

การพัฒนาผังงานวิธีการให้คะแนนสำหรับการประเมินผลการเรียนรู้ของนักเรียน

วิจิตร ชะโลปถัมภ์

วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา ชลบุรี 20131

พุลพงษ์ สุขสว่าง

วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา ชลบุรี 20131

ติดต่อผู้เขียนบทความที่ วิจิตร ชะโลปถัมภ์ วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา ชลบุรี 20131

E-mail: wijitlam@gmail.com

วันที่รับบทความ: 9 เมษายน 2562 วันที่แก้ไขบทความ: 21 พฤษภาคม 2563 วันที่ตอบรับบทความ: 30 พฤษภาคม 2563

บทคัดย่อ

วัตถุประสงค์ เพื่อพัฒนาวิธีการให้คะแนนผังงานสำหรับการประเมินผลการเรียนรู้ของนักเรียน และ ศึกษาความน่าเชื่อถือภายในผู้ประเมินผลการให้คะแนนระหว่างวิธีการให้คะแนนแบบทั่วไปกับวิธีที่พัฒนาขึ้น **วิธีการวิจัย** ระยะที่ 1 การพัฒนาผังงานวิธีการให้คะแนน 3 แบบ จำนวนแบบละ 10 ชิ้น สำหรับการประเมินผลการเรียนรู้วิชาคณิตศาสตร์ของนักเรียนระดับประถมศึกษาปีที่ 6 ประเมินความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ 5 คน อยู่ในระดับน่าพอใจ ระยะที่ 2 การหาความน่าเชื่อถือภายในระหว่างผู้ให้คะแนนและระหว่างวิธีการให้คะแนนแบบทั่วไปกับแบบผังงานที่พัฒนาขึ้นโดยกลุ่มตัวอย่างนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 3 กลุ่ม (37, 37 และ 49 คน) การวิเคราะห์ข้อมูลโดยวิธีวิเคราะห์ความน่าเชื่อถือภายในระหว่างผู้ประเมิน และระหว่างวิธีการให้คะแนน **ผลการวิจัย** ได้พัฒนาผังงานวิธีการให้คะแนนสำหรับการประเมินผลการเรียนรู้ของนักเรียน จำนวน 3 รูปแบบ มีค่าความน่าเชื่อถือภายในผู้ประเมินระหว่างวิธีทั่วไปกับวิธีที่พัฒนาขึ้นอยู่ในระดับสูง ($\alpha = 0.93$) ค่าความน่าเชื่อถือระหว่างผู้ตรวจให้คะแนน มีค่าอยู่ในระดับสูง ($\alpha = 0.96$) ด้วย **สรุป** วิธีการให้คะแนนตามผังงานที่พัฒนาขึ้น มีศักยภาพในการนำไปประยุกต์ใช้ในการประเมินผลการเรียนรู้ทางด้านคณิตศาสตร์และวิชาอื่นๆ ของนักเรียนระดับประถมศึกษา

คำสำคัญ: ผังงานวิธีการให้คะแนน การให้คะแนนแบบแยกส่วน การให้คะแนนแบบทั่วไป



Development of Scoring Procedure Flowcharts for the Assessment of Students' Learning Outcome

Wijit Chalopatham

College of Measurement and Technology in Cognitive Science, Burapha University, Chonburi 20131

Poonpong Suksawang

College of Measurement and Technology in Cognitive Science, Burapha University, Chonburi 20131

Correspondence concerning this article should be addressed to **Wijit Chalopatham** College of Measurement and Technology in Cognitive Science, Burapha University, Chonburi 20131

E-mail: wijitlam@gmail.com

Received date: April 9, 2019 Revised date: May 21, 2020 Accepted date: May 30, 2020

ABSTRACT

PURPOSES: The objectives of this study were to develop a flowcharts-based scoring method for learning evaluation, and to assess the inter-rater reliability of flowcharts scoring of selected general scoring rubrics. **METHODS:** Step 1: development of 3 types of flowchart scoring, 10 pieces for each type, for the evaluation of learning in mathematics of students at Phatomsuksa level 6. The content validity of the flowcharts as assessed by 5 experts was satisfactory. Step 2: The verification of internal reliability between the developed flowchart scoring system and general scoring system as well as among the raters for 3 sample groups of Phatomsuksa level 6 (37, 37 and 49 persons). Main analytical technique was inter-rater reliability. **RESULTS:** The instruments used as scoring criteria were 3 types of scoring flowcharts. The reliability between the developed flowchart scoring and general method was high ($\alpha = 0.93$), and the reliability among the raters were also high ($\alpha = 0.96$). **CONCLUSIONS:** The developed scoring procedure flowcharts are potentially applicable for evaluation of mathematics learning and other subjects at Phatomsuksa level.

Keywords: Scoring procedure flowchart, analytic scoring rubrics, general scoring rubrics

บทนำ

รายงานสภาวะการศึกษาไทย ปี พ.ศ. 2558-2559 ทำให้ทุกคนตระหนักดีว่า การศึกษามีหน้าที่สำคัญในการสร้างทุนมนุษย์ โดยช่วยพัฒนาให้เป็นคนที่มีประสิทธิภาพและมีคุณค่าต่อสังคม ดังจะเห็นได้จากกรอบทิศทางแผนการศึกษาแห่งชาติ พ.ศ. 2560-2574 ที่เน้นย้ำถึงความสำคัญของการพัฒนาทักษะและสมรรถนะของนักเรียน (สำนักงานเลขาธิการสภาการศึกษากระทรวงศึกษาธิการ , 2560) หลักของการวัดและการประเมินผล การเรียนรู้ของนักเรียน ประการแรก คือ การประเมินผลเพื่อการพัฒนา (Formative assessment) การประเมินผลระหว่างการเรียนการสอนอย่างต่อเนื่องที่เกิดขึ้นทุกวันเป็นการประเมินเพื่อให้รู้จุดเด่นจุดด้อย และการประเมินเพื่อตัดสินผลการเรียนรู้ (Summative assessment) เป็นการประเมินเพื่อสรุปผลการเรียนรู้ซึ่งมีหลายระดับ ในการประเมินเพื่อตัดสินผลการเรียนที่ดี ต้องให้โอกาสนักเรียนแสดงความรู้ความสามารถด้วยวิธีที่หลากหลายและตัดสินผลบนพื้นฐานของเกณฑ์การปฏิบัติมากกว่าใช้เปรียบเทียบระหว่างนักเรียน การประเมินแบบอัตนัย (Subjective scoring) แนวทางการประเมินแบบอัตนัย โดยทั่วไปมุ่งวัดระดับการเรียนรู้ชั้นสูงตั้งแต่ขั้นการนำไปใช้ขึ้นไป การสร้างแบบประเมินจะเริ่มจากการกำหนดประเด็นปัญหาที่ต้องการจะวัดซึ่งจะต้องตรงกับวัตถุประสงค์การเรียนรู้ ต่อจากนั้นจึงเขียนเกณฑ์การประเมินให้เหมาะสมกับการวัด

ระดับการเรียนรู้ของนักเรียน มีการออกแบบขั้นตอนของการกำหนดเกณฑ์การให้คะแนนไว้ โดยที่ Nitko (2001) ได้เสนอเกณฑ์การให้คะแนนแบบแยกส่วน (Analytic scoring rubrics) ได้สรุปไว้เป็นแบบ 7 ขั้นตอน เป็นเกณฑ์การให้คะแนนหลายองค์ประกอบ แต่ละระดับการให้คะแนนจะต้องมีความเหมาะสมและมีความชัดเจนในการนิยามที่ครอบคลุมเนื้อหา

