & Evaluation*. [Online]. Available from : <http://PAREonline.net/getvn.asp?v=7&n=3>. [accessed 1 August 2019]
- Nunnally, J.C. and Bernstein, I.H. (1994) The Assessment of Reliability. *Psychometric Theory*, 3, 248-292.
- Pintrich, P.R. (2000). Multiple Goals, Multiple Pathways : The Role of Goal Orientation in Learning and Achievement. *Journal of Educational Psychology*, 92(3), 544-555.
- Schraw, G. and R.S. Dennison. (1994). Assessing Metacognitive Awareness. *Contemporary Educational Psychology*, 19(4), 460-475.
- Veenman, P. (2006). Animal Physiotherapy. *Journal of Bodywork and Movement Therapies*, 10(4), 317-327.
- Yang, C. (2009). A Study of Metacognitive Strategies Employed by English Listeners in an EFL Setting. *International Education Studies*, 2(4), 134-139.