การเปรียบเทียบคุณภาพเกณฑ์การให้คะแนนแบบภาพรวมและแบบแยกองค์ประกอบในการประเมินโครงงานวิทยาศาสตร์ ระดับมัธยมศึกษาตอนปลาย

Comparison of Quality of Holistic and Analytic Scoring Rubrics in Evaluation of Higher Secondary School Science Projects

ศุภมาส ห้อยพรมราช[1*] และ จตุภูมิ เขตจัตุรัส[2]
Supphamas Hoiprommarat[1*] and Jatuphum Ketchatturat[2]
(Received : May 15, 2020 ; Revised : July 17, 2020 ; Accepted : July 21, 2020)

บทคัดย่อ

การวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ สำหรับนักเรียนระดับมัธยมศึกษา หาคุณภาพของเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ ที่มีรูปแบบของเกณฑ์การให้คะแนนต่างกัน โดยใช้โมเดลหลายองค์ประกอบของราส์ช และเปรียบเทียบคุณภาพของเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ ที่มีรูปแบบของเกณฑ์การให้คะแนนต่างกัน กลุ่มเป้าหมาย คือ ครูผู้สอนวิทยาศาสตร์ และ ครูที่เคยเป็นกรรมการตัดสินโครงงานวิทยาศาสตร์ จำนวน 8 ท่าน(จากการเลือกแบบเจาะจง) นักเรียนที่ได้รับการประเมินโดยใช้เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ ทั้งหมด 29 กลุ่ม เครื่องมือวิจัย คือ เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบภาพรวม (holistic rubric) และ เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบแยกองค์ประกอบ (analytic rubric) วิเคราะห์ข้อมูลโดยใช้โมเดลหลายองค์ประกอบของราส์ช(MFRM) ซึ่งพิจารณา ดัชนีความเที่ยง (reliability of separation index) โดยจำแนกตามประเภทของแบบประเมินทั้ง 2 ประเภท ความเข้มงวดหรือความใจดีของผู้ตรวจ (rater leniency/ severity) และความเชื่อมั่นแยกส่วน (separation reliability) ของผู้ตรวจและผู้สอบ

ผลการวิจัย พบว่า การพัฒนาเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย มี 2 รูปแบบ คือ เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบแยกองค์ประกอบ และเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบภาพรวม โดยประเด็นรายการเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ มีทั้งหมด 5 ประเด็น คือ 1) ความคิดสร้างสรรค์ 2) วิธีการศึกษาค้นคว้า 3) การเขียนรายงาน

[1] นักศึกษาปริญญาโท สาขาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น
[2] ผู้ช่วยศาสตราจารย์ สาขาการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น
[1] Master's degree student, Educational Measurement and Evaluation Program, Faculty of Education, Khon Kaen University
[2] Assistant Professor, Educational Measurement and Evaluation Program, Faculty of Education, Khon Kaen University
* Corresponding Author E-mail: supphamas_h@kkumail.com

4) การจัดแสดงผลงาน และ 5) การนำเสนอ กำหนดระดับคะแนน เป็น 5 ระดับคะแนน คือ 1 , 2 , 3 , 4 , 5 คะแนน ตามลำดับ การตรวจสอบความตรงเชิงเนื้อหา (content Validity) ของเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ มีค่า IOC อยู่ที่ 0.93 และเกณฑ์การให้คะแนนแบบภาพรวม มีค่า IOC อยู่ที่ 1.00

ผลการตรวจสอบคุณภาพของเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ พบว่าค่าดัชนีความสอดคล้องระหว่างผู้ประเมิน (Rater Agreement Index - RAI) ของเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบแยกองค์ประกอบมีค่าเท่ากับ 0.873 และเกณฑ์การให้คะแนนแบบภาพรวมมีค่าเท่ากับ 0.865 แสดงว่าเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ทั้ง 2 รูปแบบที่สร้างขึ้นมีความเที่ยงในการให้คะแนนและผู้ประเมินสามารถประเมินได้สอดคล้องกันอยู่ในระดับสูง เมื่อพิจารณาความเหมาะสมของผู้ประเมิน พบว่าผู้ประเมินสามารถประเมินได้สอดคล้องกับความสามารถของผู้ถูกประเมิน (ค่า INFIT และ OUTFIT อยู่ระหว่าง 0.6 – 1.4) ผู้ประเมินทุกคนมีคุณลักษณะเข้มงวด (ค่า Measure มากกว่า +1.00 ทุกคน)

ผลการเปรียบเทียบคุณภาพเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบแยกองค์ประกอบ (analytic) และแบบภาพรวม (holistic) วิเคราะห์ด้วยโมเดลหลายองค์ประกอบของราส์ช (MFRM) พบว่าผู้ประเมินที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic) มีค่า INFIT = 0.69 ค่า OUTFIT = 0.66 และผู้ประเมินที่ใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic) มีค่า INFIT = 1.82 ค่า OUTFIT = 1.92 แสดงว่า เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบแยกองค์ประกอบ (analytic) สามารถประเมินได้สอดคล้องกับความสามารถของผู้ถูกประเมินมากกว่าเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบภาพรวม (holistic)

คำสำคัญ : เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ เกณฑ์การให้คะแนนแบบภาพรวม
เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ โมเดลหลายองค์ประกอบของราช์ส

Abstract

The objectives of this research were to develop scoring rubrics of the science projects of secondary school students, to find the quality of the science project scoring rubrics with different scoring patterns by using multi-components of the Rasch model and compare the quality of the science project scoring rubrics with different scoring patterns. The target group consisted of 8 science teachers and teachers who used to be a judge for science projects, selected through purposive sampling, and 29 groups of students who were evaluated with science project scoring rubrics. The research instruments were a holistic science project scoring rubric (holistic scoring rubric) and a scoring rubric for modular science projects (analytic scoring rubric). The analysis employed Many-Facet Rasch Model (MFRM), considering the reliability of separation index classified by both types of the assessments, rater severity/leniency of the judge, separation reliability of the judge and the students.

The development of the science project scoring rubrics for secondary school students yielded 2 formats of rubrics: a scoring rubric for modular science projects, and a holistic science project scoring rubric. There were 5 facets in the science project scoring: 1) creative thinking, 2) method of study, 3) report writing, 4) exhibition, and 5) presentation. The scores were 1, 2, 3, 4, 5 points, respectively. In examining the content validity of the analytic scoring rubric, the IOC was 0.93 while the IOC of the holistic scoring rubric was 1.00.

The examination of the quality of the science project scoring rubrics revealed the rater agreement index (RAI) of the analytic scoring rubric at 0.873 (RAI=0.873) and the holistic scoring rubric at 0.865 (RAI=0.865). This indicates that both scoring rubrics of science projects had reliability and therefore the assessors could assess the science project with consistency at a high level. Considering the suitability of the assessors; it was found that the assessors could assess consistently with the students' ability (INFIT and OUTFIT were between 0.6 – 1.4). All of the assessors had severity characteristics (Measure value of each was over +1.00).