ผังงาน (Flowchart) เป็นกระบวนการคิดและนำเสนอวิธีการอธิบายออกมาเป็นการวางแผนอย่างเป็นขั้นตอน ในการศึกษาที่มีการนำอัลกอริทึมมาใช้ร่วมกับผังงาน (Flowchart) ถ่ายทอดแนวทางการแก้ปัญหาผ่าน ผังงานในรูปของสัญลักษณ์ที่มีทิศทางจนได้ผลลัพธ์ตามต้องการที่แสดงออกมาเพื่อให้เข้าใจได้ง่าย ประกอบด้วยโครงสร้างสามรูปแบบ คือ 1) โครงสร้างแบบลำดับ (Sequential structure) 2) โครงสร้างแบบทางเลือก (Selection structure) มีการเลือกทำตามเงื่อนไขที่กำหนด และ 3) โครงสร้างแบบวนซ้ำ (Repetition structure) โดยมีเงื่อนไขในการควบคุม อัลกอริทึมส่วนใหญ่ประกอบด้วยอัลกอริทึมทั้งสามรูปแบบนี้ภายในผังงานเดียวกันและวางแผนให้อยู่ในรูปแบบที่มีทิศทางการทำงานอย่างง่าย

จากการศึกษาเกณฑ์การให้คะแนน Rezaei and Lovorn (2010) ได้เสนอการประเมินงานเขียนของนักเรียนโดยการหาความเที่ยงและความตรงของเกณฑ์การให้คะแนน เพื่อดูว่าการใช้เกณฑ์การให้คะแนน

(Rubrics) นำไปประเมินผลจะมีความเที่ยงมากขึ้นหรือไม่ เมื่อเทียบกับการประเมินที่ไม่ใช้เกณฑ์การให้คะแนน ผลการทดลองยังบ่งชี้ว่า การใช้เกณฑ์การให้คะแนนไม่สามารถพัฒนาความเที่ยงและความตรงของการประเมินถึงการออกแบบและการใช้เกณฑ์การให้คะแนนมีประสิทธิภาพ ถ้าผู้ตรวจให้คะแนน (Rater) ไม่ได้รับการฝึกมาเป็นอย่างดี และ Panadero et al. (2012) ได้เปรียบเทียบผลของการใช้เครื่องมือการประเมินตนเองที่แตกต่างกันที่มีต่อการควบคุมตนเอง การเรียน และการรับรู้ความสามารถของตนเอง สำหรับงานวิจัยด้านเกณฑ์ Wollenschläger et al. (2016) ได้นำเสนอว่า อะไรที่ทำให้เกณฑ์การให้คะแนนมีประสิทธิภาพในการสะท้อนผลการตรวจให้คะแนนผลงานของนักเรียนและครู และต่อมา Fraile et al. (2017) ได้นำเสนอการสร้างเกณฑ์การให้คะแนน เพื่อศึกษาผลกระทบต่อการเรียนรู้การควบคุมตนเอง ความสามารถในการตนเองและประสิทธิภาพในการกำหนดเกณฑ์การประเมินกับนักเรียน จะเห็นว่าแนวทางจากเอกสารงานวิจัยเกณฑ์การให้คะแนนและการตรวจให้คะแนนมีผลต่อการสะท้อนถึงจุดเด่นและจุดด้อยของการเรียนรู้ของนักเรียนโดยตรง

จากการศึกษาเอกสารงานวิจัยที่เกี่ยวข้องกับการหาความสอดคล้องภายในของผู้ตรวจ (Rater Agreement Index: RAI) ปรากฏว่าปัจจุบันนักวิชาการให้ความสนใจในการนำวิธีการหาความน่าเชื่อถือโดยผู้วัดต่าง

คนกัน (Inter-Rater Reliability: IRR) มาใช้กับวิธีการตรวจให้คะแนนการประเมินผลอย่างกว้างขวางมากขึ้น โดยที่ Hallgren (2012) ได้ศึกษาวิจัยการหาความน่าเชื่อถือโดยผู้วัดต่างคนกันสำหรับประเมินผลการเก็บข้อมูลแบบอัตนัย และ De Groef et al. (2017) ได้ใช้วิธีการประเมินความน่าเชื่อถือโดยผู้วัดต่างคนกัน (IRR) สำหรับการตรวจวัดไหล่ทางร่างกายในหญิงวัยกลางคน ด้านอายุรเวชศาสตร์ทางการแพทย์ และ Lucas et al. (2017) ได้ศึกษาวิจัยความน่าเชื่อถือโดยผู้วัดต่างคนกัน (IRR) ประเมินความคิดไตร่ตรองของนักศึกษาเภสัชศาสตร์ เพื่อการหาความสอดคล้องภายในของผู้ตรวจและการหาความน่าเชื่อถือโดยผู้วัดต่างคนกัน

จากเหตุผลดังกล่าวผู้วิจัยได้ศึกษาการพัฒนาวิธีการให้คะแนนผังงานสำหรับการประเมินผลการเรียนรู้ของนักเรียน ซึ่งมีแนวการพัฒนาจากวิธีการให้คะแนนจากโครงสร้างเกณฑ์การให้คะแนนแบบแยกส่วน (Analytic scoring rubrics) ของ Nitko (2001) ที่ออกแบบไว้แบบ 7 ขั้นตอน โดยผู้วิจัยได้พัฒนาวิธีการให้คะแนนผังงาน 3 รูปแบบ ได้แก่ 1) ผังงานการเรียงลำดับ 2) ผังงานการทำซ้ำ และ 3) ผังงานการตรวจสอบแบบเงื่อนไขหรือตรรกศาสตร์ เพื่อพัฒนาเกณฑ์การให้คะแนนผังงาน ที่จะนำมาใช้เป็นแนวทางสำหรับประเมินผลการเรียนรู้ของนักเรียนระดับประถมศึกษาชั้นปีที่ 6

วัตถุประสงค์

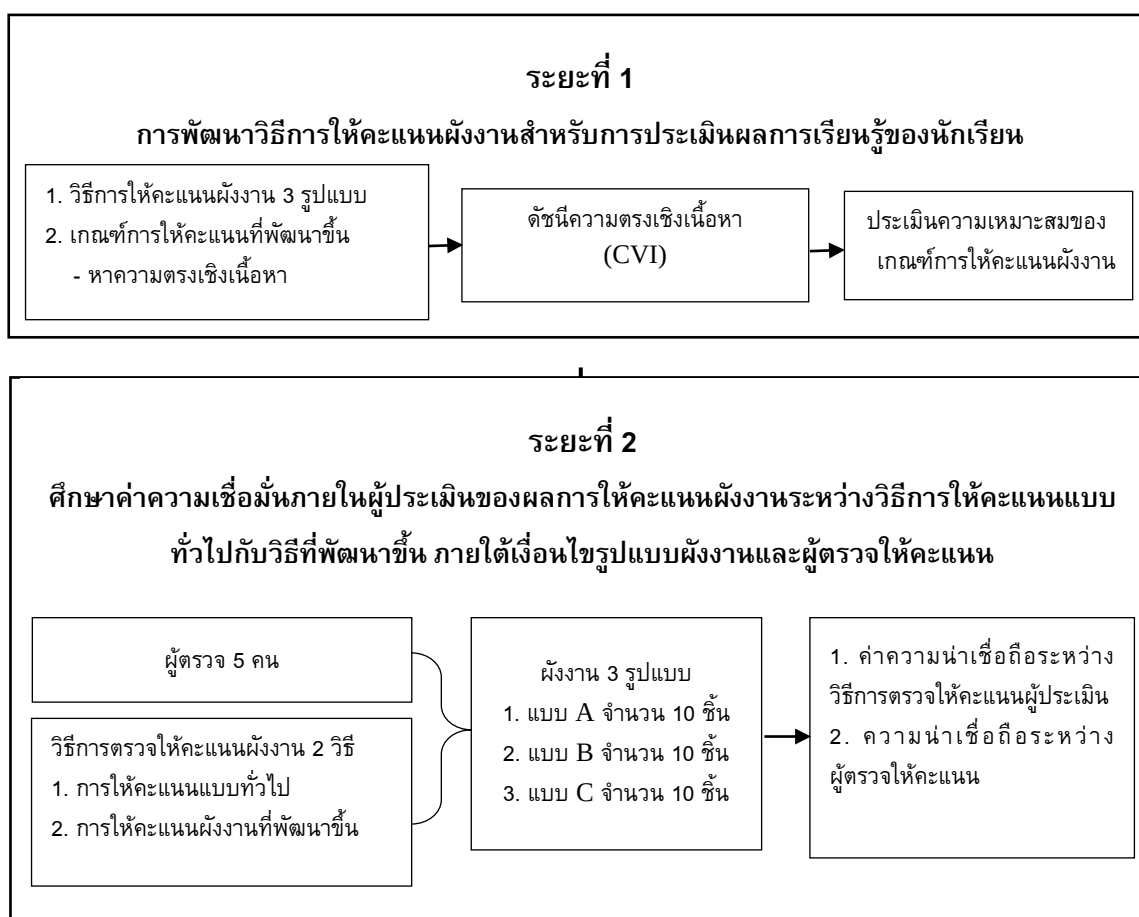
1. เพื่อพัฒนาวิธีการให้คะแนนผังงานสำหรับการประเมินผลการเรียนรู้ของนักเรียน
2. เพื่อศึกษาคำความน่าเชื่อถือภายในผู้ประเมินของผลการให้คะแนนผังงานระหว่างวิธีการให้คะแนนแบบทั่วไปกับวิธีที่พัฒนาขึ้น ภายใต้เงื่อนไขรูปแบบผังงานและผู้ตรวจให้คะแนน

