



พัฒนาแบบทดสอบวินิจฉัยสามระดับเพื่อประเมิน

การ

มโนทัศน์ที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วน ของนักเรียนชั้นประถมศึกษาปีที่ 6 ด้วยแนวคิดทฤษฎี การตอบสนองข้อสอบ (IRT)

ตรงมกล พรหมอ่วม^{1*} ปัญญวิชญ์ ใบกุหลาบ²

รับบทความ: 8 เมษายน 2569 แก้ไขบทความ: 25 พฤษภาคม 2569 ตอรับบทความ: 29 พฤษภาคม 2569

บทคัดย่อ

การวิจัยครั้งนี้เป็นการวิจัยและพัฒนา มีวัตถุประสงค์เพื่อสร้างและหาคุณภาพของแบบทดสอบวินิจฉัยสามระดับเพื่อประเมินมโนทัศน์ที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6 และศึกษามโนทัศน์ที่คลาดเคลื่อน ประชากร คือ นักเรียนชั้นประถมศึกษาปีที่ 6 สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาพิษณุโลก เขต 1 จำนวน 2,264 คน 1) กลุ่มตัวอย่างที่สุ่มแบบอย่างง่าย ซึ่งแบ่งเป็นกลุ่มวิเคราะห์คุณภาพตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม 30 คน ผลการวิเคราะห์พบว่า มีค่าความยาก (p) ระหว่าง 0.47 ถึง 0.73 ค่าอำนาจจำแนก (r) ระหว่าง 0.40 ถึง 0.60 และค่าความเชื่อมั่นทั้งฉบับ ($KR-20$) เท่ากับ 0.85 จำนวน 18 ข้อ โดยจากข้อสอบ 18 ข้อได้พัฒนาเป็นแบบทดสอบวินิจฉัยสามระดับ และปรับลดเหลือ 15 ข้อตามคำแนะนำของผู้เชี่ยวชาญเพื่อให้เหมาะสมกับเวลาสอบ และ 2) กลุ่มตัวอย่างสุ่มแบบหลายขั้นตอน เป็นกลุ่มวิเคราะห์คุณภาพตามแนวคิดทฤษฎีการตอบสนองข้อสอบ 306 คน เครื่องมือคือแบบทดสอบวินิจฉัยสามระดับ 15 ข้อ ผลการวิจัยพบว่าแบบทดสอบมีความเป็นมิติเดียว โดยองค์ประกอบหลักแรกอธิบายความแปรปรวนได้ร้อยละ 36.40 และมีความเป็นอิสระเชิงท้องถิ่น ผลการวิเคราะห์ด้วยโมเดล 2PLM ข้อสอบทุกข้อมีค่าอำนาจจำแนก (a) ระหว่าง 1.81 ถึง 3.94 ผ่านเกณฑ์สากล ค่าความยาก (b) ระหว่าง -1.30 ถึง 0.50 ฟังก์ชันสารสนเทศให้สารสนเทศสูงสุดเท่ากับ 21 ณ ระดับความสามารถระหว่าง (θ) 0.0 ถึง 0.5 แสดงว่าเครื่องมือมีความแม่นยำสูงในการวัดผู้เรียนกลุ่มความสามารถปานกลาง ผลการศึกษามโนทัศน์ พบว่า กลุ่มมโนทัศน์คลาดเคลื่อนเป็นร้อยละ 22.18 เมื่อจำแนกตามกระบวนการแก้โจทย์ปัญหา จะได้ ด้านการตีความโจทย์ปัญหา ด้านการตรวจสอบการแก้ปัญห ด้านการใช้ทฤษฎีบทหรือนิยาม และด้านการคิดคำนวณ ตามลำดับ

คำสำคัญ: แบบทดสอบวินิจฉัยสามระดับ มโนทัศน์ที่คลาดเคลื่อน โจทย์ปัญหาเศษส่วน ทฤษฎีการตอบสนองข้อสอบ (IRT)

¹⁻² มหาวิทยาลัยราชภัฏพิบูลสงคราม

* อีเมล: tpromouam8@gmail.com

Development of a Three-Level Diagnostic Test to Assess Misconceptions in Solving Fraction Word Problems of Sixth-Grade Students Using Item Response Theory (IRT)

