



Comparison of The Reliability of The Written Mathematical Communication Skills Test for Students in Grade 6: Application of Generalizability Theory

Monchaya Sakkamart, Kamontip Srihaset, and Darunee Tippayakulpairoj

Department of Evaluation and Research, Faculty of Education, Ramkhamhaeng University, Thailand

E-mail: monchaya.s@rumail.ru.ac.th, ORCID ID: <https://orcid.org/0009-0007-7392-9485>

E-mail: ksrihaset@gmail.com, ORCID ID: <https://orcid.org/0000-0002-4381-6128>

E-mail: newbobow@hotmail.com, ORCID ID: <https://orcid.org/0000-0002-4874-2607>

Received 16/04/2025

Revised 23/04/2025

Accepted 30/05/2025

Abstract

Background and Aims: Mathematical communication skills are essential for effective mathematics instruction. Therefore, the assessment of written mathematical communication skills requires high-quality instruments. This research aimed to: 1) develop an assessment tool for evaluating written mathematical communication skills of Grade 6 students and 2) compare the reliability of the assessment tool for written mathematical communication skills of Grade 6 students with the application of generalizability theory using a fully crossed design. In which all raters score all items from all examinees.

Methodology: The sample used in this research consisted of 80 Grade 6 students from the Ramkhamhaeng University Demonstration School (Bangna Campus) in the academic year 2024. The sample size was determined using the G*Power 3.1.9.7 program, and cluster sampling was employed. The research instrument was an assessment tool for written mathematical communication skills. The statistical methods used for data analysis included basic descriptive statistics, the kappa coefficient to assess inter-rater reliability, and generalizability coefficients.

Results: The research findings revealed that: 1) the developed assessment tool for written mathematical communication skills on the topic of fractions had a content validity index ranging from 0.67 to 1.00. The overall reliability of the test was 0.762. The kappa coefficients for inter-rater reliability ranged from 0.210 to 0.645. The difficulty index ranged from 0.393 to 0.668, and the discrimination index ranged from 0.379 to 0.882. 2) The assessment tool for written mathematical communication skills with a reliability coefficient of no less than 0.80 can be selected as follows: the generalizability coefficients for both relative and absolute decisions were no less than 0.80 when 2 or 3 raters scored at least 10 items.

Conclusion: The study of the Generalizability Coefficient (G-Coefficient) aimed to identify an assessment tool that is practically applicable and has a reliability of 0.80. The research findings indicated that the written mathematical communication skills assessment tool suitable for use in developing and improving the quality of instruction on fractions should consist of at least 10 items and be scored by 2 raters.

Keywords: The Written Mathematical Communication Skills Test; Application of Generalizability Theory

การเปรียบเทียบความเที่ยงของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนของนักเรียน ชั้นประถมศึกษาปีที่ 6 : การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

มณฑุมา สักกามาต, กมลทิพย์ ศรีหาเศษ และดารุณี ทิพยกุลไพโรจน์
ภาควิชาการประเมินและการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

บทคัดย่อ

ภูมิหลังและวัตถุประสงค์: ทักษะการสื่อสารทางคณิตศาสตร์มีความสำคัญต่อการจัดการเรียนรู้คณิตศาสตร์ การประเมินทักษะการสื่อสารด้านการเขียนจึงต้องใช้เครื่องมือที่มีคุณภาพ ดังนั้น การวิจัยครั้งนี้มีวัตถุประสงค์ เพื่อ 1) พัฒนาแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนของนักเรียนชั้นประถมศึกษาปีที่ 6 และ 2) เปรียบเทียบความเที่ยงของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ ด้านการเขียนของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดซึ่งผู้ตรวจตรวจให้คะแนนข้อสอบทุกข้อของผู้สอบทุกคน

ระเบียบวิธีการวิจัย: กลุ่มตัวอย่างที่ใช้ในการวิจัย คือ นักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนสาธิตมหาวิทยาลัยรามคำแหง วิทยาเขตบางนา ปีการศึกษา 2567 จำนวน 80 คน และใช้การสุ่มตัวอย่างแบบกลุ่ม เครื่องมือที่ใช้ในการวิจัย คือ แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ค่าสถิติพื้นฐาน สัมประสิทธิ์สหสัมพันธ์แคปปาของคุณภาพผู้ตรวจ และสัมประสิทธิ์การสรุปอ้างอิง

ผลการวิจัย: ผลการวิจัยพบว่า 1) แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน เรื่องเศษส่วน ที่พัฒนาขึ้น ความตรงเชิงเนื้อหาของแบบวัดมีค่าดัชนีความสอดคล้องอยู่ในช่วง 0.67 – 1.00 ค่าความเที่ยงทั้งฉบับ เท่ากับ 0.762 สัมประสิทธิ์สหสัมพันธ์แคปปาของคุณภาพผู้ตรวจ อยู่ในช่วง 0.210 – 0.645 ค่าความยากอยู่ในช่วง 0.393 – 0.668 และค่าอำนาจจำแนกอยู่ในช่วง 0.379 – 0.882 และ 2) แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน ที่มีค่าความเที่ยงไม่ต่ำกว่า 0.80 สามารถเลือกใช้แบบวัดได้ดังนี้ ค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์และเชิงสัมบูรณ์ของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน มีค่าไม่ต่ำกว่า 0.80 เมื่อมีผู้ตรวจ จำนวน 2 และ 3 คน ตรวจให้คะแนน 10 ข้อขึ้นไป

สรุปผล: ผลการศึกษาค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) เป็นการศึกษาเพื่อการตัดสินใจเลือกใช้แบบวัดที่สามารถนำไปใช้ได้จริงและมีความเที่ยงเท่ากับ 0.80 ซึ่งผลการวิจัยสะท้อนให้เห็นว่าแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนที่เหมาะสมสำหรับนำไปใช้ในการพัฒนาและปรับปรุงคุณภาพการจัดการเรียนการสอนเรื่องเศษส่วนให้ดีขึ้นต่อไป คือ จำนวนข้อสอบ 10 ข้อขึ้นไป และมีผู้ตรวจ 2 คน

คำสำคัญ: แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน; การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

บทนำ

คณิตศาสตร์เป็นศาสตร์ที่มีบทบาทสำคัญที่ช่วยพัฒนาทักษะการคิดของมนุษย์ และมีความเชื่อมโยงกับศาสตร์อื่น ๆ ที่ต้องใช้ความรู้ทางคณิตศาสตร์ในการแก้ปัญหาของศาสตร์นั้น ๆ ซึ่งการเรียนรู้คณิตศาสตร์คือการพัฒนาความสามารถทางคณิตศาสตร์ของนักเรียนเกี่ยวกับการแก้ปัญหา การให้เหตุผลและการพิสูจน์ การสื่อสาร และการเชื่อมโยง เป็นต้น โดยหนึ่งในทักษะที่สำคัญของการเรียนคณิตศาสตร์คือทักษะการสื่อสาร ซึ่งเป็นการแสดงถึงการให้เหตุผล และการกำหนดกลยุทธ์การแก้ปัญหา การใช้ภาษาและการดำเนินการเชิงสัญลักษณ์ เชิงรูปแบบ และเชิงเทคนิค ซึ่งทักษะการสื่อสารทางคณิตศาสตร์ด้วยการเขียน เป็นทักษะที่สะท้อนถึงความเชี่ยวชาญคำศัพท์ โครงสร้างทางคณิตศาสตร์ และสัญลักษณ์ ซึ่งสามารถวัดได้จากความสามารถและทักษะของนักเรียนในการแก้ปัญหาทางคณิตศาสตร์ (Dewi et al., 2020; Hermawan & Hidayat, 2018; Permatasari, 2021)

