

3) ผลการพัฒนาแบบประเมินชิ้นงาน

ผลการตรวจสอบความตรงเชิงเนื้อหา (Content Validity) ของเกณฑ์การประเมิน 9 ตัวบ่งชี้ พบว่า ผลการทดลองใช้แบบประเมินชิ้นงานครั้งที่ 1 ความตรงตามเกณฑ์ (criterion-related validity) จากค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากการประเมินรายกลุ่ม กับผลการเรียนเฉลี่ย (GPA) รายกลุ่ม อยู่ในระดับสูง ($r_{xy} = 0.777$) เมื่อพิจารณาในรายองค์ประกอบพบว่า ผลการเรียนเฉลี่ย (GPA) รายกลุ่ม มีความสัมพันธ์กับ ผลการประเมินรายองค์ประกอบ ดังนี้ 1) การคิดอย่างมีวิจารณญาณและการแก้ปัญหา อยู่ในระดับสูง ($r_{xy} = 0.760$) 2) การคิดสร้างสรรค์และนวัตกรรมอยู่ในระดับต่ำ ($r_{xy} = 0.479$) 3) การสื่อสารและการร่วมมืออยู่ในระดับต่ำ ($r_{xy} = 0.472$) ผลการตรวจสอบความเที่ยง (Reliability) จากการวิเคราะห์ความสอดคล้องของผู้ประเมิน (Rater Agree Index - RAI) 2 คน พบว่ามีความสอดคล้องในระดับมาก (RAI = 0.65)

ผลการทดลองใช้แบบประเมินชิ้นงานครั้งที่ 2 ความตรงตามเกณฑ์ (criterion-related validity) จากการหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากการประเมินรายกลุ่ม กับผลการเรียนเฉลี่ย (GPA) รายกลุ่ม อยู่ในระดับสูง ($r_{xy} = 0.767$) เมื่อพิจารณาในรายองค์ประกอบพบว่า ผลการเรียนเฉลี่ย (GPA) รายกลุ่ม มีความสัมพันธ์กับผลการประเมินรายองค์ประกอบ ดังนี้ 1) การคิดอย่างมีวิจารณญาณและการแก้ปัญหา อยู่ในระดับปานกลาง ($r_{xy} = 0.592$) 2) การคิดสร้างสรรค์และนวัตกรรมอยู่ในระดับกลาง ($r_{xy} = 0.507$) 3) การสื่อสารและการร่วมมือ อยู่ในระดับต่ำ ($r_{xy} = 0.405$) ผลการตรวจสอบความเที่ยง (Reliability) จากการวิเคราะห์ความสอดคล้องของผู้ประเมิน (Rater Agree Index - RAI) 2 คน พบว่ามีความสอดคล้องในระดับมาก (RAI = 0.88)

4) ผลการพัฒนาแบบประเมินตนเองของนักเรียน

ผลการตรวจสอบความตรงเชิงเนื้อหาของข้อคำถาม (Content Validity) ผ่านการตรวจสอบจากผู้เชี่ยวชาญและใช้ได้ 47 ข้อ

ผลการตรวจสอบความเที่ยง (Reliability) โดยหาค่าสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's alpha coefficient) จากการใช้เครื่องมือครั้งที่ 1 อยู่ในระดับสูงมาก ($\alpha = 0.911$) เมื่อพิจารณารายองค์ประกอบอยู่ในระดับสูงทุกองค์ประกอบ การใช้เครื่องมือครั้งที่ 2 อยู่ในระดับสูงมาก ($\alpha = 0.952$) เมื่อพิจารณารายองค์ประกอบอยู่ในระดับสูงทุกองค์ประกอบ

จากผลการวิจัยดังกล่าวพบว่า หากใช้แบบประเมินชิ้นงาน ในการประเมินการคิดอย่างมีวิจารณญาณ และการแก้ปัญหา และการคิดสร้างสรรค์และนวัตกรรมจะให้ค่าความตรงตามเกณฑ์ (criterion-related validity) สูงกว่าแบบสังเกตพฤติกรรม เนื่องจากองค์ประกอบเกี่ยวกับกระบวนการทางสมองที่อยู่ภายในตัวผู้เรียน ผู้สังเกตจะสามารถสังเกตได้ก็ต่อเมื่อนักเรียนแสดงสิ่งเหล่านี้ออกมาผ่านชิ้นงาน หรือผลงาน และหากใช้แบบสังเกตพฤติกรรมในการประเมินการสื่อสารและการร่วมมือจะให้ค่าความตรงตามเกณฑ์ (criterion-related validity) ที่สูงกว่าแบบประเมินชิ้นงาน เนื่องจากเป็นพฤติกรรมที่ผู้ประเมินสามารถสังเกตได้โดยตรงจากผู้เรียน