Trongkamon Promouam ^{1*}
Punnawit Baikularb ²

Abstract

This research and development aimed to construct and validate a three-tier diagnostic test for assessing misconceptions in solving fraction word problems among sixth-grade students, and to examine their specific misconceptions. The population consisted of 2,264 sixth-grade students under the Phitsanulok Primary Educational Service Area Office 1. The sample was selected using two sampling methods 1) a sample of 30 students obtained via simple random sampling for quality analysis based on Classical Test Theory (CTT); the results showed that 18 items met the criteria, with difficulty (p) ranging from 0.47 to 0.73, discrimination (r) from 0.40 to 0.60, and a total-test reliability (KR-20) of 0.85; these 18 items were then developed into a three-tier diagnostic test and subsequently reduced to 15 items based on expert recommendations to fit the testing time limit; and 2) a sample of 306 students obtained via multi-stage random sampling for quality analysis based on Item Response Theory (IRT). The research instrument was the 15-item three-tier diagnostic test. The research findings revealed that the test was unidimensional, with the first principal component accounting for 36.40% of the total variance, and demonstrated local independence. Based on the two-parameter logistic model (2PLM) analysis, all test items met international standards, with discrimination parameters (a) ranging from 1.81 to 3.94, and difficulty parameters (b) from -1.30 to 0.50. The test information function provided the maximum information value of 21 at ability levels (θ) between 0.0 and 0.5, indicating high measurement precision for students with moderate ability levels. Regarding the investigation of misconceptions, it was found that 22.18% of the students held misconceptions. When classified by word problem-solving processes, the highest prevalence of misconceptions was found in problem interpretation, followed by problem-solving verification, application of theorems or definitions, and computation, respectively.

Keywords: Three-Level Diagnostic Test, Misconceptions, Fraction Word Problems, Item Response Theory (IRT)

¹⁻² Pibulsongkram Rajabhat University

* Email: tpromouam8@gmail.com

บทนำ

คณิตศาสตร์เป็นวิชาที่มีบทบาทสำคัญต่อการพัฒนาทักษะการคิดอย่างเป็นระบบและการแก้ปัญหาในชีวิตประจำวัน โดยเนื้อหาเรื่องเศษส่วน เป็นพื้นฐานสำคัญสำหรับใช้ในวิชาคณิตศาสตร์ที่ซับซ้อนขึ้นผู้วิจัยจึงทำการศึกษาเพื่อหารูปแบบของแบบทดสอบที่ช่วยชี้ให้เห็นชัดถึงเรื่องที่นักเรียนเกิดมโนทัศน์ที่คลาดเคลื่อนในด้านใดด้านหนึ่งได้ชัดเจน เพื่อช่วยให้ผู้สอนปรับปรุง ช่อมเสริมความรู้ให้ผู้เรียนได้ตรงประเด็นมากยิ่งขึ้น

จากที่กล่าวมาข้างต้น พบว่า แบบการทดสอบที่ช่วยให้ผู้สอนทราบถึงสิ่งที่ผู้เรียนเข้าใจคลาดเคลื่อนควรใช้แบบทดสอบที่สามารถทำให้ทราบถึงเรื่องที่ผู้เรียนไม่เข้าใจได้อย่างตรงประเด็น ซึ่งพบว่าแบบทดสอบวินิจฉัยสามารถนั้นสามารถทำได้ และจากการศึกษามาตรฐาน ตัวชี้วัดรายวิชาคณิตศาสตร์ หนังสือเรียนรายวิชาพื้นฐานคณิตศาสตร์ ป.5 - ป.6 และหนังสือแบบฝึกหัดรายวิชาพื้นฐานคณิตศาสตร์ ป.5 - ป.6 ผู้วิจัยจึงแบ่งเนื้อหาการแก้โจทย์ปัญหาเศษส่วนออกเป็น 3 เรื่อง ได้แก่ 1) การแก้โจทย์ปัญหาที่ใช้เศษส่วนเป็นค่าของข้อมูล 2) การแก้โจทย์ปัญหาเศษส่วนที่มีการแก้แบบหนึ่งขั้นตอน 3) การแก้โจทย์ปัญหาเศษส่วนที่มีการแก้แบบสองถึงสามขั้นตอน ซึ่งเป็นแบบทดสอบที่มีรูปแบบการตอบในแต่ละข้อแบ่งออกเป็น 3 ระดับ อันได้แก่ 1) ระดับที่ 1 ระดับคำตอบเป็นข้อสอบแบบตัวเลือก 3 ตัวเลือก โดยแต่ละตัวเลือกจะมีการคำนวณค่าความคลาดเคลื่อนของมโนทัศน์ที่คลาดเคลื่อนในระดับที่ต่างกันเอาไว้ในแต่ละข้อ 2) ระดับที่ 2 ระดับเหตุผล เป็นข้อสอบแบบตัวเลือก 3 ตัวเลือก ซึ่งเป็นเหตุผลที่สอดคล้องกับแนวความคิด กระบวนการคิด หรือวิธีการคิดที่ทำให้เกิดคำตอบในระดับที่ 1 3) ระดับที่ 3 เป็นระดับความมั่นใจหรือระดับความเชื่อมั่นในการตอบคำตอบของระดับที่ 1 และระดับที่ 2 เป็นการแสดงความมั่นใจในการตอบของนักเรียนคือมั่นใจ และไม่มั่นใจ ซึ่งผู้วิจัยเลือกใช้รูปแบบการตรวจให้คะแนนแบบค่าคู่ (Dichotomous Scoring) คือ 1 และ 0 คะแนน โดยพิจารณาจากความสัมพันธ์ของการตอบในระดับที่ 1 และระดับที่ 2 เพื่อให้สามารถวินิจฉัยผู้ที่มีมโนทัศน์ที่ถูกต้องได้อย่างแม่นยำ