Lubis et al. (2023) กล่าวว่าทักษะการสื่อสารทางคณิตศาสตร์มีตัวบ่งชี้หลายประการและสามารถวัดได้จาก 1) การแสดงความคิด สถานการณ์ และความสัมพันธ์ทางคณิตศาสตร์ผ่านการเขียนโดยใช้ของจริง กราฟ พีชคณิต และรูปภาพ 2) การเชื่อมโยงวัตถุ รูปภาพ หรือไดอะแกรมผ่านชีวิตจริงกับรูปแบบของความคิดทางคณิตศาสตร์ และ 3) อธิบายเหตุการณ์ในชีวิตจริงโดยใช้ภาษาหรือสัญลักษณ์ทางคณิตศาสตร์ ซึ่งจากการศึกษางานวิจัยที่เกี่ยวข้องกับทักษะการสื่อสารทางคณิตศาสตร์ สะท้อนให้เห็นถึงความจำเป็นในการให้นักเรียนแสดงความคิดและเขียนแสดงวิธีคิดผ่านการทดสอบ แต่ข้อสอบแบบสร้างคำตอบมีข้อจำกัดทั้งด้านความเที่ยงต่ำ การใช้เวลา ค่าใช้จ่าย และทรัพยากรในการตรวจให้คะแนน (Chui & Wolfe, 2002; Coffman, 1971 อ้างถึงใน ชนิสร่า สงวนไว้, 2558) และข้อจำกัดอีกประการ คือ ความเป็นอัตนัยของการให้คะแนน (subjective) กล่าวคือ คะแนนที่ได้แปรเปลี่ยนไปตามลักษณะของผู้ตรวจและการให้คะแนนขึ้นอยู่กับดุลพินิจของผู้ตรวจ (Bachman et al., 1995; Linacre, 1993; Lunz et al., 1990; Turner, 2003; Smith & Kulikowich, 2004 อ้างถึงใน ชนิสร่า สงวนไว้, 2558) จากปัญหาดังกล่าวนักวิชาการจึงได้เสนอวิธีการลดข้อจำกัดเหล่านั้น Mehrens and Lehmann (1973) ที่กล่าวว่า เพื่อให้การตรวจแบบสอบมีความเที่ยงเพิ่มขึ้น ผู้ตรวจควรใช้วิธีการตรวจที่เหมาะสม ใช้เกณฑ์การตรวจกับนักเรียนทุกคน มีการออกแบบเกณฑ์การให้คะแนนแบบรูบริก (scoring rubric) และควบคุมแหล่งความคลาดเคลื่อน เช่น ข้อสอบและผู้ตรวจ โดยใช้ผู้ตรวจหลายๆ คนในการตรวจข้อสอบแต่ละข้อแล้วจึงหาคะแนนเฉลี่ย จากผลการวิจัยพบว่า เมื่อเพิ่มจำนวนผู้ตรวจจะทำให้ความเที่ยงสูงขึ้น (Brennan et al., 1995; Sudweed et al., 2004; Swartz et al., 1999) แต่จำนวนผู้ตรวจที่เหมาะสมยังไม่สามารถสรุปได้แน่ชัดเนื่องจากผลการศึกษายังพบข้อแตกต่าง

จากหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 (ฉบับปรับปรุง พ.ศ. 2560) (กระทรวงศึกษาธิการ, 2560) มุ่งพัฒนาผู้เรียนทุกคนให้เป็นมนุษย์ที่มีความสมดุลทั้งทางด้านร่างกาย ความรู้ คุณธรรม มีจิตสำนึกในความเป็นพลเมืองไทยและพลเมืองโลกที่ดี ตามมาตรฐานการเรียนรู้และตัวชี้วัดที่กำหนดขึ้น โดยหนึ่งในสมรรถนะสำคัญของผู้เรียนที่หลักสูตรต้องพัฒนาให้เกิดขึ้นกับผู้เรียน คือ ความสามารถในการสื่อสาร ซึ่งการสื่อสารเป็นกระบวนการที่มีความสำคัญต่อมนุษย์ทั้งในด้านการดำเนินชีวิตในสังคม การเมืองและการศึกษา กล่าวได้ว่าการสื่อสารเป็นฟันเฟืองของการดำรงชีวิตในสังคมให้ดำเนินไปอย่างต่อเนื่องโดยไม่หยุดยั้ง ซึ่งการสื่อสารเป็นความสามารถในการฟัง พูด อ่าน และเขียน เป็นกระบวนการแลกเปลี่ยนข่าวสารจากบุคคลหนึ่งไปยังอีกบุคคลหนึ่ง เพื่อต้องการได้รับการตอบสนองให้บรรลุวัตถุประสงค์ที่ต้องการ มุ่งก่อให้เกิดผลของการนำการสื่อสารไปใช้และเกิดการเปลี่ยนแปลงด้านการพัฒนาร่วมกัน (กิติมา สุรสนธิ, 2557) จากความสำคัญของการสื่อสารทางคณิตศาสตร์ที่ถือเป็นทักษะสำคัญที่นักเรียนจำเป็นต้องมี

เพื่อให้เข้าใจคณิตศาสตร์ ดังนั้น การจัดกิจกรรมการเรียนรู้ในชั้นเรียนจึงเน้นให้นักเรียนมีความชำนาญในด้านทักษะการสื่อสาร เพื่อนำไปใช้ได้อย่างมีประสิทธิภาพ (Pantow et al., 2020)

จากความสำคัญดังกล่าว ทางผู้วิจัยจึงมีความสนใจที่จะเปรียบเทียบความเที่ยงของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนของนักเรียนระดับชั้นประถมศึกษาปีที่ 6 ในเรื่องของ เศษส่วน เพื่อนำไปใช้ในการจัดการเรียนการสอนการเตรียมความพร้อมในการสอบทางวิชาการ ของนักเรียนระดับชั้นประถมศึกษาปีที่ 6 โดยใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด โดยผู้วิจัยศึกษาเงื่อนไขการทดสอบที่สำคัญ 2 ปัจจัย ได้แก่ จำนวนข้อสอบ และจำนวนผู้ตรวจ พิจารณาค่าความเที่ยงจากค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability coefficient) เพื่อเป็นแนวทางสำหรับผู้เกี่ยวข้องในการเลือกใช้จำนวนข้อสอบ จำนวนผู้ตรวจและรูปแบบการออกแบบการตรวจให้คะแนนที่ให้ค่าความเที่ยงของคะแนนเหมาะสม ได้แบบสอบที่มีความเที่ยงในระดับที่ต้องการ ประหยัดเวลาในการตรวจ จำนวนผู้ตรวจที่เพียงพอเหมาะสมและการนำไปใช้ที่สามารถให้ข้อมูลที่สอดคล้องกับสภาพจริงในทางปฏิบัติเพื่อให้ได้แบบสอบที่มีคุณภาพต่อการนำไปใช้และเป็นสารสนเทศประกอบการตัดสินใจในการตรวจให้คะแนนแบบวัดทักษะการสื่อสารทางคณิตศาสตร์

วัตถุประสงค์การวิจัย

1. เพื่อพัฒนาแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนของนักเรียนชั้นประถมศึกษาปีที่ 6
2. เพื่อเปรียบเทียบความเที่ยงของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ ด้านการเขียนของนักเรียนชั้นประถมศึกษาปีที่ 6 ที่มีวิธีการออกแบบโดยผู้ตรวจตรวจให้คะแนนทุกข้อของผู้สอบทุกคน ($p \times i \times r$) โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

การทบทวนวรรณกรรม

ทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน

ทักษะการสื่อสาร มี 4 ด้าน คือ ด้านการฟัง การพูด การอ่าน และการเขียน การสื่อสารทางคณิตศาสตร์ตามโมเดลทางคณิตศาสตร์เกี่ยวกับการสื่อสารของ Shannon (Shannon's mathematical model of communication) ซึ่งอัมพร ม้าคะนอง (2559) กล่าวว่า การสื่อสารทางคณิตศาสตร์เป็นการสื่อสารและสื่อความหมายที่เกี่ยวข้องกับคณิตศาสตร์ระหว่างผู้ส่งสารกับผู้รับสาร ซึ่งมีความสำคัญที่เห็นได้จากผลงานที่แสดงความสัมพันธ์ระหว่างคณิตศาสตร์โมเดลทางคณิตศาสตร์เกี่ยวกับการสื่อสารของ Shannon (Shannon's mathematical model of communication) ในปี ค.ศ.1949 และ โมเดลดังกล่าวได้ถูกกล่าวถึงและนำมาใช้ในการพัฒนาการสื่อสารทางคณิตศาสตร์ตลอดมา

การสื่อสารทางคณิตศาสตร์มีทั้งการเขียนและการพูด โดยทักษะการสื่อสารด้านการเขียนเป็นทักษะที่ต้องเน้นและให้ความสำคัญเพิ่มขึ้นในการจัดการเรียนการสอนในชั้นเรียน เนื่องจากทักษะการสื่อสารด้านการเขียนสามารถช่วยพัฒนาการเรียนรู้ ทำให้นักเรียนสามารถเรียนรู้และทำความเข้าใจเนื้อหาได้ และเพิ่มทักษะในการจัดการกับปัญหาทางคณิตศาสตร์ของนักเรียน การสื่อสารเป็นลายลักษณ์อักษรผ่านทักษะการสื่อสารด้านการเขียนจึงมีความสำคัญสำหรับนักเรียน เนื่องจากบทบาทสำคัญของการเขียนคือการเชื่อมโยงระหว่างข้อมูลที่นักเรียนทราบแล้วเรียบเรียงผ่านการเขียนด้วยข้อมูลที่เข้าใจ ทักษะนี้จึงสะท้อนความเข้าใจทางคณิตศาสตร์ของนักเรียนได้เป็นอย่างดี และช่วยประเมินความสามารถของผู้เรียนที่ทำให้ครูเข้าใจนักเรียนมากยิ่งขึ้น อีกทั้งทักษะการสื่อสารด้านการเขียนยังมีบทบาทสำคัญในกระบวนการเรียนรู้หรือการแก้ปัญหา

และสะท้อนความคิดต่างๆ ของนักเรียนในการแก้ปัญหาคณิตศาสตร์ (Meel, 1999; Meier & Rishel, 1998; Morgan, 2001; National Council Of Teachers of Mathematics (NCTM), 2000; Teledahl, 2017)

แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ที่ต้องการให้นักเรียนสามารถอธิบาย ชี้แจง แสดงความเข้าใจ หรือความคิดเกี่ยวกับคณิตศาสตร์ของตนเองให้ผู้อื่นรับรู้ โดยใช้ทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนเพื่อถ่ายทอดความเข้าใจ สามารถอธิบายหลักการ แนวคิด วิธีคิด การสรุปผล ส่งต่อสารนั้นไปให้ผู้อื่นเข้าใจได้ และการตรวจสอบการสื่อสารของนักเรียนว่าสามารถเขียนสื่อสารให้ผู้ตรวจเข้าใจได้หรือไม่นั้นจะต้องมีครูผู้ตรวจ เพื่อตรวจสอบความถูกต้องตามหลักคณิตศาสตร์ โดยได้กำหนดขั้นตอนการวัดทักษะการสื่อสาร 5 ขั้นตอน คือ ขั้นที่ 1 การเขียนเพื่อระบุแนวคิดหรือหลักการที่ใช้ในการหาคำตอบ ขั้นที่ 2 การเขียนเพื่ออธิบายเหตุผลในการเลือกใช้แนวคิดหรือหลักการในการหาคำตอบ ขั้นที่ 3 การเขียนแสดงวิธีทำ ขั้นที่ 4 การเขียนผลลัพธ์ที่ได้จากการแสดงวิธีทำ และ ขั้นที่ 5 การเขียนสรุปคำตอบที่ถูกต้อง

ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) หรือ G-Theory

ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) หรือ G-Theory เป็นทฤษฎีสำหรับวิเคราะห์ความน่าเชื่อถือของผลการวัดที่แสดงประสิทธิภาพ การสรุปอ้างอิงผลของการวัดที่ได้จากแบบสอบไปยังคะแนนเฉลี่ยที่ผู้สอบแต่ละคนควรได้รับ ภายใต้สถานการณ์สอบหรือเงื่อนไขต่างๆ ของการวัดหรือแหล่งความคลาดเคลื่อนต่างๆ ได้ ซึ่งจะเป็นประโยชน์ต่อการพัฒนาแบบวัดให้มีระดับความเที่ยงที่ต้องการภายใต้สถานการณ์ทดสอบที่ต้องการนำแบบวัดไปใช้

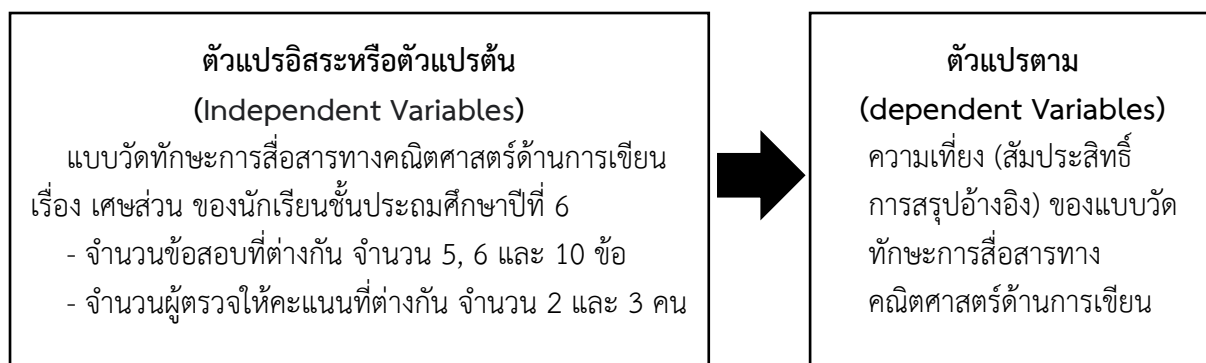
นอกจากนี้คะแนนที่ได้จากการวัดภายใต้สถานการณ์หรือเงื่อนไขเดียวกันด้วยแบบชุดเดียว และการทำข้อสอบเพียงครั้งเดียว ตามแนวคิดของทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (G-Theory) คะแนนที่ได้จะไม่สามารถให้ความน่าเชื่อถือได้อย่างเต็มที่ เนื่องจากคะแนนที่ได้จะไม่เป็นตัวแทนที่ดีของคะแนนจริงซึ่งเป็นคะแนนเฉลี่ยที่ได้จากการทดสอบหลายๆ สถานการณ์ด้วยแบบสอบหลายๆ ชุด และการทำข้อสอบหลายๆ ครั้ง เนื่องจากผู้สอบคนเดียวกันเมื่อทำการทดสอบหลายๆ ครั้งในหลายสถานการณ์ มักจะได้คะแนนในแต่ละครั้งแตกต่างกัน ความไม่คงเส้นคงวาของคะแนนเกิดจากความคลาดเคลื่อนของผลการวัด ซึ่ง G-Theory จะพิจารณาถึงแหล่งความคลาดเคลื่อนพหุ (Multiple sources of error) กล่าวคือ สามารถวิเคราะห์ความคลาดเคลื่อนจากหลายแหล่งพร้อมกัน จึงเป็นส่วนช่วยให้ผู้บริหารการทดสอบสามารถตัดสินใจได้ว่าควรใช้สถานการณ์หรือเงื่อนไขการวัดแบบใดจึงจะทำให้ได้คะแนนที่น่าเชื่อถือตามระดับที่ต้องการ จากผลการวิเคราะห์ด้วย G-Theory จะให้ค่าสัมประสิทธิ์เชิงสรุปที่แสดงถึงความน่าเชื่อถือของคะแนนที่ได้จากผลการวัด (Level of dependability) เรียกว่า สัมประสิทธิ์ การสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (generalizability Coefficient) ซึ่งคล้ายกับสัมประสิทธิ์ความเที่ยง (Reliability Coefficient) ในทฤษฎีการทดสอบแบบดั้งเดิม (ศิริชัย กาญจนวาสี, 2555)

กล่าวโดยสรุป ได้ว่าทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) หรือ G - Theory เป็นทฤษฎีที่ขยายแนวคิดมาจากทฤษฎีการทดสอบแบบดั้งเดิม ดังนี้ สามารถประมาณค่าความแปรปรวนจากแหล่งต่าง ๆ ได้จากการวัดบริบทเดียวกันและ จากการวิเคราะห์เพียงครั้งเดียว สามารถประมาณค่าสัมประสิทธิ์ความเที่ยงโดยเน้นไปที่ องค์ประกอบความแปรปรวนของแต่ละแหล่งความคลาดเคลื่อนของการวัด สามารถแบ่งการตัดสินใจออกเป็น 2 ลักษณะคือ การตัดสินใจเชิงสัมพัทธ์ (relative decision) และการตัดสินใจเชิงสมบูรณ์ (absolute decision) และ สามารถประมาณน้ำหนักของแหล่งความคลาดเคลื่อนของการวัดที่ สามารถเป็นไปได้ทั้งหมด (G-Study) และสามารถนำข้อมูลที่ได้มาใช้

ในการออกแบบการวัดในครั้งนี้ต่อไป เพื่อเป็นการลดความคลาดเคลื่อนของการวัดให้น้อยลง (D-Study) ซึ่งเป็นการออกแบบการวัด ให้ความสำคัญกับค่าทั้งค่าใช้จ่ายและเวลา และยังสามารถออกแบบการวัดได้หลายรูปแบบ ทั้งการ ออกแบบการวัดแบบไขว้ (crossed design) การออกแบบการวัดแบบแฝง (Nested design) และการออกแบบการวัดแบบผสม (confounded design)

กรอบแนวคิดการวิจัย

การศึกษาเรื่องการเปรียบเทียบความเที่ยงของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน ของนักเรียนชั้นประถมศึกษาปีที่ 6 : การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด ภายใต้จำนวนข้อสอบและจำนวนผู้ตรวจที่ต่างกัน โดยแบบสอบที่พัฒนาขึ้นเป็นแบบสอบประเภทความเรียง ซึ่งผู้ตรวจตรวจให้คะแนนข้อสอบทุกข้อของผู้สอบทุกคน โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด ซึ่งประกอบด้วย 2 ฟาเซต คือ จำนวนข้อสอบ และจำนวนผู้ตรวจ (สุรชัย มาตราข และ ประกฤติยา ทักษิณ, 2565; Bachman et al., 1995; Brennan et al., 1995; Linacre, 1993; Lunz et al., 1990; Smith & Kulikowich, 2004 อ้างถึงใน ชนิสรรา สงวนไว้, 2558; Sudweed et al., 2004; Swartz et al., 1999; Turner, 2003) มีกรอบแนวคิดในการวิจัยดังแผนภาพที่ 1



แผนภาพที่ 1 กรอบแนวคิดการวิจัย

ระเบียบวิธีการวิจัย

ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนระดับชั้นประถมศึกษาปีที่ 6 ที่เรียนในโรงเรียนสาธิตมหาวิทยาลัยรามคำแหง หลักสูตรภาคปกติ จำนวน 183 คน (โรงเรียนสาธิตมหาวิทยาลัยรามคำแหง (ฝ่ายประถม), 2567) (โรงเรียนสาธิตมหาวิทยาลัยรามคำแหง วิทยาเขตบางนา, 2567)

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนระดับชั้นประถมศึกษาปีที่ 6 หลักสูตรภาคปกติ ที่เรียนโรงเรียนสาธิตมหาวิทยาลัยรามคำแหง วิทยาเขตบางนา จำนวน 80 คน กำหนดขนาดกลุ่มตัวอย่างด้วยโปรแกรม G*Power 3.1.9.7 โดยใช้การสุ่มตัวอย่างแบบกลุ่ม จำนวนตัวอย่างดังตาราง 1