2.3 ผลการตรวจสอบคุณภาพชุดเครื่องมือประเมินทักษะการเรียนรู้และนวัตกรรมในด้านความถูกต้อง ความเหมาะสม ความเป็นไปได้ และความเป็นประโยชน์

ผลการตรวจสอบความถูกต้อง ความเหมาะสม ความเป็นไปได้ และความเป็นประโยชน์ของชุดเครื่องมือการประเมินทักษะการเรียนรู้และนวัตกรรม โดยใช้ค่าเฉลี่ยจากผู้ประเมิน 3 คน พบว่า กิจกรรมกระตุ้นทักษะการเรียนรู้และนวัตกรรม มีความถูกต้องในระดับมากที่สุด ความเหมาะสมในระดับมาก ความเป็นไปได้ในระดับมาก ความเป็นประโยชน์ในระดับมากที่สุด แบบสังเกตพฤติกรรมมีความถูกต้องในระดับมากที่สุด ความเหมาะสมในระดับ

มากที่สุด ความเป็นไปได้ในระดับมาก ความเป็นประโยชน์ในระดับมากที่สุด แบบประเมินชิ้นงานและแบบประเมินตนเองของนักเรียนมีความถูกต้อง ความเหมาะสม ความเป็นไปได้ ความเป็นประโยชน์ในระดับมากที่สุด

กิตติกรรมประกาศ

ผู้วิจัยขอขอบพระคุณ ผู้อำนวยการและครู โรงเรียนฝางวิทยาลัย และ โรงเรียนกัลยาณวัตร ที่ให้ความช่วยเหลือในการเก็บรวบรวมข้อมูล และคณาจารย์สาขาวิชาการวัดและประเมินผลการศึกษาทุกท่านที่ประสิทธิ์ประสาทความรู้ วิทยากรต่างๆ แลเอาใจใส่ และให้คำแนะนำที่ดีอันเป็นประโยชน์ต่อการวิจัยในครั้งนี้

เอกสารอ้างอิง

- จตุภูมิ เขตจัตุรัส. การวัดและประเมินผลการศึกษาในศตวรรษที่ 21. ขอนแก่น: โรงพิมพ์หาวิทยาลัยขอนแก่น; 2560.
- น้ำทิพย์ อองอาจฉิมชัย. การพัฒนาแบบวัดทักษะแห่งศตวรรษที่ 21 ตามการรับรู้ของนักเรียนมัธยมศึกษาตอนต้น: การประยุกต์ใช้แนวคิดการเข้าถึงคุณลักษณะที่มุ่งวัดของแบบสอบ. วิทยานิพนธ์ปริญญาครุศาสตรดุษฎีบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย; 2556.
- ปทุมธริกา น้อยนนท์. การพัฒนารูปแบบการประเมินการเรียนรู้ในศตวรรษที่ 21 สำหรับสาขาวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 4. วิทยานิพนธ์ปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาวิจัยและประเมินผลการศึกษา บัณฑิตวิทยาลัย มหาวิทยาลัยมหาสารคาม; 2559.
- วรพจน์ วงศ์กิจรุ่งเรือง, อธิป จิตตฤกษ์. ทักษะแห่งอนาคตใหม่: การศึกษาเพื่อศตวรรษที่ 21. กรุงเทพฯ: โอเพ่นเวิลด์ส์; 2554.
- วิภาวี ศิริลักษณ์. การพัฒนาตัวบ่งชี้ทักษะของนักเรียนในศตวรรษที่ 21. วิทยานิพนธ์ปริญญาการศึกษามหาบัณฑิต สาขาวิชาวิจัยและประเมินผลการศึกษา บัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร; 2558.
- ศิริชัย กาญจนวาสี. ทฤษฎีการทดสอบแบบดั้งเดิม. พิมพ์ครั้งที่ 6. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย; 2552.
- _____. ทฤษฎีการประเมิน. พิมพ์ครั้งที่ 7. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย; 2552.
- Hair, J. F., Jr., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. Multivariate data analysis (6 ed.). Upper Saddle River, New Jersey: Prentice Hall; 2006.
- Lai, E.R. & Viering M. Assessing 21st Century Skills: Integrating Research Findings. In annual meeting of the National Council on Measurement in Education, Vancouver, BC, Canada; 2012.
- Partnership for 21st Century Skills. (2009). P21 framework definitions document. Retrieved February 1, 2017, from <http://www.21stcenturyskills.org>