The comparison of the quality of the science project scoring rubrics between the scoring rubric for modular science projects and the holistic science project scoring rubric, using the multi-components of the Rasch model (MFRM) revealed that the assessors using the analytic scoring rubric form had the INFIT = 0.69 and the OUTFIT = 0.66. And the assessors using the holistic scoring rubric form had the INFIT = 1.82 and the OUTFIT = 1.92. This indicates that the science project analytic scoring rubric form could yield more consistency with the ability of the students than the science project holistic scoring rubric form.

Keywords: science project scoring rubric, holistic scoring rubric, analytic scoring rubric, Many-Facet Rasch Model (MFRM)

บทนำ

เกณฑ์การให้คะแนน (scoring rubric) เป็นแนวทางการให้คะแนน ประกอบด้วย รายการประเมินหรือเกณฑ์การพิจารณา และคำอธิบายคุณภาพของรายการประเมินแต่ละระดับคุณภาพ ซึ่งสามารถอธิบายหรือแยกแยะระดับต่างๆของความสำเร็จในการเรียน หรือการปฏิบัติของนักเรียนได้อย่างชัดเจน วิธีการประเมินผลงาน ผู้ประเมินจะต้องตีค่าผลการเรียนรู้ของผู้เรียนออกมาเป็นตัวเลขหรือระดับคะแนน ถ้าไม่มีเกณฑ์การให้คะแนนที่แน่นอน คะแนนที่ได้ออกมาจะมีความน่าเชื่อถือต่ำหรือมีความคลาดเคลื่อนในการตรวจให้คะแนนสูง ถ้าให้ผู้ประเมิน 2 คน ประเมินผลการเรียนรู้ของนักเรียนคนเดียวกัน ผลการประเมินอาจแตกต่างกันมาก การมีเกณฑ์การให้คะแนนเป็นเครื่องมือในการตรวจให้คะแนน จะทำให้มีการตีค่าคะแนนจากลักษณะผลการเรียนรู้มีความเชื่อมั่นสูงขึ้นและถูกต้องมากขึ้น จากการศึกษา พบว่า เกณฑ์การให้คะแนน scoring

rubric) แบ่งออกเป็น 2 รูปแบบ คือ 1) เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubrics) 2) เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubrics) (กนิษฐา สวยสด, 2554) (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2549) (ไซลัน สาและ, มปป) และ (ศุภกิจ ศรีจองแสง, 2553)

โมเดลหลายองค์ประกอบของราส์ช (Many-Facet Rasch Model: MFRM) ซึ่งพัฒนาขึ้นในปี 1989 โดยลินาเคอร์ (Linacre, 1994) เป็นโมเดลที่ปรับขยายขีดความสามารถมาจากโมเดลพื้นฐานของราส์ช (Rasch Model) แบบ 1 พารามิเตอร์ จากเดิมที่สามารถประมาณค่าพารามิเตอร์ของข้อสอบที่แสดงถึงความยากของข้อสอบและพารามิเตอร์ของผู้สอบที่แสดงถึงความสามารถของผู้สอบ พัฒนาจนสามารถที่จะประมาณค่าพารามิเตอร์ของผู้ตรวจ ซึ่งแสดงถึงความเข้มงวดของผู้ตรวจ (rater severity) หรือเพิ่มองค์ประกอบอื่น ๆ ที่สนใจศึกษาเข้าไปในโมเดลได้ จึงทำให้สามารถตรวจสอบอิทธิพลของผู้ตรวจที่มีต่อการวัดประเมินผลได้ นอกจากนี้ MFRM จะปรับแก้คะแนนที่เกิดจากความแตกต่างของความเข้มงวดของผู้ตรวจและความยากของข้อสอบ ทำให้คะแนนที่ได้มีความน่าเชื่อถือและยุติธรรมมากขึ้น โดยจุดเด่นของ MFRM คือเน้นศึกษาถึงแต่ละสมาชิกในองค์ประกอบที่เป็นแหล่งของความคลาดเคลื่อน เช่น พฤติกรรมการตรวจให้คะแนนของผู้ตรวจแต่ละคนว่ามีความเข้มงวดมากน้อยเพียงใด เพื่อพัฒนาคุณภาพของแต่ละสมาชิกในระบบการวัดประเมิน (Iramaneerat et al., 2007 อ้างอิงจาก น้ำผึ้ง อินทะเนตร, 2554) ทำให้การตรวจให้คะแนนมีความเป็นปรนัยมากขึ้น นั่นคือการวิเคราะห์โดย MFRM ทำให้เห็นภาพส่วนต่างๆ อันเป็นสารสนเทศที่มีประโยชน์ ไม่เฉพาะภาพรวมบนเส้นความสามารถเดียวกันเท่านั้น แต่ยังมีรายละเอียดย่อยของแต่ละองค์ประกอบ อีกทั้งสามารถประมาณค่าพารามิเตอร์ของผู้ตรวจ ทำให้ผู้วิจัยได้เห็นรายละเอียดและอภิปรายผลอย่างทั่วถึงของผลการวิเคราะห์คะแนนจากแบบทดสอบที่เกี่ยวข้องกับผู้ตรวจ

ในการประเมินโครงงานวิทยาศาสตร์ หน่วยงานต่างๆที่เกี่ยวข้องจะมีการพัฒนาเกณฑ์ขึ้นมาใช้ในหน่วยงานของตนเอง ซึ่งแต่ละหน่วยงานก็จะมีเกณฑ์ที่แตกต่างกัน ยังไม่มีการพัฒนาเกณฑ์การให้คะแนนที่สามารถใช้ได้ทุกหน่วยงาน และส่วนใหญ่จะใช้เกณฑ์การประเมินแบบแยกองค์ประกอบในการประเมินโครงงานวิทยาศาสตร์ สำหรับโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 31 ได้มีการประกวดแข่งขันโครงงานวิทยาศาสตร์ในหลายกิจกรรม ซึ่งแต่ละครั้งจะมีการใช้เกณฑ์ที่แตกต่างกันไป ไม่มีเกณฑ์ที่เป็นมาตรฐานในการประเมินโครงงานวิทยาศาสตร์แต่ละครั้ง

จากความเป็นมาและความสำคัญของปัญหา ผู้วิจัยได้เล็งเห็นว่า การประเมินโครงงานวิทยาศาสตร์ควรมีความถูกต้อง เชื่อถือได้ จึงต้องมีการพัฒนาเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ และหากผู้ประเมินใช้รูปแบบเกณฑ์การให้คะแนนที่ต่างกัน ผลการประเมินจะมีความแตกต่างกันมากน้อยเพียงใด ซึ่งการนำโมเดลหลายองค์ประกอบของราส์ช (MFRM) มาใช้ตรวจสอบนั้น จะเป็นวิธีการทางสถิติที่สามารถค้นหาตรวจสอบแหล่งความแปรปรวนได้หลายแหล่ง เพื่อให้ได้เกณฑ์การให้คะแนนที่มีความถูกต้อง และเชื่อถือได้ในการนำไปใช้ประเมินโครงงานวิทยาศาสตร์ต่อไป