ตาราง 1 จำนวนกลุ่มตัวอย่าง

ห้องเรียน	จำนวน(คน)
ห้อง ป.6/1	28
ห้อง ป.6/2	26
ห้อง ป.6/3	26
รวม	80

เครื่องมือที่ใช้ในการวิจัย: เครื่องมือที่ใช้ในการวิจัย คือ แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 เรื่อง เศษส่วน ผู้วิจัยได้ศึกษาเอกสาร หลักการ แนวคิดเกี่ยวกับทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 โดยให้สอดคล้องกับเนื้อหาวิชาคณิตศาสตร์ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ.2551 (ฉบับปรับปรุง 2560) โดย (1) กำหนดวัตถุประสงค์ นิยาม พฤติกรรมชี้วัด และสร้างตารางวิเคราะห์เนื้อหาตามหลักสูตร (2) สร้างข้อคำถามตามตารางวิเคราะห์ ซึ่งประกอบด้วยจุดมุ่งหมายและเนื้อหาที่จะวัด และการให้น้ำหนักในตารางโครงสร้าง พร้อมทั้งนำเครื่องมือที่สร้างขึ้นไปให้ผู้เชี่ยวชาญด้านเนื้อหาทางคณิตศาสตร์ และผู้เชี่ยวชาญด้านการวัดและประเมินผล จำนวน 3 ท่าน ดำเนินการตรวจสอบความสอดคล้องระหว่างเนื้อหาการจัดการเรียนรู้และประเมินคุณภาพของเครื่องมือ โดยมีผลการประเมินมีค่าดัชนีความสอดคล้องมากกว่าหรือเท่ากับ 0.5 ปรากฏว่าทุกข้อคำถามมีความเที่ยงตรง และมีค่า IOC ระหว่าง 0.67-1.00 พร้อมทั้งนำมาปรับปรุงแก้ไขตามข้อเสนอแนะ แล้วนำไปทดลองใช้กับนักเรียนจำนวน 30 คนเพื่อวิเคราะห์หาค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) เป็นรายข้อ พบว่าข้อสอบแต่ละข้อมีค่าค่าความยาก (p) อยู่ในช่วง 0.393 – 0.668 และค่าอำนาจจำแนก (r) อยู่ในช่วง 0.379 – 0.882 จากนั้นจึงได้คัดเลือกข้อสอบที่ใช้ได้และปรับปรุงให้มีความเหมาะสม และค่าค่าความเที่ยงทั้งฉบับ เท่ากับ 0.762

ในการวัดครั้งนี้ได้คัดเลือกผู้ตรวจที่จะมาตรวจแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน จำนวน 2 คน โดยมีคุณสมบัติ ดังนี้ ผู้มีการศึกษาอยู่ในระดับปริญญาตรีขึ้นไป กำลังสอนวิชาคณิตศาสตร์ในระดับชั้นประถมศึกษาปีที่ 6 และมีประสบการณ์ในการสอน 3 ปีขึ้นไป ในการวิจัยครั้งนี้มีเงื่อนไขการทดสอบที่สำคัญ 2 ปัจจัย คือ จำนวนข้อสอบ (i) และจำนวนผู้ตรวจ (r) ที่มีผลต่อความเที่ยงของแบบสอบการศึกษา G-study เพื่อศึกษาเชิงสรุปอ้างอิงด้วยการออกแบบและวิเคราะห์ความแปรปรวนของคะแนนจริงและความแปรปรวนของคะแนนความคลาดเคลื่อนจาก 2 แหล่ง คือ จำนวนข้อสอบ (i) และจำนวนผู้ตรวจ (r) โดยออกแบบให้ผู้ตรวจตรวจให้คะแนนทุกข้อของผู้สอบทุกคน ($p \times i \times r$)

การเก็บรวบรวมและวิเคราะห์ข้อมูล: ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลกับนักเรียน โรงเรียนสาธิตมหาวิทยาลัยรามคำแหง วิทยาเขตบางนา ระดับชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2567 ดังนี้ (1) อธิบายชี้แจงวัตถุประสงค์ ประโยชน์ที่จะได้รับ และแนะนำวิธีการขั้นตอนในการทำแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน เรื่อง เศษส่วน (2) ดำเนินการทดสอบ โดยใช้แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน เรื่อง เศษส่วน จำนวน 5 ข้อใหญ่ 33 ข้อย่อย รวมคะแนนทั้งหมด 45 คะแนน โดยใน 1 ข้อใหญ่จะเป็นแบบทดสอบที่ให้เขียนอธิบายถึงขั้นตอนการวัดทักษะการสื่อสารทั้ง 5 ขั้นตอน (3) อธิบายชี้แจงวัตถุประสงค์ ประโยชน์ที่จะได้รับ และแนะนำวิธีการในการตรวจให้คะแนนแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน เรื่อง เศษส่วน ให้กับผู้ตรวจเพื่อตรวจให้คะแนน (4) นำผลคะแนนจากการทดสอบด้วยแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน เรื่อง เศษส่วน ที่ได้มาวิเคราะห์ข้อมูลทางสถิติต่อไป

การวิเคราะห์ข้อมูล 1) วิเคราะห์คุณภาพของแบบวัด โดยหาค่าดัชนีความสอดคล้อง (Index of Item Objective Congruence--IOC) ความยาก (p) อำนาจจำแนก (r) ความเที่ยง และความเที่ยงของการตรวจให้คะแนนระหว่างผู้ตรวจ (inter-rater reliability) ด้วยสัมประสิทธิ์สหสัมพันธ์ของแคปปา (Cohen's kappa coefficient) 2) วิเคราะห์คะแนนสอบจากแบบวัด โดยใช้ค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย (mean) ส่วนเบี่ยงเบนมาตรฐาน (standard deviation) คะแนนสูงสุด (maximum) คะแนนต่ำสุด (minimum) โดยใช้โปรแกรมสำเร็จรูป และ 3) เปรียบเทียบความเที่ยงของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนของนักเรียนชั้นประถมศึกษาปีที่ 6 ที่มีวิธีการออกแบบโดยผู้ตรวจตรวจให้คะแนนทุกข้อของผู้สอบทุกคน ($p \times i \times r$) โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด การวิเคราะห์สัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือ (G-Study) ที่มีการออกแบบ Two-Facet Crossed Design คือ $p \times i \times r$ ใช้วิเคราะห์ความแปรปรวนแบบ 3 ทาง (3-way ANOVA)

ผลการวิจัย

1. ผลการพัฒนาแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน

1.1 ผลการพัฒนาแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน จากการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง ผู้วิจัยได้พัฒนาแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน โดยมีเนื้อหาที่ต้องการวัด ประกอบด้วย (1) การเรียงลำดับเศษส่วน (2) การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละแบบมีวงเล็บ (3) การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละแบบไม่มีวงเล็บ (4) โจทย์ปัญหาการบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละ แบบ 2 ขั้นตอน และ (5) โจทย์ปัญหาการบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละ แบบ 3 ขั้นตอน เมื่อนักเรียนอ่านข้อคำถามในแบบวัดแล้ว นักเรียนสามารถตอบคำถามโดยใช้ความรู้ทางคณิตศาสตร์ด้วยกลยุทธ์วิธีต่างๆ ที่หลากหลาย โดยผู้วิจัยสร้างแบบสอบจำนวน 5 ข้อและแต่ละข้อมีคำถามย่อยวัดทักษะการสื่อสารทางคณิตศาสตร์ โดยข้อคำถามแต่ละข้อคำถามย่อยจะวัดทักษะการสื่อสาร 5 ขั้นตอน คือ ขั้นที่ 1 การเขียนเพื่อระบุแนวคิดหรือหลักการที่ใช้ในการหาคำตอบ ขั้นที่ 2 การเขียนเพื่ออธิบายเหตุผลในการเลือกใช้แนวคิดหรือหลักการในการหาคำตอบ ขั้นที่ 3 การเขียนแสดงวิธีทำ ขั้นที่ 4 การเขียนผลลัพธ์ที่ได้จากการแสดงวิธีทำ และ ขั้นที่ 5 การเขียนสรุปคำตอบที่ถูกต้อง

1.2 ผลการตรวจสอบคุณภาพของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน ผลการพัฒนาแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน มีค่าดัชนีความสอดคล้อง (Index of Item Objective Congruence--IOC) โดยพิจารณาคำแนะนำในการแก้ไขปรับปรุงได้ข้อมูลที่เหมาะสม โดยผลการตรวจคุณภาพรายข้อและทั้งฉบับของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน โดยมีค่าความยาก (p) อำนาจจำแนก (r) และค่าความเที่ยง (r_{tt}) พบว่า แบบสอบมีค่าความยาก อยู่ในช่วง 0.393 – 0.668 แสดงว่า แบบสอบปานกลาง และมีค่าอำนาจจำแนก อยู่ในช่วง 0.379 – 0.882 แสดงว่า เหตุการณ์มีความสามารถในการจำแนกผู้สอบได้ดี และมีค่าความเที่ยงเท่ากับ 0.762 รายละเอียดดังตาราง 2

ตาราง 2 ค่าความยาก และค่าอำนาจจำแนก ในแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนที่ได้จากผลการทดลองใช้