เนื่องจากการสร้างแบบทดสอบวินิจฉัยสามารถ จะเริ่มจากการสร้างแบบทดสอบและหาคุณภาพของแบบทดสอบปกติก่อนผู้วิจัยจึงได้ศึกษาเกี่ยวกับการวิเคราะห์ โดยใช้ทฤษฎีการทดสอบแบบมาตรฐานเดิม (Classical Test Theory: CTT) ซึ่งเป็นเครื่องมือวัดความรู้ของผู้ทดสอบที่ใช้อย่างแพร่หลาย โดยเพื่อให้การทดสอบมีความแม่นยำและความเชื่อมั่นเพียงพอ ข้อสอบจะถูกเรียงลำดับไว้ตามเนื้อหาที่ต้องการทดสอบและมักจะประกอบด้วยข้อสอบที่มีความยากระดับปานกลางถึงยากเพื่อให้เหมาะสมกับความสามารถโดยรวมของผู้สอบ (ทิพยเกสร พรหมภักดี, 2560) เมื่อสร้างแบบทดสอบข้อที่มีคุณภาพเหมาะสมแล้ว จึงนำแบบทดสอบดังกล่าวมาयरระดับเป็นแบบวินิจฉัยสามารถ และวิเคราะห์โดยใช้โมเดลทฤษฎีการตอบสนองข้อสอบ ซึ่งมีข้อตกลงเบื้องต้นสำคัญ 3 ประการคือ 1) ความเป็นเอกมิติ (unidimensionality) 2) ความเป็นอิสระ (local independence) และ 3) การเป็นฟังก์ชันต่อเนื่อง (continuous function) สามารถอธิบายความสัมพันธ์ระหว่างคุณลักษณะแฝงกับความน่าจะเป็นในการตอบข้อสอบได้อย่างถูกต้อง (ศิริชัย กาญจนวาสี, 2555) ทั้งนี้โมเดล IRT มีการพัฒนาขึ้นมาหลายรูปแบบอย่างไรก็ตามโดยทั่วไปโมเดล IRT จะแบ่งเป็น 3 รูปแบบหลักได้แก่ 1) โมเดล IRT

ที่มีการตรวจให้คะแนนสองค่า (dichotomous scoring IRT) 2) โมเดล IRT ที่มีการตรวจให้คะแนนหลายค่า (polytomous scoring IRT) และ 3) โมเดล IRT ที่มีการตรวจให้คะแนนแบบต่อเนื่อง (continuous scoring IRT) เนื่องจากผู้วิจัยต้องการพิจารณาจากความสัมพันธ์ของการตอบทั้ง 3 ระดับ ได้แก่ คำตอบ เหตุผล และความมั่นใจ ผู้วิจัยเลือกใช้โมเดล IRT ที่มีการตรวจให้คะแนนสองค่า (Dichotomous Scoring IRT) แล้วได้ทำการรวมผลการตอบจากระดับที่ 1 และ 2 เพื่อจำแนกสถานะความรู้ของผู้เรียนเป็นรอบรู้จริง หรือ เข้าใจคลาดเคลื่อนในการวิเคราะห์คุณภาพข้อสอบ งานวิจัยนี้เลือกใช้โมเดลโลจิสติกสองพารามิเตอร์ (Two-parameter logistic model : 2PLM) เพื่อให้สอดคล้องกับลักษณะข้อมูลและขนาดกลุ่มตัวอย่าง ทั้งนี้ รายละเอียดเชิงเทคนิคของโมเดลจะนำเสนอไว้ในส่วนนิยามศัพท์เฉพาะและระเบียบวิธีวิจัยต่อไป โดยคาดหวังว่าผลการวิจัยจะนำไปสู่การพัฒนาความเข้าใจที่ถูกต้องของผู้เรียน และส่งเสริมการพัฒนาการเรียนรู้ในระดับที่สูงขึ้นอย่างยั่งยืน