ความมุ่งหมายของการวิจัย

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อการเปรียบเทียบคุณภาพเกณฑ์การให้คะแนนแบบภาพรวมและแบบแยกองค์ประกอบในการประเมินโครงงานวิทยาศาสตร์ ระดับมัธยมศึกษาตอนปลาย ซึ่งมีวัตถุประสงค์ย่อยดังนี้

1. เพื่อพัฒนาเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ สำหรับนักเรียนระดับมัธยมศึกษาตอนปลาย
2. เพื่อหาคุณภาพของเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบภาพและแบบแยกองค์ประกอบโดยใช้โมเดลหลายองค์ประกอบของราส์ช
3. เพื่อเปรียบเทียบคุณภาพของเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบภาพและแบบแยกองค์ประกอบ

วิธีดำเนินการวิจัย

1. กลุ่มเป้าหมาย

กลุ่มเป้าหมาย ในการวิจัยครั้งนี้ ผู้วิจัยใช้วิธีการเลือกแบบเจาะจง (purposive sampling) โดยกลุ่มเป้าหมายที่ผู้วิจัยได้รับจากการคัดเลือกจากผู้เชี่ยวชาญในโรงเรียนหนองบัวพิทยาคารเป็นผู้คัดเลือกให้ ซึ่งเป็นครูผู้สอนวิทยาศาสตร์ ที่มีประสบการณ์สอนไม่น้อยกว่า 3 ปี จำนวน 4 คน ครูที่เคยเป็นกรรมการตัดสินโครงงานวิทยาศาสตร์ จำนวน 4 คน โรงเรียนหนองบัวพิทยาคาร สังกัดสำนักงานเขตพื้นการศึกษามัธยมศึกษาเขต 19 และ นักเรียนที่ได้รับการประเมินโดยใช้เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ ทั้งหมด 29 กลุ่ม

2. เครื่องมือที่ใช้ในการวิจัย

2.1 คู่มือการใช้เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ทั้งแบบแยกองค์ประกอบ และแบบภาพรวม

2.2 แบบบันทึกการสนทนากลุ่ม

3. การเก็บรวบรวมข้อมูล

รายละเอียดเกี่ยวกับวิธีการดำเนินการวิจัยในแต่ละขั้นตอน มีดังนี้

ระยะที่ 1 การพัฒนาเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์

ในการพัฒนาเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ มีรายละเอียด ดังนี้

1. การร่างประเด็นที่ใช้ประกอบเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์

1.1 ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องเกี่ยวกับการสร้างและพัฒนาเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ทั้งในและต่างประเทศ

1.2 กำหนดกรอบแนวคิดการพัฒนาประเด็นที่ใช้ในประเมินโครงงานวิทยาศาสตร์

1.3 วิเคราะห์และสังเคราะห์โดยการร่างประเด็นที่ใช้ในประเมินโครงงานวิทยาศาสตร์

1.4 ผู้ทรงคุณวุฒิร่วมวิพากษ์เพื่อ เสนอแนะ ปรับปรุง ประเด็นประเมินโครงงานวิทยาศาสตร์

2. เมื่อได้ประเด็นประเมินโครงงานวิทยาศาสตร์ จึงนำประเด็นประเมินมาสร้างเกณฑ์การให้คะแนนทั้ง 2 รูปแบบ โดยผู้วิจัยร่วมวิพากษ์กับผู้ทรงคุณวุฒิ เพื่อนำข้อเสนอแนะมาปรับปรุงแก้ไขให้สมบูรณ์ที่สุด

3. นำเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ ทั้ง 2 รูปแบบที่ผู้วิจัยวิพากษ์ร่วมกับผู้ทรงคุณวุฒิ ไปตรวจสอบคุณภาพ ความเหมาะสมและความเป็นไปได้โดยผู้เชี่ยวชาญ

4. ตรวจสอบคุณภาพรูปแบบเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์

4.1 ตรวจสอบคุณภาพของในเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ เป็นการตรวจสอบความตรงเชิงเนื้อหา ความเหมาะสมและความเป็นไปได้ของประเด็นที่ใช้ในการสร้างเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ โดยระบบผู้เชี่ยวชาญ (connoisseurship) โดยให้ผู้เชี่ยวชาญ ทำการตรวจสอบคุณภาพ โดยคำนวณค่าดัชนีความสอดคล้อง (IOC) การกำหนดเกณฑ์ ข้อคำถามที่มีค่า IOC มากกว่าหรือเท่ากับ .50 ถือว่าข้อคำถามมีความเหมาะสม หากข้อคำถามมีค่าไม่ถึง .50 ควรปรับปรุงหรือตัดทิ้ง (สุวิมล ติรกานันท์, 2551)

4.2 ตรวจสอบความเหมาะสมและความเป็นไปได้ของเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ โดยสอบถามความคิดเห็นจากผู้เชี่ยวชาญ การวิเคราะห์ข้อมูลโดยการกำหนดเกณฑ์โดยใช้การพิจารณาหาค่าเฉลี่ย ได้ค่าเฉลี่ย ≥ 3.50 ขึ้นไปทั้งความเหมาะสมและความเป็นไปได้จากการกำหนดคะแนน 5 ระดับ คือ 1-5 แล้ว จะนำประเด็นดังกล่าวมาเป็นประเด็นของตัวแปรในรูปแบบเกณฑ์การให้คะแนนต่อไป

4.3 ปรับปรุงรูปแบบเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ที่ใช้ในการประเมินโครงงานวิทยาศาสตร์ให้เหมาะสม

ระยะที่ 2 การหาคุณภาพ และวิเคราะห์ขนาดความแปรปรวนในแต่ละองค์ประกอบเพื่อเปรียบเทียบคุณภาพของเกณฑ์การให้คะแนนทั้ง 2 รูปแบบ

ในการหาคุณภาพวิเคราะห์ขนาดความแปรปรวนในแต่ละองค์ประกอบ เพื่อเปรียบเทียบคุณภาพของเกณฑ์การให้คะแนนทั้ง 2 รูปแบบ มีรายละเอียด ดังนี้

1. นำรูปแบบเกณฑ์การให้คะแนนไปใช้จริง

2. หาคุณภาพของเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ และวิเคราะห์ขนาดความแปรปรวนในแต่ละองค์ประกอบของเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ เมื่อรูปแบบของเกณฑ์การให้คะแนนต่างกัน โดยใช้โมเดลหลายองค์ประกอบของราส์ช

3. เปรียบเทียบคุณภาพของเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ทั้ง 2 รูปแบบ