ข้อที่	เนื้อหา	p	r
1	การเรียงลำดับเศษส่วน	0.668	0.476
2	การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละแบบมีวงเล็บ	0.662	0.379
3	การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละแบบไม่มีวงเล็บ	0.606	0.505
4	โจทย์ปัญหา การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละ แบบ 2 ขั้นตอน	0.457	0.882
5	โจทย์ปัญหา การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละ แบบ 3 ขั้นตอน	0.393	0.752
r_{tt}		0.762	

ส่วนผลการตรวจคุณภาพคุณภาพของการตรวจให้คะแนน ในแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ ด้านการเขียน โดยค่าความเที่ยงของการตรวจให้คะแนนระหว่างผู้ตรวจ (inter-rater reliability) วิเคราะห์ด้วยสัมประสิทธิ์สหสัมพันธ์ของแคปปา (Cohen's kappa coefficient) ส่วนค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจของผู้ตรวจจำนวน 2 คน มีนัยสำคัญทางสถิติที่ระดับ 0.05 ทุกเนื้อหา โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ของแคปปา (Cohen's kappa coefficient) อยู่ในช่วง 0.210 ถึง 0.645 และเมื่อพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ในแต่ละเนื้อหา จะพบว่าเนื้อหาที่ 5 เรื่อง โจทย์ปัญหา การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละ แบบ 3 ขั้นตอนมีค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนมากที่สุด คือ 0.645 รองลงมาคือ เนื้อหาที่ 1 เรื่อง การเรียงลำดับเศษส่วน, เนื้อหาที่ 4 เรื่อง โจทย์ปัญหา การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละ แบบ 2 ขั้นตอน, เนื้อหาที่ 3 เรื่อง การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละแบบไม่มีวงเล็บ และเนื้อหาที่ 2 เรื่อง การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละแบบมีวงเล็บมีค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนน คือ 0.584, 0.398, 0.338 และ 0.210 ตามลำดับ และเมื่อพิจารณาค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของผู้ตรวจทั้ง 2 คน มีค่าใกล้เคียงกัน รายละเอียดดังตาราง 3

ตาราง 3 ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจแบบสอบของผู้ตรวจจำนวน 2 คน

ข้อที่	เนื้อหา	ผู้ตรวจคนที่ 1		ผู้ตรวจคนที่ 2		r_{xy}
		M_1	SD_1	M_2	SD_2	
1	การเรียงลำดับเศษส่วน	4.80	2.46	4.74	2.57	0.584*
2	การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละแบบมีวงเล็บ	3.59	1.36	3.41	1.58	0.210*
3	การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละแบบไม่มีวงเล็บ	2.91	1.84	2.50	1.75	0.338*
4	โจทย์ปัญหา การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละ แบบ 2 ขั้นตอน	5.59	4.15	5.24	4.22	0.398*
5	โจทย์ปัญหา การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละ แบบ 3 ขั้นตอน	5.01	5.27	4.66	5.29	0.645*

* $p < .05$

2. ผลการเปรียบเทียบความเที่ยงของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน ภายใต้จำนวนข้อสอบที่ต่างกัน ที่ใช้จำนวนผู้ตรวจที่ต่างกัน ตรวจให้คะแนนทุกข้อของผู้สอบทุกคน โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

2.1 ผลการวิเคราะห์คะแนนทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน พบว่าเมื่อพิจารณาคะแนนรวมทั้ง 5 ข้อ จากคะแนนเต็ม 45 คะแนน นักเรียนชั้นประถมศึกษาปีที่ 6 ทำแบบสอบได้คะแนนเฉลี่ย 21.226 คะแนนส่วนเบี่ยงเบนมาตรฐาน 11.628 และนักเรียนได้คะแนนสูงสุด 42.50 คะแนน รายละเอียดดังตาราง 4

ตาราง 4 ผลการวิเคราะห์คะแนนสอบของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน

เนื้อหา	ข้อคำถาม ที่	คะแนน เต็ม	ค่าสถิติพื้นฐาน				
			M	%	SD	MIN	MAX
1. การเรียงลำดับเศษส่วน	1	8	4.772	59.650	2.499	0.00	8.00
2. การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละแบบมีวงเล็บ	2	5	3.500	70.000	1.314	0.00	5.00
3. การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละแบบไม่มีวงเล็บ	3	5	2.703	54.060	1.732	0.00	5.00
4. โจทย์ปัญหา การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละ แบบ 2 ขั้นตอน	4	12	5.117	42.642	4.086	0.00	12.00
5. โจทย์ปัญหา การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละ แบบ 3 ขั้นตอน	5	15	4.834	32.227	5.232	0.00	15.00
รวมทั้งฉบับ		45	21.226	47.169	11.628	1.00	42.50

2.2 ผลการประมาณค่าความแปรปรวนขององค์ประกอบที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน ผลการวิเคราะห์ความแปรปรวนขององค์ประกอบจากแหล่งความแปรปรวนต่างๆ ที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนพบว่า ความแปรปรวนรวมทั้งหมด ประกอบด้วย ความแปรปรวนของผู้สอบ (σ_p^2) ซึ่งเป็นความแปรปรวนจริงของคะแนน (Universe score) ร้อยละ 31.45 ของความแปรปรวนทั้งหมด ความแปรปรวนของคะแนนเฉลี่ยของแต่ละข้อ (σ_i^2) ร้อยละ 9.05 ของความแปรปรวนทั้งหมด ความแปรปรวนของคะแนนเฉลี่ยของผู้ตรวจแต่ละคน (σ_r^2) เท่ากับ ร้อยละ 0.26 ของความแปรปรวนทั้งหมด ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างผู้สอบกับข้อสอบ (σ_{pi}^2) ร้อยละ 52.77 ของความแปรปรวนทั้งหมด ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างผู้สอบกับการตรวจให้คะแนนของผู้ตรวจแต่ละครั้ง (σ_{pr}^2) ร้อยละ 0.50 ของความแปรปรวนทั้งหมดความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างข้อสอบของผู้ตรวจแต่ละคน (σ_{ir}^2) ร้อยละ 0.00 ของความแปรปรวนทั้งหมด และความแปรปรวนของคะแนนความคลาดเคลื่อนหรือส่วนที่เหลือ ($\sigma_{pir,e}^2$) ซึ่งเป็นความแปรปรวนของความคลาดเคลื่อนที่เหลือจาก 3 แหล่ง ร้อยละ 5.97 ของความแปรปรวนทั้งหมด เขียนผลการวิเคราะห์ความแปรปรวน G-Study for $p \times i \times r$ ได้ดังนี้

$$\sigma_{p \times i \times r}^2 = \sigma_p^2 + \sigma_i^2 + \sigma_r^2 + \sigma_{pi}^2 + \sigma_{pr}^2 + \sigma_{ir}^2 + \sigma_{pir,e}^2$$

$$12.707 = 3.996 + 1.150 + 0.033 + 6.706 + 0.063 + 0.001 + 0.758$$

รายละเอียดดังตาราง 5

ตาราง 5 ผลการวิเคราะห์ความแปรปรวนจาก G-Study ($p \times i \times r$) ของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน

SOURCE OF VARIANCE	DF	SS	MS	ESTIMATED VARIANCE COMPONENT	% OF TOTAL VARIANCE
P	79	4302.213	54.458	3.996	31.45
I	4	793.697	198.424	1.150	9.05
R	1	14.634	14.634	0.033	0.26
PI	316	4478.208	14.171	6.706	52.77

SOURCE OF VARIANCE	DF	SS	MS	ESTIMATED VARIANCE COMPONENT	% OF TOTAL VARIANCE
PR	79	85.139	1.077	0.063	0.50
IR	4	3.603	0.900	0.001	0.00
PIR, e	316	239.542	0.758	0.758	5.97
TOTAL	799	9917.039		12.707	100%

2.3 ผลการศึกษาเพื่อการตัดสินใจการสรุปอ้างอิง (D-Study) ของคะแนนแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน ผลการศึกษา D (D-Study of p x i x r design) ค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ในการศึกษา D (D-Study) เมื่อจำนวนข้อสอบของแบบสอบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนที่มีจำนวนข้อสอบและจำนวนผู้ตรวจที่ต่างกัน คือ 5, 6 และ 10 ข้อ และจำนวนผู้ตรวจ 2 และ 3 คน พบว่า

แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนที่มีจำนวนข้อสอบ 5 ข้อ และจำนวนผู้ตรวจ 2 คน พบว่า สัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ (Relative Coefficient) เท่ากับ 0.733 และสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมบูรณ์ (Absolute Coefficient) เท่ากับ 0.702 และแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนที่มีจำนวนข้อสอบ 5 ข้อ และจำนวนผู้ตรวจ 3 คน พบว่า สัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ (Relative Coefficient) เท่ากับ 0.738 และสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมบูรณ์ (Absolute Coefficient) เท่ากับ 0.707

แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนที่มีจำนวนข้อสอบ 6 ข้อ และจำนวนผู้ตรวจ 2 คน พบว่า สัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ (Relative Coefficient) เท่ากับ 0.767 และสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมบูรณ์ (Absolute Coefficient) เท่ากับ 0.737 และแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนที่มีจำนวนข้อสอบ 6 ข้อ และจำนวนผู้ตรวจ 3 คน พบว่า สัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ (Relative Coefficient) เท่ากับ 0.771 และสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมบูรณ์ (Absolute Coefficient) เท่ากับ 0.742

แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนที่มีจำนวนข้อสอบ 10 ข้อ และจำนวนผู้ตรวจ 2 คน พบว่า สัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ (Relative Coefficient) เท่ากับ 0.843 และสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมบูรณ์ (Absolute Coefficient) เท่ากับ 0.820 และแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนที่มีจำนวนข้อสอบ 10 ข้อ และจำนวนผู้ตรวจ 3 คน พบว่า สัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ (Relative Coefficient) เท่ากับ 0.847 และสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมบูรณ์ (Absolute Coefficient) เท่ากับ 0.825 รายละเอียดดังตาราง 6

ตาราง 6 ผลการศึกษา D (D-Study of p x i x r design)

EFFECT	ESTIMATED VARIANCE COMPONENTS IN D-STUDY					
	$n_i=5$		$n_i=6$		$n_i=10$	
	$n_r=2$	$n_r=3$	$n_r=2$	$n_r=3$	$n_r=2$	$n_r=3$
$\sigma_p^2 = 3.996$	3.996	3.996	3.996	3.996	3.996	3.996
$\sigma_i^2 = 1.150$	0.230	0.230	0.191	0.191	0.115	0.115
$\sigma_r^2 = 0.033$	0.016	0.011	0.016	0.011	0.016	0.011
$\sigma_{pi}^2 = 6.706$	1.341	1.341	1.117	1.117	0.670	0.670
$\sigma_{pr}^2 = 0.063$	0.031	0.021	0.031	0.021	0.031	0.021
$\sigma_{ir}^2 = 0.001$	0.000	0.000	0.000	0.000	0.000	0.000
$\sigma_{pir,e}^2 = 0.758$	0.075	0.050	0.063	0.042	0.037	0.025
$\hat{\rho}_g^2$	0.733	0.738	0.767	0.771	0.843	0.847

ESTIMATED VARIANCE COMPONENTS IN D-STUDY

EFFECT	$n_i=5$		$n_i=6$		$n_i=10$	
	$n_r=2$	$n_r=3$	$n_r=2$	$n_r=3$	$n_r=2$	$n_r=3$
$\hat{\rho}_A^2$	0.702	0.707	0.737	0.742	0.820	0.825

อภิปรายผล

จากการศึกษาการเปรียบเทียบความเที่ยงของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนของนักเรียนชั้นประถมศึกษาปีที่ 6 : การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดสามารถสรุปและอภิปรายผล ดังนี้

1. แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน เรื่องเศษส่วน ที่พัฒนาขึ้น มีจำนวน 5 ข้อ โดยเนื้อหาประกอบด้วย (1) การเรียงลำดับเศษส่วน (2) การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละแบบไม่มีวงเล็บ (3) การบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละแบบไม่มีวงเล็บ (4) โจทย์ปัญหาการบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละ แบบ 2 ขั้นตอน และ (5) โจทย์ปัญหาการบวก ลบ คูณ หารระคน เศษส่วนและจำนวนคละ แบบ 3 ขั้นตอน ในแต่ละข้อประกอบด้วย ขั้นตอนการวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน จำนวน 5 ขั้นตอน ได้แก่ ขั้นที่ 1 การเขียนเพื่อระบุแนวคิดหรือหลักการที่ใช้ในการหาคำตอบ ขั้นที่ 2 การเขียนเพื่ออธิบายเหตุผลในการเลือกใช้แนวคิดหรือหลักการในการหาคำตอบ ขั้นที่ 3 การเขียนแสดงวิธีทำ ขั้นที่ 4 การเขียนผลลัพธ์ที่ได้จากการแสดงวิธีทำ และขั้นที่ 5 การเขียนสรุปคำตอบที่ถูกต้อง พบว่าค่าความเที่ยงที่ได้ทั้งฉบับสูงกว่า 0.70 ซึ่งสะท้อนให้เห็นว่า แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนที่พัฒนาขึ้นมีค่าความเที่ยงสูง (มีค่าอยู่ระหว่าง 0.70 – 1.00) การที่แบบวัดทักษะการสื่อสารฯ ฉบับนี้มีค่าความเที่ยงสูง เนื่องมาจากคุณภาพของเครื่องมือที่ได้รับการพัฒนาและตรวจสอบคุณภาพอย่างต่อเนื่อง เก็บข้อมูลจากตัวอย่างที่ครอบคลุมความสามารถ ความยาวของแบบวัดทักษะการสื่อสารที่มีความยาวพอเหมาะ ไม่มากเกินไปหรือน้อยเกินไป รวมถึงระยะเวลาในการทำข้อสอบที่กำหนดเวลาไว้อย่างเหมาะสม ซึ่งเป็นไปตามแนวคิดของ ศิริชัย กาญจนวาสี (2556) ที่ระบุว่า ปัจจัยที่มีผลต่อสัมประสิทธิ์ความเที่ยง ได้แก่ ความเป็นเอกพันธ์ของผู้สอบ ความยาวของแบบสอบ ความสัมพันธ์ระหว่างข้อสอบ เวลาที่ใช้ในการทำข้อสอบ และวิธีที่ใช้ในการประมาณค่าความเที่ยงสัมประสิทธิ์สหสัมพันธ์แบบค่าของคุณภาพผู้ตรวจอยู่ในช่วง 0.210 – 0.645 ค่าความยากอยู่ในช่วง 0.393 – 0.668 และค่าอำนาจจำแนกอยู่ในช่วง 0.379 – 0.882 ผลการวิจัยสะท้อนให้เห็นว่า แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนที่พัฒนาขึ้นมีค่าความยากง่ายและค่าอำนาจจำแนก อยู่ในช่วงตามเกณฑ์ที่กำหนด คือ สามารถวัดเด็กเก่งเด็กอ่อน และจำแนกผู้ที่มีการคิดเชิงคณิตศาสตร์ในระดับที่สูงกับต่ำได้ ซึ่งเป็นไปตามแนวคิดของ ล้วน สายยศ และ อังคณา สายยศ (2541) ที่ระบุว่า ค่าความยากง่ายควรมีค่าตั้งแต่ 0.2 – 0.8 และค่าอำนาจจำแนกมีค่าระหว่าง -1 ถึง +1 ถ้าคำนวณออกมาได้มีค่าบวกสูงแปลว่าอำนาจจำแนกดี แต่ถ้าติดลบ หรือ 0 ถือว่าอำนาจจำแนกใช้ไม่ได้ค่าอำนาจจำแนกที่ถือว่าจำแนกคนเก่งและอ่อนได้จะใช้ค่าอยู่ตั้งแต่ 0.20 ขึ้นไป อีกทั้งผลการวิจัยสอดคล้องกับงานวิจัยของสุรัชย์ มาตราและประภุติยา ทักษิณ (2565) ที่พบว่า แบบสอบอัตนัยวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของผู้ตรวจให้คะแนนภายใต้รูปแบบการตรวจให้คะแนนต่างกัน มีค่าความยากตั้งแต่ 0.22 – 0.69 และค่าอำนาจจำแนกตั้งแต่ 0.56 – 0.80 และยังสอดคล้องกับงานวิจัยของดร.ณิ อภัยกาวิ และคณะ (2563) ที่พบว่า แบบสอบอัตนัยการคิดทางคณิตศาสตร์ ค่าดัชนีความยากง่ายอยู่ระหว่าง 0.37 ถึง 0.67 ซึ่งทุกข้อมีค่าดัชนีอำนาจจำแนก อยู่ระหว่าง 0.27 ถึง 0.56

ผลการตรวจสอบคุณภาพของการตรวจให้คะแนนแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนที่พัฒนาขึ้น เป็นข้อสอบแบบเขียนวัดทักษะการสื่อสารทางคณิตศาสตร์ โดยพิจารณาจากความเที่ยงในการตรวจให้คะแนนจากผู้ตรวจจำนวน 2 คน (inter rater reliability) พบว่า มีค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจของผู้ตรวจจำนวน 2 คน แสดงว่าแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนที่พัฒนาขึ้น มีความเป็นปรนัยในการตรวจให้คะแนนของผู้ตรวจทั้ง 2 คน คะแนน