การวิจัยครั้งนี้เป็นการวิจัยและพัฒนา (Research and Development) ที่มุ่งเน้นการสร้างและตรวจสอบคุณภาพของแบบทดสอบวินิจฉัยสามระดับ โดยดำเนินการพัฒนาเครื่องมืออย่างเป็นระบบตั้งแต่การสร้างข้อสอบ การตรวจสอบความตรงเชิงเนื้อหา การทดลองใช้ตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม และการวิเคราะห์คุณภาพข้อสอบด้วยทฤษฎีการตอบสนองข้อสอบ เพื่อให้ได้เครื่องมือที่มีคุณภาพ สามารถใช้วินิจฉัยมโนทัศน์ที่คลาดเคลื่อนได้อย่างแม่นยำ

วัตถุประสงค์การวิจัย

1. เพื่อสร้างและหาคุณภาพของแบบทดสอบวินิจฉัยสามระดับสำหรับประเมินโน้ตทัศน์ที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6
2. เพื่อศึกษามโนทัศน์ที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6 ใน 4 ด้าน ได้แก่ ด้านการตีความ ด้านการใช้ทฤษฎีบทหรือนิยามด้านการคิดคำนวณ และด้านการตรวจสอบการแก้ปัญหา โดยใช้แบบวินิจฉัยสามระดับที่พัฒนาขึ้น

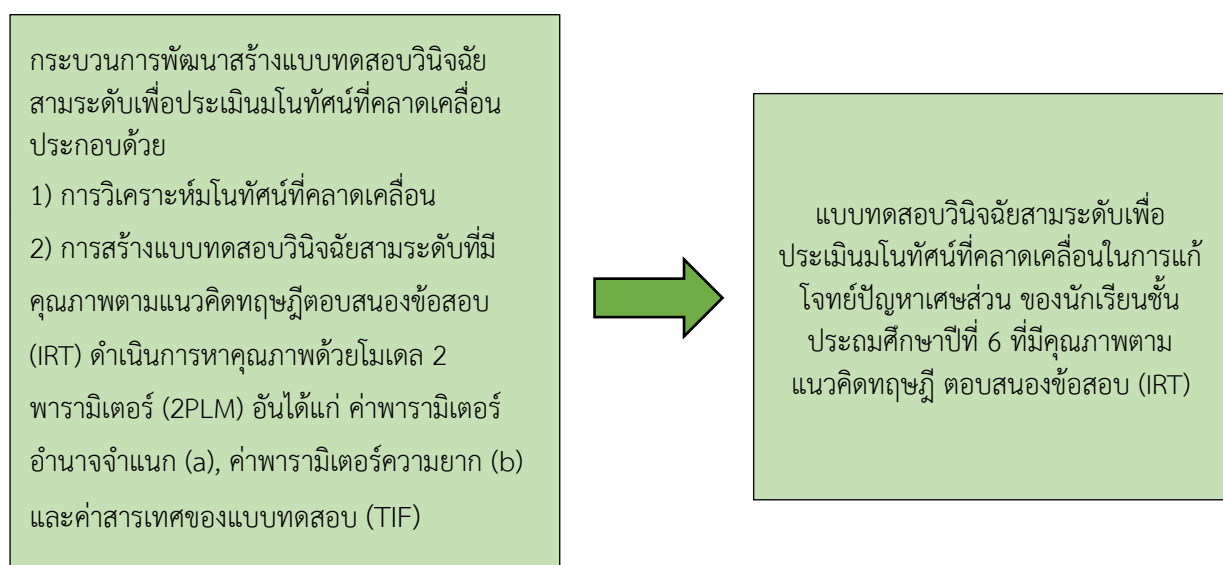
วรรณกรรมและทฤษฎีที่เกี่ยวข้อง

ความสำคัญของมโนทัศน์ที่คลาดเคลื่อนทางคณิตศาสตร์ คณิตศาสตร์มีบทบาทสำคัญในการพัฒนาการคิดอย่างเป็นระบบและการแก้ปัญหาในชีวิตจริง, อย่างไรก็ตาม เรื่อง เศษส่วน เป็นเนื้อหาที่มีความเป็นนามธรรมสูงและซับซ้อนสำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ซึ่งมักนำไปสู่การเกิดมโนทัศน์ที่คลาดเคลื่อน (Misconceptions), โดยนิยามของมโนทัศน์ที่คลาดเคลื่อนคือ ความเข้าใจหรือความเชื่อที่ผิดแผกไปจากหลักการทางวิชาการที่ยอมรับกันในสังคม ซึ่งเกิดขึ้นอย่างเป็นระบบและฝังลึกในโครงสร้างความคิดของผู้เรียนจากการสังเคราะห์งานวิจัยพบว่าความคลาดเคลื่อนนี้ปรากฏใน 4 ด้านหลัก ได้แก่ ด้านการตีความจากโจทย์ ด้านการใช้ทฤษฎีบท สูตร นิยาม ด้านการคิดคำนวณ และด้านการตรวจสอบ การแก้ปัญหา