สมมติฐาน

ผลการให้คะแนนผังงานระหว่างวิธีการให้คะแนนแบบทั่วไปกับวิธีที่พัฒนาขึ้น ภายใต้เงื่อนไขรูปแบบผังงานและผู้ตรวจให้คะแนน มีความน่าเชื่อถือภายในผู้ประเมิน

กระบวนการวิจัย

กำหนดการรายละเอียดของการดำเนินงานตามระยะ แสดงดังภาพที่ 1



คำจำกัดความเชิงปฏิบัติการ

วิธีการให้คะแนนที่พัฒนาขึ้น หมายถึง แนวทางการให้คะแนนผังงานแบบแยกส่วน สำหรับประเมินผลการเรียนรู้ของนักเรียน มีส่วนประกอบ คือ แบบประเมินผังงาน แนวทางการตอบผังงานและเกณฑ์การให้คะแนน โดยใช้ร่วมกับแนวทางการตอบแบบประเมินผังงาน กำหนดเกณฑ์การให้คะแนนแบบ 10 ระดับคะแนน คือ 1, 2, 3, 4, 5, 6, 7, 8, 9 และ 10 คะแนน ที่ประยุกต์มาจากเกณฑ์การให้คะแนนแบบ Task-Specific Rubrics ของ Nitko (2011) ที่อธิบายถึงคุณภาพของผลการปฏิบัติงานที่เฉพาะเจาะจงของการแก้ปัญหาคณิตศาสตร์

วิธีการให้คะแนนแบบทั่วไป หมายถึง แนวทางการให้คะแนนคณิตศาสตร์แบบทั่วไป จะใช้เกณฑ์การให้คะแนนทักษะหรือกระบวนการทางคณิตศาสตร์ 4 ด้าน ได้แก่ 1) การแก้ปัญหา 2) การให้เหตุผล 3) การสื่อความหมายทางคณิตศาสตร์ และ 4) การเชื่อมโยง กำหนดเกณฑ์การให้คะแนนแบบ 5 ระดับคะแนน คือ 0, 1, 2, 3, และ 4 คะแนน ตามระดับความถูกต้องจากน้อยไปหามาก

ผังงาน (Flowchart) หมายถึง แผนภาพที่แสดงถึงขั้นตอนวิธีการขั้นตอนการทำงานหรือกระบวนการขั้นตอนการแสดงสัญลักษณ์ชนิดต่างๆและทิศทางการเชื่อมต่อด้วยลูกศรแสดงให้เห็นถึงการแสดงแผนภาพรูปแบบวิธีการแก้ปัญหาที่กำหนด มี 3 รูปแบบ ได้แก่ 1) ผังงานรูปแบบ A (การเรียงลำดับ) 2) ผังงานรูปแบบ B (การทำซ้ำ)

และ 3) ผังงานรูปแบบ C (การตรวจสอบเงื่อนไข)

ผู้ตรวจ (Raters) หมายถึง ผู้ตรวจให้คะแนนผังงานสำหรับประเมินผลการเรียนรู้ของการเป็นครูสอนรายวิชาคณิตศาสตร์ ซึ่งสำเร็จการศึกษาไม่ต่ำกว่าปริญญาโท ที่มีประสบการณ์ทางด้านการสอนในระดับประถมศึกษาอย่างน้อย 10 ปี

กลุ่มผู้ตรวจให้คะแนน (Raters group) หมายถึง ผู้ตรวจให้คะแนนผังงานสำหรับประเมินผลการเรียนรู้ของนักเรียนเป็นครูสอนรายวิชาคณิตศาสตร์ จำแนกเป็น 2 กลุ่ม ได้แก่ กลุ่มผู้ตรวจที่มีประสบการณ์การสอนต่ำกว่า 10 ปี และ กลุ่มผู้ตรวจที่มีประสบการณ์การสอน 10 ปี ขึ้นไป

ความสอดคล้องของผู้ตรวจ (Inter-Raters) หมายถึง ระดับความสอดคล้องของการให้คะแนนคำตอบของผู้ตรวจให้คะแนนผังงาน ในแต่ละวิธีการตรวจของผู้ตรวจแต่ละคน ซึ่งตรวจสอบค่าดัชนีความสอดคล้องภายในผู้ประเมิน ซึ่งค่าแปรเปลี่ยนตั้งแต่ 0.0-1.0

ค่าความน่าเชื่อถือภายในผู้ตรวจ (Inter-Raters Reliability) หมายถึง การประเมินค่าความน่าเชื่อถือของระดับความสอดคล้องการให้คะแนนของผู้ตรวจให้คะแนนผังงานต่างคนกันในการตรวจให้คะแนนผังงานสำหรับการประเมินผลการเรียนรู้

วิธีดำเนินการวิจัย

การวิจัยนี้มีวิธีการดำเนินการวิจัยเป็น 2 ระยะ ดังนี้

ระยะที่ 1 การพัฒนาวิธีการให้คะแนนผังงาน สำหรับการประเมินผลการเรียนรู้ของนักเรียน ดังต่อไปนี้

การพัฒนาวิธีการให้คะแนนผังงาน สำหรับการประเมินผลการเรียนรู้ของนักเรียน มีรายละเอียด ดังนี้

1. ศึกษาเอกสารที่เกี่ยวข้อง และ สร้างกรอบแนวทางแบบพหุมิติ โดยกำหนด มิติที่ต้องการวัดตามเนื้อหาต้องการได้แบบ 2 มิติ คือ 1) มิติกระบวนการพุทธิปัญญา และ 2) มิติความรู้การประยุกต์วิธีดำเนินการ และ หลักการใช้วิธีการของผังงานกับทักษะ กระบวนการแก้ปัญหา กำหนดขอบเขตเนื้อหา ตามมาตรฐานการเรียนรู้และสาระการเรียนรู้ วิชาคณิตศาสตร์ และมีวิธีการให้คะแนนตาม แนวคิดเชิงทฤษฎี วิธีการให้คะแนนแบบแยก ส่วน ของ Nitko (2001) ลำดับ 7 ขั้นตอน มา พัฒนาวิธีการตรวจให้คะแนนผังงานขึ้น ครั้งนี้ มุ่งการประเมินแบบอัตโนมัติที่แสดงทักษะหรือ กระบวนการทำงานของนักเรียน