ที่ได้มีค่าใกล้เคียงกัน ผลการวิจัยสะท้อนให้เห็นว่า แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน ที่พัฒนาขึ้น เกณฑ์การให้คะแนนมีความชัดเจน มีรายละเอียดที่ทำให้ผู้ประเมินมีความเข้าใจที่ตรงกัน แนวคำตอบและเฉลยมีความชัดเจน ซึ่งเป็นไปตามแนวคิดของ Saal et al. (1980) ที่ระบุว่า 1) ผู้ให้คะแนนที่เชื่อถือได้จะเห็นด้วยกับการให้คะแนน "อย่างเป็นทางการ" ของประสิทธิภาพการทำงาน 2) ผู้ให้คะแนนที่เชื่อถือได้จะตกลงกันเองเกี่ยวกับคะแนนที่แน่นอนที่จะได้รับ 3) ผู้ให้คะแนนที่เชื่อถือได้เห็นด้วยว่า ประสิทธิภาพดีกว่าและประสิทธิภาพต่ำกว่า อีกทั้ง ผลการวิจัยสอดคล้องกับงานวิจัยของดุจฤดี ประเสริฐศิลป์พลมา และประกฤติยา ทักษิณ (2564) ที่พบว่า ผลการวิเคราะห์ความสอดคล้อง ด้วยค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intraclass Correlation Coefficient: ICC) โดยความเที่ยงในการตรวจให้คะแนน ระหว่างผู้ตรวจ (Inter Rater Reliability) มีความสอดคล้องกันในระดับดีมาก แสดงให้เห็นว่าผู้ตรวจให้คะแนนสอดคล้องไปในทิศทางเดียวกัน เนื่องมาจากมีการอธิบายการตรวจและเกณฑ์การตรวจให้คะแนน รวมถึงอธิบายตัวอย่างหรือแนวคำตอบในการตรวจให้คะแนน เพื่อลดความคลาดเคลื่อนอันเนื่องมาจากการให้คะแนน พร้อมทั้งให้คำแนะนำในการตรวจให้คะแนน

2. การเปรียบเทียบความเที่ยงของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน ภายใต้จำนวนข้อสอบที่ต่างกัน ที่ใช้จำนวนผู้ตรวจที่ต่างกัน ตรวจให้คะแนนทุกข้อของผู้สอบทุกคน โดยการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด ซึ่งผลการวิเคราะห์ความแปรปรวนขององค์ประกอบจากแหล่งความแปรปรวนต่างๆ ที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ของคะแนนแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน สรุปผลได้ว่า ความแปรปรวนของคะแนน ปฏิสัมพันธ์ระหว่างผู้สอบกับข้อคำถามจากในแต่ละข้อ ($\sigma^2_{pi} : o$) มีผลต่อสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) มากที่สุด (ร้อยละ 52.77) ผลการวิจัยสะท้อนให้เห็นว่า ปฏิสัมพันธ์ระหว่างผู้สอบกับข้อสอบ แสดงว่าความสามารถของนักเรียนมีความแตกต่างกันและมีความถนัดในข้อสอบแต่ละข้อต่างกัน และ (σ^2_{ir}) เป็นศูนย์ ซึ่งเป็นการจัดความแปรปรวนปฏิสัมพันธ์ระหว่างจำนวนข้อสอบกับผู้ตรวจได้อย่างสิ้นเชิง ซึ่งสะท้อนให้เห็นว่า ข้อสอบสามารถวัดได้ครอบคลุมกับพฤติกรรมที่ต้องการวัด เนื่องจากผู้ตรวจได้ตรวจให้คะแนนตามเกณฑ์ที่ผู้วิจัยกำหนดไว้ ซึ่งเป็นไปตามแนวคิดของ ศิริชัย กาญจนวาศี (2563) ที่ระบุว่าผลการประเมินในแต่ละแหล่งมีความสอดคล้องกันเนื่องมาจากค่าความแปรปรวนของผู้ประเมิน และยังสอดคล้องกับ วนิดา ภูเอี่ยม (2550) อ้างถึงใน ดุจฤดี ประเสริฐศิลป์พลมา และประกฤติยา ทักษิณ, (2564) ซึ่งกล่าวว่า ปฏิสัมพันธ์ระหว่างผู้สอบกับข้อสอบมีผลต่อนักเรียนเนื่องจากนักเรียนแต่ละคนมีความสามารถต่างกันทำให้เกิดความแปรปรวน และแหล่งที่ความแปรปรวนน้อยที่สุดคือ แหล่งปฏิสัมพันธ์ระหว่างข้อสอบและผู้ตรวจ (IR) อีกทั้ง ผลการวิจัยสอดคล้องกับงานวิจัยของดุจฤดี ประเสริฐศิลป์พลมา และประกฤติยา ทักษิณ (2564) ที่พบว่า การวิเคราะห์แหล่งความแปรปรวนของแหล่งความแปรปรวนต่าง ๆ ที่มีผลต่อการสรุปอ้างอิงความน่าเชื่อถือผลการประมาณค่าความแปรปรวนของทั้ง 3 องค์ประกอบ คือ ผู้สอบ (P) ข้อสอบ (I) และผู้ตรวจ (R) เมื่อใช้รูปแบบการตรวจข้อสอบทุกข้อของผู้สอบทุกคน [$P \times I \times R$] พบว่า แหล่งความแปรปรวนที่มากที่สุด คือ แหล่งปฏิสัมพันธ์ระหว่างผู้สอบกับข้อสอบ (PI) แสดงว่าความสามารถของนักเรียนมีความแตกต่างกันและมีความถนัดในข้อสอบแต่ละข้อต่างกัน

ผลการศึกษาค่าสัมประสิทธิ์ การสรุปอ้างอิง (G-Coefficient) ในการศึกษา D (D-Study) เป็นการศึกษาเพื่อตัดสินใจเลือกใช้แบบสอบที่สามารถนำไปปฏิบัติได้จริงในการสอบและมีความเที่ยงตามที่ต้องการ ซึ่งงานวิจัยในครั้งนี้ ผู้วิจัยเลือกค่าความเที่ยงอย่างน้อย 0.80 และมีจำนวนข้อสอบของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน คือ 10 ข้อ และมีจำนวนผู้ตรวจให้คะแนน 2 และ 3 คน โดยมีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพันธ์และเชิงสัมพันธ์ของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน มีค่าไม่ต่ำกว่า 0.80 เมื่อมีผู้ตรวจ จำนวน 2 และ 3 คน ตรวจให้คะแนน 10 ข้อขึ้นไป ผลการวิจัยสะท้อนให้เห็นว่า ถึงแม้ว่าผู้ตรวจข้อสอบทั้ง 10 ข้อจะให้ค่าความเที่ยงเกิน 0.80 แต่การที่ผู้ตรวจ 2 และ 3 คน โดยเฉพาะการตรวจ 3 คน ให้ค่าความเที่ยงมากกว่า เป็นเพราะยังจำนวนผู้ตรวจมาก

ค่าความเชื่อมั่นของแบบทดสอบจะมีค่าสูงขึ้น การที่ผู้ตรวจ 2 และ 3 คน ตรวจข้อสอบ 10 ข้อขึ้นไป ความเที่ยงสูงกว่าจำนวนข้อสอบ 5 และ 6 ข้อ เป็นเพราะแบบทดสอบที่มีจำนวนข้อมากจะวัดพฤติกรรมที่ต้องการวัดได้เพียงพอ คะแนนที่ได้จากการเดาจึงลดน้อยลง และผู้ตรวจแต่ละคนมีการตรวจให้คะแนนที่มีความสอดคล้องกัน มีความแตกต่างในการตรวจให้คะแนนค่อนข้างน้อย ซึ่งเป็นไปตามแนวคิดของ ศิริชัย กาญจนวาลี (2550) ที่ระบุว่า ในทุกรูปแบบการตรวจให้คะแนนที่เหมือนกันและจำนวนผู้ตรวจต่างกันมีผลทำให้ค่าความเชื่อมั่นแตกต่างกัน และถ้าใช้จำนวนผู้ตรวจหลายๆ คน ในการตรวจให้คะแนนผู้สอบแต่ละคน แล้วหาคะแนนเฉลี่ยแทนความสามารถของผู้สอบจะช่วยลดความแปรปรวนอันเนื่องมาจากผู้ตรวจและความคลาดเคลื่อนที่เหลือซึ่งจะทำให้ค่าความเชื่อมั่นสูงขึ้น จำนวนข้อสอบเพิ่มขึ้น จำนวนผู้ตรวจเพิ่มขึ้น ความเที่ยงจะเพิ่มขึ้น อีกทั้ง ผลการวิจัยสอดคล้องกับงานวิจัยของดุจฤดี ประเสริฐศิลป์พลมา และประภฤติยา ทักษิณ (2564) ที่พบว่า การศึกษาเพื่อการตัดสินใจสรุปอ้างอิง (D-Study) ของผลการตรวจ ด้วยรูปแบบการตรวจข้อสอบทุกข้อของผู้สอบทุกคน $[P \times I \times R]$ และ รูปแบบการตรวจให้คะแนนแบบการตรวจข้อสอบบางข้อของผู้สอบทุกคน $[P \times (I : R)]$ เมื่อมีผู้ตรวจให้คะแนน 2 และ 3 คนมีสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ (Relative Coefficient) และสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมบูรณ์ (Absolute Coefficient) เพิ่มขึ้นเมื่อผู้ตรวจเพิ่มขึ้น สอดคล้องกับผลการวิจัยของ ของดุจฤดี อภัยกาวิ และคณะ (2563) ที่พบว่า ค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับตัดสินใจเชิงสมบูรณ์ของผลการให้คะแนน มีค่าเพิ่มขึ้นเมื่อผู้ตรวจตั้งแต่ 3 คนหรือจำนวนข้อสอบเพิ่มขึ้น 3 ข้อขึ้นไป จะทำให้ผลการตรวจยังมีความเที่ยงเพิ่มมากขึ้น และสอดคล้องกับงานวิจัยของ อิศระ สระแก้ว (2565) พบว่า เมื่อความยาวของแบบทดสอบและจำนวนผู้ตรวจเพิ่มขึ้น ค่าความเชื่อมั่นก็จะสูงขึ้นตามลำดับ นอกจากนี้ ยังสอดคล้องกับงานวิจัยของ Kocaoğlu, S. & Şahin, M.G. (2024) พบว่า การเพิ่มจำนวนข้อและแบบทดสอบจะส่งผลดีต่อความน่าเชื่อถือ แต่การเพิ่มจำนวนแบบทดสอบจะส่งผลให้ความน่าเชื่อถือเพิ่มขึ้น