แบบทดสอบวินิจฉัยสามระดับ (Three-Tier Diagnostic Test) การวินิจฉัยจุดบกพร่องรายบุคคลมีความสำคัญต่อการจัดการเรียนรู้ซ่อมเสริม, ผู้วิจัยจึงเลือกพัฒนา แบบทดสอบวินิจฉัยสามระดับ เพื่อก้าวข้ามข้อจำกัดของแบบทดสอบปกติและแบบทดสอบสองระดับที่ไม่สามารถแยกแยะระหว่างผู้ที่มีมีโนทัศน์คลาดเคลื่อน กับผู้ที่ไม่มีความรู้ (เดา) ได้ชัดเจน โดยโครงสร้างของข้อสอบประกอบด้วย 3 ส่วน ได้แก่ 1) ระดับที่ 1 (Content Tier) ประเมินความรู้เชิงเนื้อหา 2) ระดับที่ 2 (Reason Tier) ประเมินเหตุผลเบื้องหลังการเลือกคำตอบ โดยใช้ตัวลงเชิงวินิจฉัยที่สร้างจากคำตอบผิดจริงของนักเรียน และ 3) ระดับที่ 3 (Confidence Tier) ประเมินความเชื่อมั่นในการตอบเพื่อยืนยันว่านักเรียนมีความเข้าใจผิดที่ฝังลึกหรือเพียงแค่เดาคำตอบ

การวิเคราะห์คุณภาพด้วยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT) เพื่อให้ได้เครื่องมือที่มีมาตรฐานสากล การวิจัยนี้ประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) ซึ่งมีข้อดีคือค่าพารามิเตอร์ของข้อสอบจะเป็นอิสระจากกลุ่มผู้สอบ โดยใช้โมเดลโลจิสติก 2 พารามิเตอร์ (2-Parameter Logistic Model : 2PLM) ในการวิเคราะห์ค่าคุณภาพ ได้แก่ ค่าอำนาจจำแนก (a) คือ ความชันของเส้นโค้งลักษณะข้อสอบที่บ่งบอกประสิทธิภาพในการจำแนกผู้สอบ ค่าความยาก (b) คือ ระดับความสามารถที่ทำให้มีโอกาสตอบถูกร้อยละ 50 นอกจากนี้ยังมีการวิเคราะห์ตามทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) เพื่อหาค่าความตรงเชิงเนื้อหา (IOC) ค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) เพื่อเป็นการตรวจสอบคุณภาพเบื้องต้น

กรอบแนวคิดการวิจัย



ภาพที่ 1 กรอบแนวคิดการวิจัย

วิธีการดำเนินการวิจัย

ประชากรและผู้ให้ข้อมูลสำคัญ

ประชากรที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ในโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษาพิษณุโลก เขต 1 กระทรวงศึกษาธิการ ซึ่งผู้วิจัยได้ทำการสำรวจในปีการศึกษา 2568 พบว่ามีประชากรทั้งหมด 2,264 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ในโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษาพิษณุโลก เขต 1 กระทรวงศึกษาธิการ โดยผู้วิจัยดำเนินการสุ่มกลุ่มตัวอย่าง 2 กลุ่ม ดังนี้

กลุ่มตัวอย่างที่ 1 กลุ่มตัวอย่างสำหรับการวิเคราะห์คุณภาพตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม (Classical Test Theory: CTT) ได้แก่ นักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 30 คน ซึ่งได้มาโดยวิธีการสุ่มอย่างง่าย (Simple Random Sampling) ซึ่งมีวัตถุประสงค์เพื่อนำข้อมูลผลการสอบมาใช้วิเคราะห์ หาค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) เป็นรายข้อในขั้นการทดลองใช้เครื่องมือ (Try-out) เพื่อปรับปรุงโครงสร้างแบบทดสอบเบื้องต้น

กลุ่มตัวอย่างที่ 2 กลุ่มตัวอย่างสำหรับการวิเคราะห์คุณภาพตามแนวคิดทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ได้แก่ นักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 306 คน ซึ่งได้มาโดยการสุ่มแบบหลายขั้นตอน (Multi-stage Sampling) มีวัตถุประสงค์เพื่อนำผลการตอบแบบทดสอบวินิจัยสามระดับมาใช้ตรวจสอบข้อตกลงเบื้องต้นและประมาณค่าพารามิเตอร์รายข้อด้วยโมเดลโลจิสติกส์แบบ 2 พารามิเตอร์ (2PLM)