4. ประชุมสนทนากลุ่มผู้ใช้รูปแบบประเมิน

5. ปรับปรุงและเผยแพร่รูปแบบเกณฑ์การให้คะแนนวิทยาศาสตร์

4. การวิเคราะห์ข้อมูล

ผู้วิจัยวิเคราะห์ข้อมูลเพื่อตอบวัตถุประสงค์ของการวิจัย ดังนี้

4.1 นำข้อมูลที่ได้จากการสังเคราะห์งานวิจัย บันทึกการสนทนากลุ่มจากการวิพากษ์มาสร้างเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ทั้งสองรูปแบบ

4.2 นำเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ทั้งสองรูปแบบ มาวิเคราะห์หาความตรงเชิงเนื้อหา (IOC) วิเคราะห์ความเหมาะสมและความเป็นไปได้ของรูปแบบเกณฑ์การให้คะแนน และประเด็นรายการและรายละเอียดของประเด็นการสร้างเกณฑ์การให้คะแนน

4.3 วิเคราะห์หาความเที่ยงของเกณฑ์การให้คะแนน โดยใช้วิธีหาดัชนีความสอดคล้องระหว่างผู้ประเมิน (Rater Agreement Index – RAI)

4.4 ตรวจสอบคุณภาพของเครื่องมือ (เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์) โดยใช้โมเดลหลายองค์ประกอบของราส์ช (MFRM) เพื่อตรวจสอบความเหมาะสมกลมกลืนของคะแนนเชิงประจักษ์กับค่าที่คาดหวังโมเดล Partial Credit Model ซึ่งเป็นการตรวจสอบว่าผู้ประเมินสามารถประเมินได้สอดคล้องกับผู้ถูกประเมินหรือไม่ และตรวจสอบคุณลักษณะของผู้ประเมิน (ความเข้มงวด/ใจดี)

ผลการวิจัย

1. ผลการพัฒนาเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์

1.1 ผลการพัฒนารูปแบบและประเด็นรายการเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์

รูปแบบและประเด็นรายการที่ใช้ในเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ทั้งสองรูปแบบได้จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง และนำไปวิพากษ์ร่วมกับผู้ทรงคุณวุฒิและผู้เชี่ยวชาญ ซึ่งได้ผลสรุปรูปแบบเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ และประเด็นรายการและรายละเอียดในแต่ละประเด็นดังนี้

1.1.1 รูปแบบเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์

เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย มี 2 รูปแบบ คือ เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบแยกองค์ประกอบ และ เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบภาพรวม

1.1.2 ประเด็นรายการเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์

ประเด็นรายการเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย ได้จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง และร่วมกันพัฒนาประเด็นการให้คะแนนร่วมกับผู้เชี่ยวชาญ ได้ประเด็นรายการเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ ทั้งหมด 5 ประเด็น คือ 1) ความคิดสร้างสรรค์ 2) วิธีการศึกษาค้นคว้า 3) การเขียนรายงาน 4) การจัดแสดงผลงาน และ 5) การนำเสนอ โดยกำหนดระดับคะแนน เป็น 5 ระดับคะแนน คือ 1 , 2 , 3 , 4 , 5 คะแนน ตามลำดับ

1.2 ผลการตรวจสอบความตรงเชิงเนื้อหา (IOC) ความเหมาะสมความเป็นไปได้ของรูปแบบเกณฑ์และประเด็นรายการให้คะแนนโครงงานวิทยาศาสตร์

การตรวจสอบคุณภาพของเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ เป็นการตรวจสอบความตรงเชิงเนื้อหา ความเหมาะสมและความเป็นไปได้ของประเด็นที่ใช้ในเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ โดยระบบผู้เชี่ยวชาญ (connoisseurship) ด้วยการประเมินเชิงเหตุผล (logical evaluation) ดังนี้

1.2.1 การตรวจสอบความตรงเชิงเนื้อหา (content validity) ของประเด็น และรายละเอียดของประเด็นที่ใช้ในเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ โดยให้ผู้เชี่ยวชาญทำการตรวจสอบคุณภาพโดยคำนวณค่าดัชนีความสอดคล้อง (IOC) ผลการตรวจสอบพบว่าเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบแยกองค์ประกอบ มีค่า IOC อยู่ที่ 0.93 และเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบภาพรวม มีค่า IOC อยู่ที่ 1.00 ซึ่งอยู่ในเกณฑ์ที่รับได้ โดยอ้างอิงจาก (สุวิมล ติรกานันท์, 2551) ค่า IOC $\geq$ .50 ถือว่ามีความเหมาะสม ค่า IOC $<$ .50 ควรปรับปรุงหรือตัดทิ้ง และเมื่อได้คะแนนมาแล้ว จึงนำไปวิพากษ์จากผู้เชี่ยวชาญ จะเห็นได้ว่า การแบ่งแบบ 5 ระดับ จะสามารถประเมินได้มีประสิทธิภาพมากกว่า 4 ระดับและ 3 ระดับ ผู้วิจัยจึงได้ตัดสินใจเลือกรูปแบบเกณฑ์การให้คะแนนทั้งแบบแยกองค์ประกอบ (analytic) และแบบภาพรวม (holistic) โดยใช้ระดับการประเมิน 5 ระดับ มาใช้ในการเปรียบเทียบคุณภาพเกณฑ์การให้คะแนน

1.2.2 ตรวจสอบความเหมาะสมและความเป็นไปได้ของประเด็น และรายละเอียดของประเด็นที่ใช้ในเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ โดยสอบถามความคิดเห็นจากผู้เชี่ยวชาญ ทำการตรวจสอบโดยใช้การพิจารณาค่าเฉลี่ย หากค่าเฉลี่ย ≥ 3.50 ขึ้นไป ทั้งความเหมาะสมและความเป็นไปได้จากการกำหนดคะแนน 5 ระดับ คือ 1-5 แล้วนำประเด็นดังกล่าวมาเป็นประเด็นของตัวแปรในรูปแบบเกณฑ์การให้คะแนนต่อไป ซึ่งผลการตรวจสอบดังตาราง 1

ตาราง 1 แสดงสถิติของความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับความเหมาะสมและความเป็นไปได้ของประเด็นที่ใช้ในเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์

ข้อ	ประเด็นที่ใช้ในเกณฑ์การให้คะแนน	ค่าเฉลี่ย	ระดับความเหมาะสมและความเป็นไปได้
1	ความคิดสร้างสรรค์	4.2	มาก
2	วิธีการศึกษาค้นคว้า	4.65	มากที่สุด
3	การเขียนรายงาน	3.95	มาก
4	การจัดแสดงผลงาน	4	มาก
5	การนำเสนอ	4.9	มากที่สุด

จากตาราง 1 แสดงว่า ผู้เชี่ยวชาญให้ความคิดเห็นเกี่ยวกับประเด็นที่ใช้ในเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ มีความเหมาะสมและความเป็นไปได้ในระดับมาก และมากที่สุด แสดงว่า ประเด็นที่ใช้ในเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ สามารถนำไปใช้ในการให้คะแนนโครงงานวิทยาศาสตร์ได้