2. สังเคราะห์แนวโน้มที่เป็นไปได้ สำหรับการให้คะแนนผังงานสำหรับ ประเมินผลการเรียนรู้ที่พัฒนาขึ้น โดยการ สังเคราะห์ขึ้นจากการทบทวนวรรณกรรมจาก องค์ประกอบผังงานจะมีลำดับการใช้ สัญลักษณ์ผังงาน 5 ขั้นตอน (ประกอบผล, 2558) กับเกณฑ์การให้คะแนนของ Nitko (2001) สำหรับการแก้ปัญหาคณิตศาสตร์ 3 ด้าน ได้แก่ ด้านความรู้คณิตศาสตร์ ด้าน ความรู้เชิงกลยุทธ์ และด้านการสื่อสาร ได้ รูปแบบการเขียนผังงาน จำนวน 3 รูปแบบ จะ

มีกระบวนการแก้ปัญหาแสดงวิธีการของผัง งานใช้คุณลักษณะตามรูปสัญลักษณ์ที่และการ แก้ปัญหาคณิตศาสตร์

3. การพัฒนาแบบประเมินผังงาน และแนวทางการตอบ ผู้วิจัยใช้ลักษณะการใช้ ผังงานกับแนวทาง การพิจารณารายวิชา คณิตศาสตร์พื้นฐาน ในการประเมินผลการ เรียนรู้ที่จะสะท้อนความรู้ความเข้าใจในแต่ละ ส่วนการเรียนรู้ โดยใช้ตามแบบฝึกหัดรายวิชา คณิตศาสตร์พื้นฐาน ชั้นประถมศึกษาปีที่ 6 พ.ศ. 2551 กระบวนการประเมินผลการเรียนรู้ ของนักเรียน เกณฑ์ให้คะแนนผังงานที่ พัฒนาขึ้น มีที่เหมาะสมกับการประเมินแบบ อัตโนมัติ แสดงการแก้ปัญหาผ่านการใช้รูป สัญลักษณ์ ตัวแปร และ Pseudo code และจะ ไม่มีการประเมินด้านความคิดริเริ่มสร้างสรรค์

4. การตรวจสอบคุณภาพของแบบ ประเมินและเกณฑ์การให้คะแนนผังงาน สำหรับการประเมินผลการเรียนรู้ของนักเรียน ผู้เชี่ยวชาญตรวจสอบคุณภาพของแบบ ประเมินและเกณฑ์การให้คะแนนผังงานด้าน ความตรงเชิงเนื้อหา (Content validity) โดย ผู้เชี่ยวชาญจำนวน 5 คน เป็นผู้ตรวจสอบ เกณฑ์การให้คะแนนชิ้นงานกับเนื้อหาวิชา คณิตศาสตร์พื้นฐาน ตามเกณฑ์การให้คะแนน ดังนี้ 1) ประเมินผังงานรูปแบบ A (การ เรียงลำดับ) 2) ประเมินผังงานรูปแบบ B (การ ทำซ้ำ) และ 3) ประเมินผังงานรูปแบบ C (การ ตรวจสอบเงื่อนไข)

5. การถ่วงน้ำหนักคะแนน (Weighted scoring) ของเกณฑ์การให้

คะแนนผลงาน ผู้วิจัยนำองค์ประกอบของผลงานทั้ง 6 องค์ประกอบ กำหนดค่าตัวเลขที่จะตัดสิน แต่ละองค์ประกอบ ขึ้นอยู่กับจุดประสงค์การเรียนรู้ของเนื้อหาในแต่ละรูปแบบผลงาน มี 3 รูปแบบ โดยใช้ Weighted scoring model นำผลการตรวจให้คะแนนจากเกณฑ์ที่พัฒนาขึ้น มาบันทึกแล้วนำไปวิเคราะห์หาค่าสถิติต่อไป

6. การประเมินความเหมาะสมของเกณฑ์การให้คะแนนผลงาน โดยผู้วิจัยนำแบบประเมินผลงานให้ผู้เชี่ยวชาญ จำนวน 5 คน ทั้ง 3 รูปแบบ ได้แก่ แบบประเมินผลงานรูปแบบ A (การเรียงลำดับ) แบบประเมินผลงานรูปแบบ B (การทำซ้ำ) และ แบบประเมินผลงานรูปแบบ C (การตรวจสอบเงื่อนไข) มีเกณฑ์การพิจารณาประเมินความเหมาะสมพิจารณาความสอดคล้องระหว่างแบบประเมินผลงานกับแนวทางการตอบ

ระยะที่ 2 การศึกษาคำความน่าเชื่อถือภายในผู้ประเมินของผลการให้คะแนนผลงานระหว่างวิธีการให้คะแนนแบบทั่วไปกับวิธีที่พัฒนาขึ้นระหว่างผู้ตรวจให้คะแนน

มีลำดับขั้นตอนของการดำเนินการ ดังนี้

1. ขั้นตอนการเก็บข้อมูลชิ้นงานและการคัดเลือกชิ้นงาน

ขั้นตอนนี้เป็นการเก็บข้อมูลชิ้นงานคัดเลือกชิ้นงานและจัดกลุ่มชิ้นงานนักเรียนเพื่อเป็นตัวแทนแยกตามระดับกลุ่มคะแนนออกเป็น 3 กลุ่มคะแนน ได้แก่ 1) กลุ่มคะแนน

สูง 2) กลุ่มคะแนนปานกลาง และ 3) กลุ่มคะแนนต่ำ โดยดำเนินการ ดังต่อไปนี้

1.1 ผู้วิจัยทำหนังสือขอความอนุเคราะห์ขอข้อมูลเพื่อการวิจัยจากวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา ถึงผู้อำนวยการโรงเรียนสาธิตมหาวิทยาลัยศรีนครินทรวิโรฒ (ฝ่ายประถม) โดยมีความประสงค์ขอความอนุเคราะห์ ในการเก็บรวบรวมข้อมูลนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2561 โดยได้รับความอนุเคราะห์จากครูสอนคณิตศาสตร์ 2 คน ดำเนินการเก็บรวบรวมข้อมูลระหว่างวันที่ 11 - 31 มกราคม พ.ศ. 2562

1.2 การกำหนดขอบเขตของเนื้อหา โดยขอคำปรึกษาและการไปอธิบายกับครูสอนคณิตศาสตร์ 2 คน เพื่อกำหนดขอบเขตของเนื้อหาวิชาคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 6 ให้ทราบถึงวัตถุประสงค์ของการวิจัยครั้งนี้ เริ่มจากอธิบายวิธีการเขียนผลงาน และวิธีการประเมินให้คะแนนชิ้นงานนักเรียน แล้วลำดับวิธีการเขียนผลงานกับการแก้ปัญหาทางคณิตศาสตร์ กำหนดเป็นโจทย์ปัญหาทางคณิตศาสตร์ที่สอดคล้องกับการเขียนใช้ผลงานได้ 3 แบบ

- 1) ผลงานรูปแบบ A (การเรียงลำดับ)
- 2) ผลงานรูปแบบ B (การทำซ้ำ)
- 3) ผลงานรูปแบบ C (การตรวจสอบเงื่อนไข)

กำหนดเป็นโจทย์คณิตศาสตร์จำนวน 3 ข้อ ใช้เป็นข้อคำถามสำหรับการเก็บชิ้นงานของนักเรียน