การรวบรวมข้อมูล เครื่องมือและการวิเคราะห์ข้อมูล

การพัฒนาเครื่องมือที่ใช้ในการวิจัย สำหรับเครื่องมือที่ใช้ในงานวิจัยครั้งนี้มี 1 ฉบับ 2 รูปแบบ คือ 1) รูปแบบที่ 1 เป็นแบบทดสอบโดยการวิเคราะห์ตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) เรื่อง โจทย์ปัญหาเศษส่วน จำนวน 1 ฉบับ และ 2) รูปแบบที่ 2 แบบทดสอบวินิจัยสามระดับเพื่อประเมินโน้ตที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วน ของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยการวิเคราะห์ตามแนวคิดทฤษฎีการตอบสนองข้อสอบ (IRT) ที่ถูกพัฒนามาจากแบบทดสอบโดยวิเคราะห์ตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) โดยผู้วิจัยเป็นผู้สร้างขึ้นเอง มีรายละเอียดขั้นตอนการสร้าง ดังนี้

1. ศึกษาวิธีสร้างแบบทดสอบโดยวิเคราะห์ตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) เรื่อง โจทย์ปัญหาเศษส่วน จากตำรา เอกสาร และงานวิจัยที่เกี่ยวข้อง

2. ศึกษาหลักสูตร คู่มือครู หนังสือเรียนสาระการเรียนรู้พื้นฐาน คณิตศาสตร์ เล่ม 1 กลุ่มสาระการเรียนรู้คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ตามหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ฉบับปรับปรุง 2560 (สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ, 2560) เพื่อเป็นแนวทางวางแผนการสร้างแบบทดสอบโดยวิเคราะห์ตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) เรื่อง โจทย์ปัญหาเศษส่วน

3. วิเคราะห์เนื้อหาเรื่องโจทย์ปัญหาเศษส่วน ของนักเรียนชั้นประถมศึกษาปีที่ 6 ซึ่งผู้วิจัยสังเคราะห์รูปแบบโจทย์ปัญหาเศษส่วนออกมาได้ 3 ลักษณะ ได้แก่ ลักษณะที่ 1 โจทย์ปัญหาที่ใช้เศษส่วน เป็นตัวบอกค่าของข้อมูลที่ใช้ถาม ลักษณะที่ 2 เป็นโจทย์ปัญหาแบบ 1 ขั้นตอน ลักษณะที่ 3 เป็นโจทย์ปัญหาแบบ 2-3 ขั้นตอน

4. ศึกษาโน้ตค้นที่คลาดเคลื่อนเกี่ยวกับวิชาคณิตศาสตร์จากการสังเคราะห์งานวิจัยของ เวชฤทธิ์ อังชนะภัทรขจร (2546) ที่สังเคราะห์งานวิจัยเกี่ยวข้องกับโน้ตค้นที่คลาดเคลื่อนในวิชาคณิตศาสตร์ซึ่งสรุปได้ว่าโน้ตค้นที่คลาดเคลื่อน แบ่งออกเป็น 4 ด้านดังนี้ 4.1) ด้านการตีความจากโจทย์ 4.2) ด้านการใช้ทฤษฎีบท สูตร กฎ บทนิยาม และสมบัติ 4.3) ด้านการคิดคำนวณ 4.4) ด้านการตรวจสอบการแก้ปัญหา

5. วิเคราะห์ด้านทั้ง 4 ของมโนทัศน์ที่คลาดเคลื่อน แล้วสร้างข้อคำถามที่สอดคล้องกับเนื้อหาเรื่องโจทย์ปัญหาเศษส่วน ของนักเรียนชั้นประถมศึกษาปีที่ 6 ที่มีความเกี่ยวข้องกับมโนทัศน์ที่คลาดเคลื่อนทั้ง 4 ด้าน

6. สร้างแบบทดสอบที่สอดคล้องตามการวิเคราะห์ที่ดำเนินการมาแล้ว

7. นำแบบทดสอบโดยวิเคราะห์ตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) ที่ผู้วิจัยสร้างขึ้นให้อาจารย์ที่ปรึกษาตรวจสอบและให้คำปรึกษา เพื่อแก้ไขปรับปรุงให้ถูกต้องแล้วนำไปให้ผู้ทรงคุณวุฒิจำนวน 3 ท่าน เพื่อพิจารณาตรวจสอบความเที่ยงตรงเชิงเนื้อหา โดยดัชนีความสอดคล้อง (IOC) ได้ผลว่า จากข้อสอบ 30 ข้อ มีข้อสอบ 27 มีค่า IOC ระหว่าง 0.67 ถึง 1.00 ถือว่าผ่านเกณฑ์