จากการตรวจสอบความตรงเชิงเนื้อหา ความเหมาะสมและความเป็นไปได้ของประเด็นที่ใช้ในเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ โดยระบบผู้เชี่ยวชาญ (connoisseurship) ผู้วิจัยจึงได้จัดทำ เกณฑ์การประเมินโครงงานวิทยาศาสตร์ระดับชั้นมัธยมศึกษา เพื่อใช้การประเมินโครงงานวิทยาศาสตร์ ซึ่งประเด็นประเมินและรายละเอียดในแต่ละระดับคะแนนของรายการประเด็น มีรายละเอียดดังต่อไปนี้

ตาราง 2 แสดงประเด็นประเมินและรายละเอียดในแต่ละระดับของรายการประเมิน

ประเด็นประเมิน	ระดับคะแนน
ความคิดสร้างสรรค์	5---เป็นประเด็นที่มีความแปลกใหม่ แตกต่าง และน่าสนใจ สามารถนำผลการศึกษาค้นคว้าไปต่อยอดจนเกิดสิ่งประดิษฐ์ใหม่ นวัตกรรมใหม่ ความรู้ใหม่ วิธีการใหม่ 4---เป็นประเด็นที่มีความแปลกใหม่ แตกต่าง แต่ไม่น่าสนใจ สามารถนำผลการศึกษาค้นคว้าไปต่อยอดจนเกิดสิ่งประดิษฐ์ใหม่ นวัตกรรมใหม่ ความรู้ใหม่ วิธีการใหม่ 3---เป็นประเด็นที่มีความแปลกใหม่ ไม่มีความแตกต่าง ไม่น่าสนใจ แต่สามารถนำผลการศึกษาค้นคว้าไปต่อยอดจนเกิดสิ่งประดิษฐ์ใหม่ นวัตกรรมใหม่ ความรู้ใหม่ วิธีการใหม่ 2---ไม่ใช่ประเด็นที่มีความแปลกใหม่ แตกต่าง ไม่น่าสนใจ แต่สามารถนำผลการศึกษาค้นคว้าไปต่อยอดจนเกิดสิ่งประดิษฐ์ใหม่ นวัตกรรมใหม่ ความรู้ใหม่ วิธีการใหม่ 1---ไม่ใช่ประเด็นที่มีความแปลกใหม่ แตกต่าง ไม่น่าสนใจ ไม่สามารถนำผลการศึกษาค้นคว้าไปต่อยอดจนเกิดสิ่งประดิษฐ์ใหม่ นวัตกรรมใหม่ ความรู้ใหม่ วิธีการใหม่ ได้
วิธีการศึกษาค้นคว้า	5---มีวิธีการศึกษาค้นคว้าครบ 5 ข้อ คือ การกำหนดประเด็นปัญหา ตั้งสมมติฐาน ตรวจสอบสมมติฐาน วิเคราะห์ข้อมูล และสรุปผลการศึกษาค้นคว้า 4---มีวิธีการศึกษาค้นคว้า 4 ใน 5 ข้อ 3---มีวิธีการศึกษาค้นคว้า 3 ใน 5 ข้อ 2---มีวิธีการศึกษาค้นคว้า 2 ใน 5 ข้อ 1---มีวิธีการศึกษาค้นคว้า 1 ใน 5 ข้อ
การเขียนรายงาน	5---การเขียนรายงาน ครอบคลุมทั้ง 5 ข้อ ดังนี้ แบบฟอร์มถูกต้อง ครอบคลุมทุกหัวข้อ มีการสรุปและอภิปรายผล มีข้อเสนอแนะ มีการแสดงหลักฐานการบันทึกข้อมูลเกี่ยวกับสิ่งที่ศึกษา มีเอกสารอ้างอิง 4---มีการเขียนรายงานครอบคลุม 4 ใน 5 ข้อ 3---มีการเขียนรายงานครอบคลุม 3 ใน 5 ข้อ 2---มีการเขียนรายงานครอบคลุม 2 ใน 5 ข้อ 1---มีการเขียนรายงานครอบคลุม 1 ใน 5 ข้อ
การจัดแสดงผลงาน	5---การจัดแสดงผลงาน ครอบคลุมทั้ง 5 ข้อ ดังนี้ เนื้อหาในการจัดแสดงครบถ้วน ลำดับขั้นตอนการจัดแสดงผลงานมีความต่อเนื่อง เชื่อมโยง การจัดองค์ประกอบสื่อและเนื้อหามีความชัดเจนสอดคล้องกับประเด็นที่ศึกษ

ตาราง 2 (ต่อ)

ประเด็นประเมิน	ระดับคะแนน
	สื่อมีเทคนิคที่หลากหลาย สร้างสรรค์ ชวนติดตาม แผงโครงงานมีขนาดตามที่กำหนด 4---การจัดแสดงผลงานครอบคลุม 4 ใน 5 ข้อ 3---การจัดแสดงผลงานครอบคลุม 3 ใน 5 ข้อ 2---การจัดแสดงผลงานครอบคลุม 2 ใน 5 ข้อ 1---การจัดแสดงผลงานครอบคลุม 1 ใน 5 ข้อ
การนำเสนอ	5---การนำเสนอครอบคลุมทั้ง 5 ข้อ ดังนี้ สามารถใช้ภาษาในการสื่อสารได้ถูกต้อง ชัดเจนและเข้าใจง่าย นำเสนอครบทุกองค์ประกอบ เนื้อหาครบถ้วน ตรงตามประเด็น ตอบคำถามตรงประเด็นพร้อมให้เหตุผลและยกตัวอย่างประกอบ มีความมั่นในใจการนำเสนอ และแก้ปัญหาเฉพาะหน้าได้ดี ใช้เวลาตรงตามที่กำหนด 4---การนำเสนอครอบคลุม 4 ใน 5 ข้อ 3---การนำเสนอครอบคลุม 3 ใน 5 ข้อ 2---การนำเสนอครอบคลุม 2 ใน 5 ข้อ 1---การนำเสนอครอบคลุม 1 ใน 5 ข้อ

2. ผลการตรวจสอบคุณภาพของเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์

2.1 ผลการตรวจสอบความเที่ยงของเกณฑ์การให้คะแนนโดยใช้วิธีหาดัชนีความสอดคล้องระหว่างผู้ประเมิน (Rater Agreement Index: RAI)

ผลการตรวจสอบความเที่ยงของเกณฑ์การให้คะแนน โดยใช้วิธีหาดัชนีความสอดคล้องระหว่างผู้ประเมิน (Rater Agreement Index : RAI) จากผู้ประเมินจำนวน 8 คน โดยผู้ประเมินแต่ละคนประเมินโครงงานวิทยาศาสตร์ จำนวน 29 กลุ่ม ผลการประเมินดังแสดงในตาราง 3

ตาราง 3 ผลการตรวจสอบความเที่ยงของเกรฑ์การให้คะแนนโครงงานวิทยาศาสตร์ โดยใช้วิธีการหาดัชนีความสอดคล้องระหว่างผู้ประเมิน (Rater Agreement Index : RAI)