1.3 การเก็บชิ้นงานของนักเรียน โดยเริ่มจากให้นักเรียนฝึกเขียนผังงานแสดง การหาพื้นที่รูปสามเหลี่ยมมุมฉากแบบเดิม แล้วจึงอธิบายเกี่ยวกับการใช้วิธีการเขียนผัง งานและการสื่อความหมายของผังงาน เมื่อใช้ องค์ประกอบของผังงาน โดยครูผู้สอนและ ผู้วิจัยช่วยกันอธิบายขั้นตอนจนนักเรียนเขียน ผังงานหาพื้นที่รูปสามเหลี่ยมมุมฉากได้ แล้ว เริ่มการเก็บชิ้นงานจากนักเรียน ด้วยข้อ คำถาม 3 ข้อ จนได้ผังงานที่นักเรียนเขียนขึ้น จากโจทย์ที่กำหนด

1.4 การคัดเลือกชิ้นงานของ นักเรียนจากการเขียนผังงาน 3 รูปแบบ โดย ผู้สอน 2 คน ทำการตรวจให้คะแนนชิ้นงานทุก ชิ้นงาน และเลือกชิ้นงานของนักเรียน ดังนี้

1.4.1 ผังงานรูปแบบ A เป็น ผังงานการแก้โจทย์ปัญหาร้อยละ จำนวน 1 ข้อ จากนักเรียน จำนวน 49 คน ได้ชิ้นงาน 49 ชิ้นงาน โดยตรวจให้คะแนนทุกชิ้นงาน แล้ว เลือกมา 10 ชิ้นงาน

1.4.2 ผังงานรูปแบบ B เป็น ผัง งานการคิดคำนวณผลรวมตั้งแต่ 1-10 จำนวน 1 ข้อ จากนักเรียน จำนวน 37 คน ได้ชิ้นงาน 37 ชิ้นงาน โดยตรวจให้คะแนนทุกชิ้นงาน แล้วเลือกมา 10 ชิ้นงาน

1.4.3 ผังงานรูปแบบ C เป็น ผัง งานการคิดคำนวณสูตรคูณแม่ 5 จำนวน 1 ข้อ จากนักเรียน จำนวน 37 คน ได้ชิ้นงาน 37 ชิ้นงาน โดยตรวจให้คะแนนทุกชิ้นงาน แล้วเลือกมา 10 ชิ้นงาน

นำผลการคัดเลือกชิ้นงานในแต่ละ แบบมาละชิ้นงาน และลบชื่อนักเรียนออก แล้วกำหนดเป็นชุดผังงานตามเนื้อหาแบบละ ระดับคะแนน ได้ผังงาน 3 รูปแบบ จัดเก็บให้ เป็นแฟ้มข้อมูลรูปภาพ เพื่อนำเข้าสู่ขั้นตอน การพัฒนาโปรแกรมแบบออนไลน์ต่อไป

2.กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ครั้งนี้ เป็นครูทำ หน้าที่ผู้ตรวจ (Raters) ให้คะแนนผังงาน สำหรับประเมินผลการเรียนรู้ของนักเรียนที่ สอนรายวิชาคณิตศาสตร์ จำนวน 5 คน จะ เป็นการประเมินให้คะแนนชิ้นงานของนักเรียน ตรวจตามเกณฑ์การให้คะแนนแบบทั่วไปและ ตรวจตามเกณฑ์การให้คะแนนผังงานที่ พัฒนาขึ้น ใช้กลุ่มตัวอย่างแบบเจาะจง (Purposive sampling) ลักษณะของกลุ่มที่ เลือกเป็นไปตามวัตถุประสงค์ของการวิจัย

3. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย ครั้งนี้ ผู้วิจัยพัฒนาขึ้นเป็นแบบประเมินการให้ คะแนนผังงานสำหรับการประเมินผลการ เรียนรู้ของนักเรียน เป็นวิธีการให้คะแนนผัง งานสำหรับการประเมินผลการเรียนรู้ของ นักเรียนที่พัฒนาขึ้น ที่ผ่านขั้นตอนการหา คุณภาพของเครื่องมือการวิจัย ได้เกณฑ์การ ให้คะแนนผังงานสำหรับการประเมินผลการ เรียนรู้ของนักเรียน มี 3 รูปแบบ ได้แก่ 1) เกณฑ์การประเมินผังงานรูปแบบ A 2) เกณฑ์การประเมินผังงานรูปแบบ B และ 3) เกณฑ์การประเมินผังงานรูปแบบ C

4. การเก็บรวบรวมข้อมูล

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูล โดยการนำแบบประเมินผังงานให้ผู้ตรวจจำนวน 5 คน ตรวจให้คะแนนแบบทั่วไปและตรวจให้คะแนนผังงานตามเกณฑ์ที่พัฒนาขึ้น จะตรวจให้คะแนนผังงานตามเกณฑ์ทั้ง 2 วิธีจากผู้ตรวจจำนวน 5 คน โดยกำหนดให้ตรวจให้คะแนนผังงานของนักเรียนทุกชั้นงาน ดังนี้

- 1) ตรวจให้คะแนนผังงานรูปแบบ A จำนวน 10 ชั้นงาน
- 2) ตรวจให้คะแนนผังงานรูปแบบ B จำนวน 10 ชั้นงาน
- และ 3) ตรวจให้คะแนนผังงานรูปแบบ C จำนวน 10 ชั้นงาน

นำผลการตรวจให้คะแนนจากผู้ตรวจแต่ละคนแต่ละชั้นงาน จะได้ผลการตรวจให้คะแนนได้ผลของผลการตรวจให้คะแนน มาบันทึกแล้วนำไปวิเคราะห์ค่าทางสถิติในขั้นตอนต่อไป

5. การวิเคราะห์ข้อมูลและสถิติที่ใช้ในการวิเคราะห์ข้อมูล

5.1 วิเคราะห์ค่าสถิติพื้นฐานของคะแนนที่ได้จากเกณฑ์ที่พัฒนาขึ้น และเกณฑ์การให้คะแนนแบบทั่วไป โดยคำนวณค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ด้วยโปรแกรมวิเคราะห์ทางสถิติ

5.2 การหาความเชื่อมั่น (Reliability) ของคะแนนที่ได้จากการตรวจแบบทดสอบการเขียนผังงานตามเกณฑ์การตรวจให้คะแนนผัง

6. การทดสอบโดยวิธีสัมประสิทธิ์แอลฟา (Alpha coefficient) เพื่อหาค่าความสอดคล้องภายใน ให้ผลลัพธ์เป็นค่าสัมประสิทธิ์แอลฟา (Alpha coefficient) ที่

บ่งชี้ถึงความสัมพันธ์ระหว่างวิธีการให้คะแนนผังงานที่พัฒนาขึ้นกับวิธีการให้คะแนนแบบทั่วไป ว่ามีความสัมพันธ์ดีหรือไม่ดี วิธีนี้จึงเป็นวิธีการหาค่าความสอดคล้องภายในของผู้ประเมินหรือความสอดคล้องภายในของผู้ตรวจให้คะแนน การหาค่าความเชื่อมั่นโดยใช้สูตรของครอนบัคได้ โดยที่คะแนนของแบบทดสอบเป็นคะแนนแบบมาตรฐานหรืออันดับหรืออันดับภาคด้วยสัมประสิทธิ์แอลฟา (Cronbach's alpha coefficient) มีการพิจารณาจากเกณฑ์ค่าสถิติที่ได้ไม่ต่ำกว่า 0.8 จึงยอมรับได้ว่ามีความสอดคล้องภายใน