8. นำแบบทดสอบโดยวิเคราะห์ตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) ที่ปรับปรุงแก้ไขแล้วไปใช้กับนักเรียนชั้นประถมศึกษาปีที่ 6 ที่เป็นกลุ่มตัวอย่างที่ได้จากการสุ่มอย่างง่าย 30 คน

9. นำผลการทดสอบมาวิเคราะห์ข้อมูลโดยวิเคราะห์ตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) เพื่อหาค่าความยาก ค่าอำนาจจำแนก และความเชื่อมั่น

10. นำแบบทดสอบโดยวิเคราะห์ตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) มายกระดับเป็นแบบทดสอบวินิจฉัยสามระดับ โดยออกแบบตัวเลือกในระดับที่ 2 ระดับเหตุผล ให้เป็นตัวลวงเชิงวินิจฉัย (Diagnostic Distractors) ที่สะท้อนมโนทัศน์คลาดเคลื่อนใน 4 ด้าน ได้แก่ 1) ด้านการตีความโจทย์ 2) ด้านการใช้ทฤษฎีบท สูตร กฎ บทนิยาม และสมบัติ 3) ด้านการคิดคำนวณ และ 4) ด้านการตรวจสอบการแก้ปัญหา และระดับที่ 3 เป็นระดับความมั่นใจ มีคำตอบคือ มั่นใจ และไม่มั่นใจ

11. นำแบบทดสอบวินิจฉัยสามระดับที่พัฒนาขึ้น เสนอต่อผู้ทรงคุณวุฒิจำนวน 5 ท่าน เพื่อประเมินความตรงเชิงเนื้อหา (Content Validity) ความเหมาะสมของภาษา และความสอดคล้องของตัวลวงเชิงวินิจฉัยกับมโนทัศน์ที่คลาดเคลื่อนทั้ง 4 ด้าน โดยพิจารณาจากค่าดัชนีความสอดคล้อง (Index of Item-objective Congruence : IOC)

12. การนำแบบทดสอบวินิจฉัยสามระดับฉบับสมบูรณ์ ไปทดสอบกับกลุ่มตัวอย่าง 306 คน ซึ่งแบบทดสอบวินิจฉัยสามระดับสำหรับประเมินโน้ตสัที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6 เรื่องโจทย์ปัญหาเศษส่วน

13. นำผลการทดสอบทั้งหมดมาหาผลการวิเคราะห์ตามแนวคิดทฤษฎีตอบสนองข้อสอบ (IRT) เพื่อวิเคราะห์หาค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าสารสนเทศของแบบทดสอบ (TIF) และนำผลคำตอบที่ได้ไปแปลผลเพื่อจัดกลุ่มโน้ตสัที่คลาดเคลื่อนตามด้านทั้ง 4 ด้านต่อไป

การเก็บรวบรวมข้อมูล

1. ประสานโรงเรียนที่จะเก็บข้อมูลเพื่อส่งเอกสารขอความอนุเคราะห์เก็บรวบรวมข้อมูล
2. ดำเนินการเก็บข้อมูลด้วยตนเอง โดยเป็นผู้ชี้แจงการใช้เครื่องมือ และรวบรวมข้อมูลด้วยตนเอง ร่วมกันครูประจำวิชา