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. ครูผู้ใช้ควรศึกษาแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน และทำความเข้าใจเกี่ยวกับการบริหารการสอบ วิธีตรวจให้คะแนน ระยะเวลาในการสอบ โดยศึกษาคู่มือการใช้แบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน ด้านเศษส่วน สำหรับนักเรียนระดับชั้นประถมศึกษาปีที่ 6 ที่ผู้วิจัยสร้างขึ้น
2. ผลการศึกษาค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) เป็นการศึกษาเพื่อการตัดสินใจเลือกใช้แบบวัดที่สามารถนำไปใช้ได้จริงและมีความเที่ยงเท่ากับ 0.80 ซึ่งผู้วิจัยมีจำนวนข้อของแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียน คือ 10 ข้อขึ้นไป มีผู้ตรวจ 2 คน

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. เนื่องจากแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านการเขียนที่พัฒนาขึ้นในครั้งนี้ใช้กับนักเรียนระดับชั้นประถมศึกษาปีที่ 6 จึงควรมีการพัฒนาแบบวัดทักษะการสื่อสารทางคณิตศาสตร์ด้านอื่น สำหรับนักเรียนระดับชั้นประถมศึกษาปีที่ 6 เพื่อนำผลวิจัยไปใช้ในการพัฒนาและปรับปรุงคุณภาพการจัดการเรียนการสอนเรื่องเศษส่วนให้ดีขึ้นต่อไป
2. ควรศึกษาองค์ประกอบอื่นๆ ที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิง เช่น ศึกษาการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิง ของการวัดรูปแบบความสัมพันธ์องค์ประกอบแบบไขว้ (Nested) ในด้านจำนวนข้อและจำนวนผู้ตรวจให้คะแนนจากแหล่งอื่น $p \times (i : r)$

เอกสารอ้างอิง

- กระทรวงศึกษาธิการ. (2560). *ตัวชี้วัดและสาระการเรียนรู้แกนกลางกลุ่มสาระการเรียนรู้คณิตศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานพุทธศักราช 2551*. โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย จำกัด.
- กิตติมา สุรสนธิ. (2557). *ความรู้ทางการสื่อสาร* (พิมพ์ครั้งที่ 5). คณะวารสารศาสตร์และสื่อสารมวลชน มหาวิทยาลัยธรรมศาสตร์.
- ชนิสรา สงวนไว. (2558). การเปรียบเทียบความเที่ยงของแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด. (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ, มหาวิทยาลัยศิลปากร).
- ดร.ณิ อภัยกาวิ, กิตติศักดิ์ พันธุ์เกษม, & จิตติพงศ์ พุ่มพวง. (2563). การพัฒนาแบบวัดทักษะการคิดทางคณิตศาสตร์ระดับชั้นประถมศึกษา. *วารสารศึกษาศาสตร์มหาวิทยาลัยขอนแก่น*, 43(1), 115–132.
- ดุจฤดี ประเสริฐศิลป์พลมา, & ประกฤติยา ทักษิโน. (2564). ความเที่ยงของแบบวัดทักษะการคิดเชิงวิเคราะห์ทางคณิตศาสตร์. *วารสารศึกษาศาสตร์มหาวิทยาลัยขอนแก่น*, 44(2), 30–44.
- โรงเรียนสาธิตมหาวิทยาลัยรามคำแหง วิทยาเขตบางนา. (2567). *รายงานผลการดำเนินงานประจำปีการศึกษา 2567*. กรุงเทพฯ: โรงเรียนสาธิตมหาวิทยาลัยรามคำแหง วิทยาเขตบางนา.
- ล้วน สายยศ, & อังคณา สายยศ. (2541). *เทคนิคการวัดผลการเรียนรู้*. จุฬาลงกรณ์มหาวิทยาลัย.
- วนิดา ภู่อี่ยม. (2550). *การสร้างและการพัฒนาเครื่องมือวัดผลการศึกษา*. กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2555). *การวัดผลการศึกษา* (พิมพ์ครั้งที่ 8). โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- สุรัช มาตรา, & ประกฤติยา ทักษิโน. (2565). การพัฒนาแบบวัดทักษะการแก้ปัญหาทางคณิตศาสตร์ด้วยข้อสอบอัตนัย. *วารสารครุศาสตร์*, 49(2), 43–56.
- อัมพร ม้าคะนอง. (2559). การประยุกต์ใช้โมเดลทางคณิตศาสตร์เพื่อพัฒนาทักษะการสื่อสาร. *วารสารครุศาสตร์*, 48(4), 33–44.
- อิสระ สระแก้ว. (2565). ผลของการเพิ่มจำนวนข้อสอบและจำนวนผู้ตรวจต่อค่าความน่าเชื่อถือของผลการสอบ. *วารสารครุศาสตร์สาร*, 44(4), 77–90.
- Bachman, L. F., Davidson, F., Ryan, K., & Choi, I. C. (1995). *An investigation into the comparability of two tests of English as a foreign language*. Cambridge University Press.
- Bennett, N., & Lemoine, G. J. (2014). What a difference a word makes: Understanding threats to performance in a VUCA world. *Business Horizons*, 57(3), 311–317. <https://doi.org/10.1016/j.bushor.2014.01.001>
- Brennan, R. L., & Johnson, E. G. (1995). Generalizability theory and classical test theory. *Journal of Educational Measurement*, 32(4), 347–368. <https://doi.org/10.1111/j.1745-3984.1995.tb00473.x>
- Chui, M. M., & Wolfe, E. W. (2002). Essay scoring reliability: A review. *Journal of Educational Measurement*, 39(3), 243–270.
- Coffman, W. E. (1971). Essay examinations. In R. L. Thorndike (Ed.), *Educational Measurement* (2nd ed., pp. 271–302). American Council on Education.
- Dewi, N. K., Marsigit, M., & Sugiman, S. (2020). Mathematical communication ability: Case study on students' responses in mathematics learning. *International Journal of Instruction*, 13(2), 335–350.



- Hermawan, D., & Hidayat, W. (2018). Developing mathematical communication ability through learning model. *Journal of Physics: Conference Series*, 983(1), 012074.
- Kocaoğlu, S., & Şahin, M. G. (2024). The effect of rubric usage on assessment reliability in mathematics education. *Education and Science*, 49(220), 92–107.
- Linacre, J. M. (1993). Generalizability: Theory and practice. *Journal of Applied Measurement*, 4(4), 384–389.
- Lubis, A. H., Sari, D. F., & Siregar, E. (2023). Development of students' mathematical communication ability. *Journal of Education and Learning*, 12(1), 78–86.
- Lunz, M. E., Stahl, J. A., & Wright, B. D. (1990). Interrater reliability with judges rated by Rasch measurement. *Educational and Psychological Measurement*, 50(3), 703–709.
- Meel, D. E. (1999). Prospective secondary teachers' emerging knowledge of proof. *Issues in Mathematics Education*, 7, 67–101.
- Mehrens, W. A., & Lehmann, I. J. (1973). *Measurement and evaluation in education and psychology*. Holt, Rinehart and Winston.
- Meier, D. R., & Rishel, T. (1998). Writing to learn mathematics. *Mathematics Teaching in the Middle School*, 4(4), 236–239.
- Morgan, C. (2001). *Mathematical communication: Analyzing the language of mathematical learning and teaching*. Open University Press.
- National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. NCTM.
- Pantow, D., Ratu, D. D. L., & Poluakan, C. (2020). The influence of communication ability on mathematics learning achievement. *International Journal of Instruction*, 13(2), 121–138.
- Permatasari, P. (2021). Mathematical communication skills in problem-based learning. *Journal of Physics: Conference Series*, 1833(1), 012022.
- Saal, F. E., Downey, R. G., & Lahey, M. A. (1980). Rating the ratings: Assessing the psychometric quality of rating data. *Psychological Bulletin*, 88(2), 413–428.
- Smith, J. K., & Kulikowich, J. M. (2004). The effect of rater training on writing assessment. *Assessing Writing*, 9(2), 229–247.
- Sudweeks, R. R., Reeve, S., & Bradshaw, W. S. (2004). A comparison of rater training strategies in writing assessment. *Assessing Writing*, 9(3), 227–242.
- Swartz, C. W., Norman, K. L., & Penfield, J. (1999). Reliability of scores from essay tests: Interrater agreement and the factors that influence it. *Educational and Psychological Measurement*, 59(3), 451–466.
- Teledahl, A. (2017). Written mathematical communication: Analyzing students' texts in mathematics. *Nordic Studies in Mathematics Education*, 22(1), 21–44.
- Turner, R. C. (2003). A model for measuring writing assessment scoring reliability. *Assessing Writing*, 8(2), 147–166.