การทดสอบความสอดคล้องการให้คะแนนผังงานระหว่างวิธีการให้คะแนนแบบทั่วไปกับวิธีที่พัฒนาขึ้น ภายใต้เงื่อนไขรูปแบบผังงานและผู้ตรวจให้คะแนน โดยใช้สูตรสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน และทดสอบสหสัมพันธ์รายคู่ของผู้ประเมินของผลการให้คะแนนผังงาน ภายใต้เงื่อนไขการตรวจให้คะแนนผัง ซึ่งค่าสถิติที่ได้ไม่ต่ำกว่า 0.8 จึงยอมรับได้ว่ามีความสอดคล้องภายใน

ผลการวิจัย

ตอนที่ 1 ผลการพัฒนาวิธีการให้คะแนนผังงานสำหรับการประเมินผลการเรียนรู้ของนักเรียน

การวิจัยในตอนนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีการให้คะแนนผังงานสำหรับการประเมินผลการเรียนรู้ของนักเรียน ผลการสังเคราะห์ประเด็นแนวโน้มความเป็นไปได้

ของวิธีการเขียนผังงานประกอบกับวิธีการให้คะแนนแบบแยกส่วน ของ Nitko (2001) มาประยุกต์เป็นวิธีการให้คะแนนผังงานได้ประเด็นแนวโน้มความเป็นไปได้ของกรอบเกณฑ์การให้คะแนนทั้งหมด จำนวน 60 ประเด็น เพื่อนำไปใช้ประเมินชิ้นงานของนักเรียน ได้ 3 รูปแบบ ได้แก่ 1) เกณฑ์การตรวจให้คะแนนผังงานรูปแบบ A (การเรียงลำดับ) จำนวน 50 ประเด็น 2) เกณฑ์การตรวจให้คะแนนผังงานรูปแบบ B (การทำซ้ำ) จำนวน 60 ประเด็น และ 3) เกณฑ์การตรวจให้คะแนนผังงานรูปแบบ C (การตรวจสอบเงื่อนไข) จำนวน 60 ประเด็น

ผลการวิเคราะห์ความตรงเชิงเนื้อหา (Content validity) โดยกลุ่มผู้เชี่ยวชาญจำนวน 5 คน ให้ความคิดเห็นและปรับปรุงแก้ไขตามคำแนะนำ พร้อมทั้งหาดัชนีความตรงเชิงเนื้อหา (Content Validity Index: CVI) ผลปรากฏว่า ค่าดัชนีความตรงเชิงเนื้อหา จากความคิดเห็นของผู้เชี่ยวชาญ

จากวิธีการตรวจให้คะแนนผังงานสำหรับการประเมินผลการเรียนรู้ของนักเรียน โดยภาพรวมมีดัชนีความตรงเชิงเนื้อหาทั้งฉบับ (CVI) เท่ากับ 1.00 ซึ่งมีความเหมาะสม ดัชนีความตรงตามเนื้อหาของเครื่องมือมากกว่าเกณฑ์ที่ยอมรับได้คือ 0.80 ขึ้นไป ลำดับความถูกต้องตามเกณฑ์ที่กำหนด

ตอนที่ 2 ผลการศึกษาค่าความสอดคล้องภายในผู้ประเมินของผลการให้คะแนนผังงานระหว่างวิธีการให้คะแนนแบบทั่วไปกับวิธีที่พัฒนาขึ้น ภายใต้เงื่อนไขรูปแบบผังงานและระหว่างผู้ตรวจให้คะแนน

1. ผลการเปรียบเทียบค่าความสอดคล้องระหว่างผู้ประเมิน ภายใต้เงื่อนไขการตรวจให้คะแนนผังงาน 2 วิธี คือ 1) เกณฑ์การให้คะแนนผังงานแบบทั่วไป และ 2) เกณฑ์การให้คะแนนผังงานแบบที่พัฒนา จำนวน 30 ชิ้นงาน และผู้ตรวจ จำนวน 5 คน

ตารางที่ 1 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผู้ประเมินของผลการให้คะแนนผังงาน และค่าสัมประสิทธิ์แอลฟา ภายใต้เงื่อนไขการตรวจให้คะแนนจากเกณฑ์การตรวจให้คะแนนแบบทั่วไป

ผู้ตรวจ	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5
คนที่ 1	1.00	-	-	-	-
คนที่ 2	.90	1.00	-	-	-
คนที่ 3	.72	.73	1.00	-	-
คนที่ 4	.64	.76	.78	1.00	-
คนที่ 5	.70	.80	.85	.84	1.00
ค่าสัมประสิทธิ์แอลฟา	.94				

จากตารางที่ 1 ความสอดคล้องภายในผู้ประเมินของวิธีการให้คะแนนผังงานแบบทั่วไป มีค่าความสัมพันธ์ระหว่างผู้ตรวจ

ให้คะแนน อยู่ระหว่าง .64 ถึง .90 เมื่อทดสอบความสอดคล้องภายในแบบสัมประสิทธิ์แอลฟามีค่าเท่ากับ .94

ตารางที่ 2 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผู้ประเมินของผลการให้คะแนนผังงาน และค่าสัมประสิทธิ์แอลฟาภายใต้เงื่อนไขการตรวจให้คะแนนจากวิธีการให้คะแนนผังงานแบบที่พัฒนา

ผู้ตรวจ	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5
คนที่ 1	1.00	-	-	-	-
คนที่ 2	.93	1.00	-	-	-
คนที่ 3	.92	.93	1.00	-	-
คนที่ 4	.67	.71	.70	1.00	-
คนที่ 5	.91	.91	.98	.77	1.00
ค่าสัมประสิทธิ์แอลฟา	.96				

จากตารางที่ 2 ค่าความสอดคล้องภายในผู้ประเมินของวิธีการให้คะแนนผังงานที่พัฒนาขึ้น มีค่าสหสัมพันธ์ระหว่างผู้ตรวจให้

คะแนน อยู่ระหว่าง .67 ถึง .98 เมื่อทดสอบความสอดคล้องภายในแบบสัมประสิทธิ์แอลฟามีค่า เท่ากับ .96

ตารางที่ 3 ค่าอำนาจจำแนกของผู้ตรวจให้คะแนน จากวิธีการให้คะแนนผังงาน 2 วิธี

ผู้ตรวจ	ค่าอำนาจจำแนก	
	แบบทั่วไป	แบบที่พัฒนา
คนที่ 1	.81	.92
คนที่ 2	.90	.94
คนที่ 3	.83	.96
คนที่ 4	.83	.73
คนที่ 5	.88	.96

จากตารางที่ 3 วิธีการให้คะแนนผังงานแบบทั่วไป มีค่าอำนาจจำแนกของผู้ตรวจอยู่ระหว่าง .81 ถึง .90 และวิธีการให้คะแนนผังงานแบบที่พัฒนา มีค่าอำนาจจำแนกของ

ผู้ตรวจอยู่ระหว่าง .73 ถึง .96 เมื่อพิจารณาระดับของค่าอำนาจจำแนกมากกว่า .20 แสดงว่า การตรวจให้คะแนนของผู้ตรวจให้คะแนนทุกคนมีความสอดคล้องในทิศทางเดียวกัน

2. ผลการเปรียบเทียบการให้คะแนนผังงานระหว่างวิธีการให้คะแนนแบบ

ทั่วไปกับวิธีที่พัฒนาขึ้น ภายใต้เงื่อนไขรูปแบบผังงานและผู้ตรวจให้คะแนน

ตารางที่ 4 ค่าสัมประสิทธิ์สหสัมพันธ์ผู้ประเมินของผลการให้คะแนนผังงาน ระหว่างวิธีแบบทั่วไปกับแบบที่พัฒนาขึ้น

วิธีการให้คะแนน	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5
แบบทั่วไป vs แบบที่พัฒนาขึ้น	.87	.93	.74	.77	.82
ค่าสัมประสิทธิ์แอลฟา	0.93				