รูปแบบเกณฑ์การให้คะแนน	ดัชนีความสอดคล้องระหว่างผู้ประเมิน
เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ	0.873
เกณฑ์การให้คะแนนแบบภาพรวม	0.865

จากตาราง 3 การตรวจสอบความเที่ยงของเกณฑ์การให้คะแนนโรงงานวิทยาศาสตร์ โดยการหาดัชนีความสอดคล้องระหว่างผู้ประเมินพบว่า พบว่า เมื่อผู้ประเมินใช้เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบแยกองค์ประกอบ ดัชนีความสอดคล้องระหว่างผู้ประเมิน อยู่ที่ 0.873 เมื่อผู้ประเมินใช้เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบภาพรวม ดัชนีความสอดคล้องระหว่างผู้ประเมิน อยู่ที่ 0.865 แสดงว่าผู้ประเมินให้คะแนนสอดคล้องกันสูงมาก

2.2 ผลการตรวจสอบคุณภาพเครื่องมือด้วยโมเดลหลายองค์ประกอบของราส์ช (Many-Facet Rasch Model : MFRM)

2.2.1 การตรวจสอบความเหมาะกลมกลืนของคะแนนเชิงประจักษ์กับค่าที่คาดหวังโมเดล Partial Credit Model

จากการวิเคราะห์ พบว่า ค่า infit ของเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubrics) และ เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubrics) เท่ากับ 0.69 และ 1.82 ตามลำดับ ค่า outfit ของเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubrics) และเกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubrics) เท่ากับ 0.66และ 1.96 ตามลำดับ ซึ่งค่าทั้งสองอยู่ในช่วง 0.5 ถึง 2.0 แสดงว่า ข้อมูลเชิงประจักษ์ (คะแนนจากการตรวจของผู้ตรวจ) กับค่าที่คาดหวังมีความสอดคล้องกัน

2.2.2 ผลการตรวจสอบความสอดคล้อง / คุณลักษณะของผู้ประเมิน

การระบุความเข้มงวด /ใจดี ในขั้นแรกนั้นต้องพิจารณาว่าผู้ตรวจให้คะแนนแต่ละคนให้คะแนนเชิงประจักษ์เหมาะสมกลมกลืนกับโมเดล Partial Credit Model หรือไม่ โดยพิจารณาค่า Weighted mean square residual : INFIT และ Unweighted mean square residual : OUTFIT ที่กำหนด ขอบเขตของการยอมรับได้ ของโมเดลที่เหมาะสมกับข้อมูลอยู่ระห่าง 0.5 ถึง 1.5 จากนั้นจึงระบุความเข้มงวด / ใจดี จากค่าศูนย์ (0) เครื่องหมายบวก (+) ลบ (-) ของค่า Logit โดยกำหนดให้ผู้ตรวจให้คะแนน ที่เข้มงวด จะมีค่า Logit มากกว่า +1.00 ให้คะแนนเป็นกลาง จะมีค่า Logit อยู่ระหว่าง ± 1 และให้คะแนนใจดี จะมีค่า Logit น้อยกว่า -1.00 (บุษวรรษ์ แสนปลื้ม, 2556) ผลการตรวจสอบดังแสดงในตาราง 4

ตาราง 4 แสดงผลคุณลักษณะของผู้ประเมิน (raters)

ผู้ประเมิน Raters	INFIT MnSq	OUTFIT MnSq	Logit (Measure)	คุณลักษณะ
R01	1.00	1.14	1.16	เข้มงวด
R02	1.17	1.29	2.13	เข้มงวด
R03	1.16	1.17	1.01	เข้มงวด
R04	.84	1.26	1.61	เข้มงวด
R05	.44	.32	3.10	เข้มงวด
R06	.95	.93	1.76	เข้มงวด
R07	.53	.49	2.27	เข้มงวด
R08	1.27	1.21	1.16	เข้มงวด

จากตาราง 4 พิจารณาความเหมาะสมของผู้ประเมินโดยพิจารณาจากค่า Weighted mean square residual : INFIT และ Unweighted mean square residual : OUTFIT พบว่า ผู้ประเมินส่วนใหญ่

มีค่า INFIT และ OUTFIT อยู่ในช่วง 0.6 ถึง 1.4 แสดงว่า ผู้ประเมินสามารถประเมินได้สอดคล้องกับความสามารถของผู้ถูกประเมิน (Linacre, 1994)

สำหรับคุณลักษณะของผู้ประเมิน ซึ่งจากตารางจะเห็นว่า ผู้ประเมินทุกคน มีค่า Measure มากกว่า +1.00 ทุกคน แสดงว่าผู้ประเมินทุกคน มีคุณลักษณะเข้มงวด

3. เปรียบเทียบคุณภาพของเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์

จากการทดลองใช้แบบประเมินโครงงานวิทยาศาสตร์ทั้งแบบแยกองค์ประกอบ (Analytic) และแบบภาพรวม (Holistic) นำผลไปวิเคราะห์แล้วพิจารณาว่าผู้ตรวจให้คะแนนแต่ละคนให้ คะแนนเชิงประจักษ์เหมาะสมกลมกลืนกับโมเดล Partial Credit Model หรือไม่ โดยพิจารณาค่า Weighted mean square residual : INFIT และ Unweighted mean square residual : OUTFIT ที่กำหนด ขอบเขตของการยอมรับได้ของโมเดลที่เหมาะสมกับข้อมูลอยู่ระหว่าง 0.5 ถึง 1.5 (บุษวรรษ์ แสนปลื้ม, 2556) ผลการเปรียบเทียบดังแสดงในตาราง 5

ตาราง 5 แสดงค่าดัชนีทดสอบความสอดคล้องของโมเดล

Criteria	INFIT MnSq	OUTFIT MnSq
Analytic	.69	.66
Holistic	1.82	1.96

จากตาราง 5 เมื่อพิจารณาค่า Weighted mean square residual : INFIT และ Unweighted mean square residual : OUTFIT พบว่า ผู้ประเมินที่ใช้แบบประเมินแบบแยกองค์ประกอบ (analytic) มีค่า INFIT = 0.69 ค่า OUTFIT = 0.66 (อยู่ในช่วงที่รับได้) และผู้ประเมินที่ใช้แบบประเมินแบบภาพรวม (holistic) มีค่า INFIT = 1.82 ค่า OUTFIT = 1.92 (มากกว่าค่าที่รับได้) แสดงว่า แบบประเมินโครงงานวิทยาศาสตร์แบบแยกองค์ประกอบ (analytic) สามารถประเมินได้สอดคล้องกับความสามารถของผู้ถูกประเมินมากกว่าแบบประเมินโครงงานวิทยาศาสตร์แบบภาพรวม (holistic)