จากตารางที่ 4 ผลการทดสอบความสอดคล้องการให้คะแนนผังงานระหว่างวิธีการให้คะแนนแบบทั่วไปกับวิธีที่พัฒนาขึ้น ภายใต้เงื่อนไขรูปแบบผังงานโดยผู้ตรวจให้คะแนน 5 คน ตรวจชิ้นงาน จำนวน 30 ชิ้นงาน ทดสอบความสัมพันธ์รายคู่ของผู้ตรวจ มีค่าสัมประสิทธิ์สหสัมพันธ์ อยู่ระหว่าง .74 ถึง .93 แสดงว่า วิธีการให้คะแนน 2 วิธี มีแนวโน้มการตรวจให้คะแนนสอดคล้องกัน เมื่อทดสอบความสอดคล้องภายในผู้ตรวจ ทั้ง 5 คน โดยใช้สูตรของครอนบัต ได้ค่าสัมประสิทธิ์สหสัมพันธ์ (α) เท่ากับ 0.93 แสดงว่าวิธีการให้คะแนนผังงาน 2 วิธี มีความน่าเชื่อถือภายในผู้ประเมิน ดังนั้นจึงเป็นไปตามสมมติฐาน

สรุปและอภิปรายผล

1. การพัฒนาเกณฑ์การให้คะแนนคะแนนผังงานสำหรับการประเมินผลการเรียนรู้ของนักเรียน ปรากฏผลดังนี้

1) เกณฑ์การให้คะแนนผังงานสำหรับการประเมินผลการเรียนรู้ของนักเรียน

มี 3 รูปแบบ คือ 1) ผังงานรูปแบบ A (การเรียงลำดับ) 2) ผังงานรูปแบบ B (การทำซ้ำ) และ 3) ผังงานรูปแบบ C (การตรวจสอบเงื่อนไข)

2) การตรวจให้คะแนนผังงานสำหรับการประเมินผลการเรียนรู้ของนักเรียน กำหนดให้ตรวจความถูกต้องของคำตอบในแต่ละผังงานและกำหนดเกณฑ์การให้คะแนนผังงานทั้ง 3 รูปแบบ โดยการให้คะแนนตามสัญลักษณ์ของการใช้งาน คือ 1) การเริ่มต้นผังงาน ให้คะแนนใน 5 ประเด็น 2) การกำหนดค่าเริ่มต้นของตัวแปร ให้คะแนนใน 10 ประเด็น 3) การประมวลผล 3.1) ผังงานรูปแบบ A (การเรียงลำดับ) ให้คะแนนใน 10 ประเด็น 3.2) ผังงานรูปแบบ B (การทำซ้ำ) ให้คะแนนใน 30 ประเด็น และ 3.3) ผังงานรูปแบบ C (การตรวจสอบเงื่อนไข) ให้คะแนนใน 30 ประเด็น 4) การแสดงผลลัพธ์ของตัวแปร ให้คะแนนใน 10 ประเด็น และ การกำหนดสิ้นสุด ให้คะแนนใน 5 ประเด็น ตามลำดับความถูกต้องตามเกณฑ์ที่กำหนด

เกณฑ์การตรวจให้คะแนนผังงานที่สร้างขึ้น ได้ผ่านการตรวจสอบคุณภาพเบื้องต้นจากผู้เชี่ยวชาญ 5 คน และการพัฒนาเกณฑ์การตรวจให้คะแนนผังงานแต่ละประเด็นการตรวจเป็นผลมาจากการกำหนดจุดประสงค์การเรียนรู้ ประเด็นการตรวจ และเกณฑ์การตรวจให้คะแนนดีมีความชัดเจนแน่นอนทั้งเนื้อหาที่ไม่กำกวมและมีความเป็นปรนัยเป็นเหตุผลหนึ่งที่ทำให้ผู้ตรวจทั้ง 5 คนให้คะแนนสอดคล้องกัน เพราะผู้ตรวจได้ศึกษาประเด็นการตรวจและเกณฑ์การตรวจที่ผู้วิจัยสร้างขึ้น จึงทำให้ผู้ตรวจทำความเข้าใจประเด็นการตรวจและเกณฑ์การตรวจให้คะแนนได้ดี ทำให้ตรวจสอบความรู้ความเข้าใจและข้อบกพร่องของการเรียนรู้ตามทักษะได้ชัดเจน แต่การตรวจให้คะแนนแบบแยกส่วน (Analytic scoring rubrics) จะเสียเวลาในการตรวจต้องรวมคะแนนเป็นคะแนนรวมทั้งหมดของผังงาน การตรวจให้คะแนนตามเกณฑ์นี้ต้องใช้เวลามาก และผู้ตรวจต้องมีความรู้อย่างแท้จริง ทำให้มีขั้นตอนมากกว่า ซึ่งแตกต่างกับเกณฑ์การให้คะแนนผังงานแบบภาพรวม (Holistic scoring rubrics) ที่มีวิธีการให้คะแนนผังงานแบบภาพรวม โดยการจัดกลุ่มคุณภาพของผังงานออกกลุ่มแล้วให้คะแนน จะชี้ให้เห็นคุณภาพของการปฏิบัติงานการประเมินให้คะแนนการเรียนรู้ได้อย่างรวดเร็ว และสอดคล้องกับงานวิจัยของ กาญจนวาศี (2550) ที่ว่า ไม่ว่าจะใช้รูปแบบการตรวจแบบใด ถ้าเป็นไปได้ควรให้ผู้ตรวจหลาย ๆ คน ใน

การให้คะแนนผู้สอบแต่ละคนแล้วการหาค่าเฉลี่ยหรือคะแนนรวมแทนความสามารถของผู้สอบ จะช่วยลดความแปรปรวนอันเนื่องมาจากผู้ประเมินและความคลาดเคลื่อนที่เหลือน ซึ่งจะทำให้ความเที่ยงสูงขึ้น

2. ผลการศึกษาค่าความน่าเชื่อถือระหว่างวิธีการให้คะแนนผังงาน ภายใต้เงื่อนไขรูปแบบผังงานและจำนวนผู้ตรวจให้คะแนน

การเปรียบเทียบการให้คะแนนผังงานระหว่างวิธีการให้คะแนนแบบทั่วไปกับวิธีที่พัฒนาขึ้น ภายใต้เงื่อนไขรูปแบบผังงานจากผลทดสอบสหสัมพันธ์รายคู่ของผู้ตรวจมีค่าสหสัมพันธ์กันสูง มีแนวโน้มการตรวจให้คะแนนสอดคล้องกัน ผลการวิจัยเป็นไปตามสมมติฐาน เมื่อจำแนกตามรูปแบบ A รูปแบบ B และรูปแบบ C การให้คะแนนผังงานระหว่างวิธีที่พัฒนาขึ้นและวิธีการให้คะแนนผังงานแบบทั่วไป รูปแบบละ 10 ชิ้นงาน ทั้งหมดมีค่าความน่าเชื่อถือภายในได้ไม่ต่ำกว่า 0.8 จึงยอมรับได้ว่ามีค่าความน่าเชื่อถือสูงทุกกรณี แสดงว่า วิธีการตรวจให้คะแนนผังงานที่พัฒนาขึ้น มีประสิทธิภาพที่จะนำไปใช้ได้จริง เมื่อพิจารณาผลการตรวจภายใต้เงื่อนไขของผลการให้คะแนนแบบทั่วไป และผลการหาค่าความน่าเชื่อถือภายในวิธีการให้คะแนนผังงานที่พัฒนาขึ้น จำแนกตามกลุ่มระดับคะแนน 3 กลุ่ม คือ 1) ชิ้นงานกลุ่มระดับคะแนนสูง ผลการตรวจให้คะแนนของผังงานรูปแบบ A มีค่าต่ำสุด เท่ากับ .67 ของการหาค่าความน่าเชื่อถือของการวิจัยครั้งนี้ 2) ชิ้นงานกลุ่ม