อภิปรายผล

จากการพัฒนาเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ มี 2 รูปแบบ คือ เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบแยกองค์ประกอบ (analytic rubrics) และ เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบภาพรวม (holistic scoring) ซึ่งสอดคล้องกับ รูปแบบของของเกณฑ์การให้คะแนนของ กนิษฐา สวยสด (2554) สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (2549) ไชลัน สาและ (มปป) ศุภกิจ ศรีจองแสง (2553) และ Stenmark (1991) ซึ่งได้กำหนดเกณฑ์การให้คะแนนมีวิธีการ 2 แบบ คือ การกำหนดเกณฑ์โดยภาพรวม (holistic scoring) และ การกำหนดเกณฑ์โดยแยกองค์ประกอบ (analytic scoring)

ประเด็นที่ใช้ในการสร้างเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ ซึ่งได้จากการสังเคราะห์และการหาความตรงเชิงเนื้อหา จะได้ประเด็นที่ใช้ในการให้คะแนนโครงงานวิทยาศาสตร์ทั้งหมด 5 รายการ คือ 1) ความคิดสร้างสรรค์ 2) วิธีการศึกษาค้นคว้า 3) การเขียนรายงาน 4) การจัดแสดงผลงาน และ 5) การนำเสนอ ซึ่งสอดคล้องกับศูนย์วิทยาศาสตร์เพื่อการศึกษา (2559) ซึ่งได้กำหนดเกณฑ์การประกวดโครงงานวิทยาศาสตร์ ประจำปีงบประมาณ 2559 ดังนี้ ความคิดสร้างสรรค์ การใช้วิธีการทางวิทยาศาสตร์ การเขียนรายงาน/ สื่อความหมาย การจัดแสดงผลงาน และการนำเสนอ

การตรวจสอบคุณภาพเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ holistic และ analytic scoring rubric ด้วยโมเดลหลายองค์ประกอบของราส์ช (MFRM) เมื่อพิจารณาความเหมาะสมของผู้ประเมินพบว่า ผู้ประเมินสามารถประเมินได้สอดคล้องกับความสามารถของผู้ถูกประเมิน และจากการพิจารณาคุณลักษณะของผู้ประเมินพบว่า ผู้ประเมินทุกคน มีคุณลักษณะเข้มงวด เนื่องจากการประเมินในครั้งนี้เป็นการประเมินเพื่อนำคะแนนที่ได้ไปไปใช้ในการตัดสินผลการเรียนในรายวิชาของผู้เรียน ทำให้ผู้ประเมินทุกคนมีความเข้มงวดในการให้คะแนน

การเปรียบเทียบคุณภาพของเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์แบบแยกองค์ประกอบ (analytic) และแบบภาพรวม (holistic) นำผลไปวิเคราะห์ด้วยโมเดลหลายองค์ประกอบของราส์ช (MFRM) เมื่อพิจารณาค่า Weighted mean square residual : INFIT และ Unweighted mean square residual : OUTFIT พบว่า ผู้ประเมินที่ใช้แบบประเมินแบบแยกองค์ประกอบ (analytic) มีค่า INFIT = 0.69 ค่า OUTFIT = 0.66 (อยู่ในช่วงที่รับได้) และผู้ประเมินที่ใช้แบบประเมินแบบภาพรวม (holistic) มีค่า INFIT = 1.82 ค่า OUTFIT = 1.92 (มากกว่าค่าที่รับได้) แสดงว่า แบบประเมินโครงงานวิทยาศาสตร์แบบแยกองค์ประกอบ (analytic) สามารถประเมินได้สอดคล้องกับความสามารถของผู้ถูกประเมินมากกว่าแบบประเมินโครงงานวิทยาศาสตร์แบบภาพรวม (holistic) เนื่องจากการประเมินโครงงานวิทยาศาสตร์ เป็นการประเมินที่แยกออกเป็น 5 ด้าน ซึ่งแต่ละด้านจะมีรายละเอียดที่แตกต่างกันไป ทำให้ได้รับข้อมูลที่มีรายละเอียดซึ่งแสดงถึงจุดเด่นและจุดด้อยของผู้ถูกประเมินในแต่ละองค์ประกอบ ช่วยให้ครูเน้นหรือปรับปรุงจุดที่ผู้ถูกประเมินต้องปรับปรุงแก้ไขได้ตรงกับความต้องการของผู้ถูกประเมินมากขึ้น นำไปใช้ในการวางแผนและปรับปรุงการเรียนการสอน รวมถึงการให้ข้อมูลย้อนกลับแก่ผู้ถูกประเมินเพื่อพัฒนาหรือปรับปรุงตนเองได้อย่างชัดเจน (Stenmark, 1991) ซึ่งสอดคล้องกับ (ประทีป กาลเขว้า, 2559) ซึ่งได้พัฒนารูปแบบการประเมินทักษะการปฏิบัติทางวิชาชีพ ด้านการจัดการอนามัยสิ่งแวดล้อม การตรวจสอบคุณภาพเกณฑ์การให้คะแนนใช้โมเดลหลายองค์ประกอบของราส์ช และประยุกต์กรอบแนวคิดการประเมินของศูนย์ควบคุมและป้องกันโรคแห่งสหรัฐอเมริกา พบว่า การตรวจสอบคุณภาพเครื่องมือ ด้วยโมเดลหลายองค์ประกอบของราส์ช (Many – Facet Rasch Model : MFRM) ค่าความเข้มงวด/ใจดี (severity/leniency) ของการให้คะแนนแบบแยกส่วน (analytic scoring rubric) สูงกว่า การให้คะแนนแบบภาพรวม (holistic scoring rubric) (เท่ากับ 2.96 และ 2.14 ตามลำดับ) ดัชนีความเชื่อมั่นแยกส่วน (reliability of separation index) และความเชื่อมั่นแยกส่วนของผู้ตรวจ (rater separation reliability) (2.09 และ 0.92 ตามลำดับ) สูงกว่า การให้คะแนนแบบภาพรวม (holistic scoring rubric) (เท่ากับ 1.26 และ 0.61 ตามลำดับ) และสอดคล้องกับ Chi (2001) ได้ศึกษา

เปรียบเทียบวิธีการตรวจให้คะแนนแบบรวม (holistic method) กับการตรวจให้คะแนนแบบแยกส่วน (analytic method) ในการวัดประเมินการปฏิบัติโดยใช้โมเดลหลายองค์ประกอบของราส์ช พบว่า การให้คะแนนแบบแยกส่วนมีความคงเส้นคงวาของเข้มงวด มากกว่าการให้คะแนนแบบรวม ช่วงของความเข้มงวดของผู้ตรวจให้คะแนนแบบรวมกว้างกว่าการให้คะแนนแบบแยกส่วนและความเชื่อมั่นแยกส่วนของผู้ตรวจ (rater separation reliability) ในการให้คะแนนแบบรวมมีค่าสูงกว่าการให้คะแนนแบบแยกส่วน