ระดับคะแนนปานกลาง และ 3) ชี้นำกลุ่มระดับคะแนนสูง ทั้งหมด มีค่าความน่าเชื่อถือได้ไม่ต่ำกว่า 0.8 จึงยอมรับได้ว่ามีค่าความน่าเชื่อถือสูงทุกกรณี แสดงว่า วิธีการตรวจให้คะแนนผังงานที่พัฒนาขึ้น มีประสิทธิภาพที่จะนำไปใช้ได้จริง

ผลการทดสอบความสอดคล้องการให้คะแนนผังงานระหว่างวิธีที่พัฒนาขึ้นกับวิธีการให้คะแนนแบบทั่วไป ภายใต้เงื่อนไขรูปแบบผังงานและผู้ตรวจให้คะแนน ทดสอบความสัมพันธ์รายคู่ ของผู้ตรวจ มีค่าความน่าเชื่อถือ เท่ากับ .93 และสัมประสิทธิ์สหสัมพันธ์ความสอดคล้องของผู้ตรวจให้คะแนน มีค่าอยู่ระหว่าง .74 - .93 จึงยอมรับได้ว่ามีค่าความน่าเชื่อถือ ดังนั้นวิธีการให้คะแนนผังงานที่พัฒนาขึ้นมีประสิทธิภาพที่จะนำไปใช้ได้จริง

จากการหาความสอดคล้องภายในของผู้ตรวจ ซึ่งให้เห็นถึงระดับความสัมพันธ์ของการตรวจที่เป็นไปในทิศทางเดียวกัน เป็นผลมาจากเกณฑ์การให้คะแนนที่ผู้ตรวจนำไปประเมินชิ้นงานของนักเรียนมีความเที่ยงของเครื่องมือวัด มีวิธีการหาความเชื่อมั่นโดยผู้วัดต่างคนกัน (Inter-Rater Reliability: IRR) มาใช้กับวิธีการตรวจให้คะแนนการประเมินผลอย่างกว้างขวางมากขึ้นและสอดคล้องกับการศึกษาวิจัยของ Hallgren (2012) ได้ศึกษาวิจัยการหาความน่าเชื่อถือโดยผู้วัดต่างคนกันสำหรับประเมินผลการเก็บข้อมูลแบบอัตนัย และ De Groef et al. (2017) ได้ใช้วิธีการประเมินความเที่ยงโดยผู้วัดต่างคนกัน

(IRR) สำหรับการตรวจวัดไหล่ทางร่างกายในหญิงวัยกลางคน ด้านอายุรเวชศาสตร์ทางการแพทย์ และ Lucas et al. (2017) ได้ศึกษาวิจัยความเที่ยงโดยผู้วัดต่างคนกัน (IRR) โดยใช้เกณฑ์การประเมิน ความเที่ยงของเครื่องมือ และค่าความเที่ยงของผู้ตรวจ (Rater Reliability) สรุปว่าการใช้เกณฑ์การให้คะแนนของผู้ตรวจ 4 คน มีความตรงกันสูง และคุณภาพของเครื่องมือวัดมีความเป็นเครื่องมือประเมิน Reflective thinking process ของนักศึกษาเภสัชได้

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. วิธีการให้คะแนนผังงานสำหรับการประเมินผลการเรียนรู้ของนักเรียนที่พัฒนาขึ้น เป็นเกณฑ์การให้คะแนนที่สามารถนำไปประยุกต์ได้กับรายวิชาอื่น แต่ควรศึกษาคู่่มือการใช้เกณฑ์การให้คะแนนเพื่อทราบถึงส่วนประกอบของเกณฑ์

2. ผู้ที่สนใจนำแบบประเมินการให้คะแนนผังงานสำหรับการประเมินผลการเรียนรู้ของนักเรียน จะเป็นครูผู้สอนรายวิชาวิทยาศาสตร์ คอมพิวเตอร์และรายวิชาอื่นทุกระดับชั้น สามารถนำข้อมูลที่ได้จากการวิจัยนี้ไปใช้ตัดสินใจจำนวนผู้ประเมินและจำนวนชิ้นงานที่เหมาะสม เมื่อเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ และเกณฑ์การให้คะแนนแบบภาพรวม

3. ผู้ที่สนใจนำวิธีการให้คะแนนฝั่งงานสำหรับการประเมินผลการเรียนรู้ของนักเรียน ไปพัฒนาโปรแกรมในรูปแบบ Web application หรือแบบออนไลน์ที่พัฒนาขึ้น บางครั้งระบบสารสนเทศที่สร้างขึ้นเกิดความล้มเหลวและไม่ได้รับการยอมรับจากผู้ใช้งาน ดังนั้นการพัฒนาโปรแกรมคอมพิวเตอร์ จะต้องผ่านการทดลองใช้ ทดสอบ และการยอมรับจากผู้ใช้งานก่อน

ข้อเสนอแนะเพื่อการทำวิจัยครั้งต่อไป

1. ควรมีการพัฒนาเกณฑ์การตรวจให้คะแนน ให้มีคุณภาพและเหมาะสมกับผู้เรียนในระดับต่าง ๆ สอดคล้องกับการเรียนการสอนในรายวิชาอื่น ๆ โดยไม่จำเป็นจะต้องเน้นในรายวิชาที่มีการปฏิบัติอย่างเดียว เช่น วิชาวิทยาศาสตร์ คอมพิวเตอร์

2. ควรมีการเพิ่มเงื่อนไขในการศึกษาเพื่อควบคุมแหล่งที่คาดว่าจะมีความคลาดเคลื่อนไว้ล่วงหน้า เช่น จำนวนครั้งในการตรวจให้คะแนน ประสิทธิภาพของผู้ตรวจให้คะแนน และประสิทธิภาพของผู้เชี่ยวชาญ

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนอุดหนุนการวิจัยประเภทบัณฑิตศึกษา จากสำนักงานคณะกรรมการการวิจัยแห่งชาติประจำปี 2562

References

- De Groef, A., Kampen, V., Vervloesem, M., Clabau, N., Christiaens, E., Neven, M. R., & Devoogdt, N. (2017). Inter-rater reliability of shoulder measurements in middle-aged women. *Physiotherapy*, 103(2), 222-230.
- Fraile, J., Panadero, E., & Pardo, R. (2017). Co-creating rubrics: The effects on self-regulated learning, self-efficacy and performance of establishing assessment criteria with students. *Studies in Educational Evaluation*, 53(1), 69-76.
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23-34.
- Kanchanawasi S. (2011). *Classical test theory* (7th ed.). Chulalongkorn University Publishing.
- Lucas, C., Bosnic-Anticevich, S., Schneider, C. R., Bartimote-Aufflick, K., McEntee, M., & Smith, L. (2017). Inter-rater reliability of a reflective rubric to assess pharmacy students' reflective thinking. *Currents in Pharmacy Teaching and Learning*, 9(6), 989-995.
- Nitko, A. J. (2001). *Educational assessment of students* (3rd ed.). Prentice-Hall, Inc.
- Office of the Education Council Secretariate, Ministry of Education. (2560). *Thailand Education Report 2015/2016*. <http://www.backoffice.onec.go.th/uploads/Book/1530-file.pdf>.

- Panadero, E., Tapia, J. A., & Huertas, J. A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and Individual Differences*, 22(6), 806-813.
- Prakobbhol, D. (2558). *Structure of data and algorithm*. Success Media.
- Rezaei, A. R., & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing*, 15(1), 18-39.
- Wollenschlager, M., Hattie, J., Machts, N., Moller, J., & Harms, U. (2016). What makes rubrics effective in teacher-feedback? Transparency of learning goals is not enough. *Contemporary Educational Psychology*, 44(1), 1-11.