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 การนำเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ไปใช้ต้องศึกษารายละเอียดในรายการประเด็นประเมินให้เข้าใจ อาจมีการจัดอบรมให้กับผู้ใช้เกณฑ์การให้คะแนน เพื่อให้ผู้ใช้เกณฑ์เข้าใจรายละเอียดของแต่ละรายการประเมินมากขึ้น เพื่อให้มีความเที่ยงตรงสูง และเพื่อป้องกันการเกิดความคลาดเคลื่อนเนื่องมาจากประเมิน

1.2 เกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ที่พัฒนาขึ้น สามารถนำไปประยุกต์ใช้กับรายวิชาอื่นได้

1.3 การนำเกณฑ์การให้คะแนนโครงงานวิทยาศาสตร์ไปใช้ ควรเลือกใช้รูปแบบการให้คะแนนให้เหมาะสมกับสิ่งต้องการวัด หากต้องการประเมินเพื่อแสดงถึงจุดเด่นจุดด้อยในแต่ละองค์ประกอบ รวมถึงการให้ข้อมูลย้อนกลับเพื่อพัฒนาหรือปรับปรุงตนเองก็สามารถใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ หากต้องการประเมินโดยใช้เวลาน้อยและง่ายต่อการนำไปใช้ เน้นภาพรวมของงานหรือสิ่งที่ต้องการวัดก็สามารถใช้เกณฑ์แบบภาพรวมได้

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

ในการทำวิจัยครั้งต่อไป ผู้วิจัยควรมีการจัดกลุ่มผู้ประเมินและผู้ถูกประเมิน ซึ่งหากมีการจำแนกกลุ่มผู้ประเมิน เช่น ครูที่ปรึกษาโครงงาน ครูที่เป็นกรรมการในการตัดสินโครงงานวิทยาศาสตร์ หรือครูผู้สอนวิทยาศาสตร์ จัดกลุ่มผู้ถูกประเมินเป็น กลุ่มเก่ง กลุ่มปานกลาง และกลุ่มอ่อน แล้วนำผลการประเมินที่ได้มาช่วยการหาความเชื่อมั่นของเกณฑ์การให้คะแนน เพราะหากผู้ประเมินทุกกลุ่มสามารถประเมินได้สอดคล้องกับความสามารถของผู้ถูกประเมิน จะทำให้เกณฑ์การให้คะแนนมีความเชื่อมั่นและความน่าเชื่อถือมากที่สุด

เอกสารอ้างอิง

กนิษฐา สวยสด. (2554). *การเปรียบเทียบความเที่ยงและความเป็นปรนัยของการตรวจให้คะแนนผลงาน โดยใช้กฎเกณฑ์การให้คะแนนและไม่ใช้กฎเกณฑ์การให้คะแนน.* ปริญญาครุศาสตรมหาบัณฑิต สาขาวิชาการวัดประเมิน และวิจัยทางการศึกษา, มหาวิทยาลัยราชภัฏลำปาง.

จันทร์ฉาย สุวลัย. (2552). *การเปรียบเทียบความเที่ยงของการประเมินผลงานโดยใช้กฎเกณฑ์การให้คะแนน ที่นักเรียนร่วมสร้างกับกฎเกณฑ์การให้คะแนนที่ครูสร้าง.* ปริญญาครุศาสตรมหาบัณฑิต สาขาวิชาการวัดประเมิน และวิจัยทางการศึกษา, มหาวิทยาลัยราชภัฏลำปาง.

ไซลัน สาและ. (ม.ป.ป.). *เกณฑ์การให้คะแนน.* ค้นเมื่อ 19 สิงหาคม 2560, สืบค้นจาก http://www.ded.edu.kps.ku.ac.th

น้ำผึ้ง อินทะเนตร. (2554). *การศึกษาคุณลักษณะของคะแนนแบบทดสอบปลายเปิดวิชาคณิตศาสตร์เมื่อจำนวนผู้ตรวจและรูปแบบการตรวจให้คะแนนต่างกัน โดยใช้โมเดลการสรุปอ้างอิงและโมเดลหลายองค์ประกอบของราส์ช.* ปริญญาการศึกษาดุษฎีบัณฑิต สาขาวิชาการทดสอบและวัดผลการศึกษา บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.

บุษวรรษ์ แสนปลื้ม. (2556). *การใช้วิธีการตรวจคุณลักษณะ และสัดส่วนจำนวนผู้ตรวจให้คะแนนที่มีผลต่อความเที่ยงตรงของการวัดความสามารถในการเขียนของนักเรียนชั้นประถมศึกษาปีที่ 3.* ปริญญาการศึกษาดุษฎีบัณฑิต สาขาวิชาการทดสอบและวัดผลการศึกษา บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.

ประทีป กาลเขว้า. (2559). *การพัฒนารูปแบบการประเมินทักษะการปฏิบัติทางวิชาชีพ ด้านการจัดการอนามัยสิ่งแวดล้อม : การตรวจสอบคุณภาพเกณฑ์การให้คะแนนโดยใช้โมเดลหลายองค์ประกอบของราส์ช และประยุกต์ใช้กรอบแนวคิดการประเมินศูนย์ควบคุมและป้องกันโรคแห่งสหรัฐอเมริกา.* วิทยานิพนธ์ปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาการวัดและประเมินผลทางการศึกษา บัณฑิตวิทยาลัย มหาวิทยาลัยขอนแก่น.

ศุภกิจ ศรีจองแสง. (2553). *การสร้างและพัฒนาเกณฑ์ประเมินการอ่าน คิดวิเคราะห์และเขียนของนักเรียนช่วงชั้นที่ 4 สังกัดสำนักงานเขตพื้นที่การศึกษาเชียงราย เขต 2.* วิทยานิพนธ์ครุศาสตรมหาบัณฑิต สาขาการวิจัยและประเมินผลการศึกษา มหาวิทยาลัยราชภัฏเชียงราย.

ศูนย์วิทยาศาสตร์เพื่อการศึกษา. (2559). *เกณฑ์การประกวดโครงงานวิทยาศาสตร์ ประจำปีงบประมาณ 2559*. กรุงเทพฯ: ศูนย์วิทยาศาสตร์เพื่อการศึกษา.

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2549) *คู่มือวัดผลประเมินผลวิทยาศาสตร์ สสวท.* . กรุงเทพฯ: โรงพิมพ์คุรุสภาลาดพร้าว.

สุวิมล ติรกานันท์. (2551). *สถิติและการวิจัยเบื้องต้นทางการศึกษา.* กรุงเทพฯ : สำนักพิมพ์มหาวิทยาลัยรามคำแหง.

Chi, Eunlim. (2001)**.** *Comparison Holistic and Analytic Scoring foe Performance Assessment with Many – Facet Rasch Model.* Journal of Applied Measurement. 2(4):379-388.

Linacre, John M.; et al. (1994). *Measurement with Judges : Many – Faceted Conjoint Measurement.* International Journal Research. *21*(6). 569-577.

Stenmark, Jean K. (1991). *Mathematics Assessment : Myths, Models, Good Questions, and Practical Suggestions.* 6th ed. Reston, Virginia : National Council of Teachers of Mathematics.