การวิเคราะห์ข้อมูลและการใช้สถิติ

1. การวิเคราะห์ข้อมูลเชิงคุณภาพ ผู้วิจัยนำข้อมูลที่ได้จากการทำแบบทดสอบโดยวิเคราะห์ตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) และแบบทดสอบวินิจฉัยสามระดับสำหรับประเมินโน้ตสัที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยใช้การวิเคราะห์ตามแนวคิดทฤษฎีตอบสนองข้อสอบ (IRT) ของกลุ่มตัวอย่างมาวิเคราะห์ ดังต่อไปนี้ 1.1) ผู้วิจัยตรวจสอบแบบทดสอบโดยวิเคราะห์ตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) แล้วนำผลการทดสอบมาวิเคราะห์หาคุณภาพของเครื่องมือ โดยวิเคราะห์ค่าความยากง่าย (p) และค่าอำนาจจำแนก (r) เป็นรายข้อ และหาค่าความเชื่อมั่น (Reliability) ของแบบทดสอบทั้งฉบับ โดยใช้สูตร Kuder-Richardson (KR-20) 1.2) ผู้วิจัยตรวจสอบแบบทดสอบวินิจฉัยสามระดับสำหรับประเมินโน้ตสัที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยใช้การวิเคราะห์ตามแนวคิดทฤษฎีตอบสนองข้อสอบ (IRT) 1.3) ผู้วิจัยกำหนดเกณฑ์การให้คะแนนแบบแยกตามระดับ เพื่อให้สามารถจำแนกโน้ตสัของนักเรียนได้อย่างละเอียด ดังนี้ 1.3.1) การให้รหัสคะแนนรายระดับในแต่ละข้อคำถาม นักเรียนจะได้รับรหัสคะแนนในแต่ละระดับ ดังนี้ระดับที่ 1 (การเลือกคำตอบ) ตอบถูกรหัสคะแนน 1 ตอบผิดรหัสคะแนน 0 ระดับที่ 2 (การเลือกเหตุผล) เลือกเหตุผลถูกต้องรหัสคะแนน 1 เลือกเหตุผลผิดแต่สอดคล้องกับระดับคำตอบรหัสคะแนน 0* เลือกเหตุผลผิดและไม่สอดคล้องกับระดับคำตอบรหัสคะแนน 0 ระดับที่ 3 (ความมั่นใจ) ระบุว่ามั่นใจ รหัสคะแนน 1 ระบุว่าไม่มั่นใจ รหัสคะแนน 0 คะแนน 1.3.2) การให้คะแนนเพื่อการวิเคราะห์คุณภาพข้อสอบ (Scoring for IRT) เนื่องจากโมเดลโลจิสติก 2 พารามิเตอร์ (2PLM) กำหนดให้ใช้คะแนนค่าฐานสอง (Dichotomous Scoring) ผู้วิจัยจึงกำหนดเกณฑ์การแปลผลคะแนนดังนี้ได้ 1 คะแนน เมื่อนักเรียนตอบได้รหัสคะแนนทั้งระดับที่ 1 (เลือกคำตอบ) และระดับที่ 2 (เลือกเหตุผล) เป็น 1 ได้ 0 คะแนน เมื่อนักเรียนได้มีรหัสคะแนนระดับที่ 1 (เลือกคำตอบ) หรือระดับที่ 2 (เลือกเหตุผล) เป็น 0 หรือ 0* สำหรับระดับ

ที่ 3 (ความมั่นใจ) จะไม่นำมาคำนวณเป็นคะแนนในการวิเคราะห์พารามิเตอร์ของข้อสอบ แต่จะใช้เพื่อการจำแนกสถานะมโนทัศน์ของผู้เรียนในขั้นตอนการวินิจฉัยเท่านั้น 1.4) ผู้วิจัยนำผลการตรวจแบบทดสอบวินิจฉัยสามระดับสำหรับประเมินมโนทัศน์ที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วน ของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยใช้การวิเคราะห์ตามทฤษฎีตอบสนองข้อสอบ (IRT) มาวิเคราะห์ค่าความยาก อำนาจจำแนก และค่าความเชื่อมั่นตามเกณฑ์ 1.5) ผู้วิจัยนำผลการตรวจแบบทดสอบวินิจฉัยสามระดับสำหรับประเมินมโนทัศน์ที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วน ของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยใช้การวิเคราะห์ตามทฤษฎีตอบสนองข้อสอบ (IRT) มาแปลผลวิเคราะห์มโนทัศน์ที่คลาดเคลื่อน โดยผู้วิจัยจำแนกมโนทัศน์ของนักเรียนออกเป็น 4 กลุ่ม ตามแนวคิดของ Pesman & Eryilmaz (2010) ซึ่งมีการให้รหัสคะแนน ตามข้อ 1.3.1 ซึ่งทำให้เกิดรูปแบบการตอบได้ทั้งหมด 10 รูปแบบ และจำแนกตามกลุ่มมโนทัศน์ได้ตามตารางต่อไปนี้

ตารางที่ 1 รูปแบบผลการตอบที่สอดคล้องกับกลุ่มมโนทัศน์ของนักเรียนตามแนวคิดของ Pesman & Eryilmaz (2010)

กลุ่มสถานะมโนทัศน์	ระดับที่ 1 (คำตอบ)	ระดับที่ 2 (เหตุผล)	ระดับที่ 3 (ความมั่นใจ)	การแปลความหมาย
1. มโนทัศน์ถูกต้อง	1	1	1	มีความรู้ถูกต้องสมบูรณ์
2. มโนทัศน์คลาดเคลื่อน	0	1	1	มีความเข้าใจผิดอย่างมั่นใจ
	0	0*	1	
3. ไม่มีความรู้	1	0	1	ไม่มีความรู้ในเนื้อหา/เดา
	1	0	0	
	0	1	0	
	0	0	1	
	0	0*	0	
	0	0	0	
4. ข้อผิดพลาด	1	1	0	มีความรู้แต่ไม่มั่นใจ/สะเพร่า

การจำแนกมโนทัศน์ของนักเรียนได้แบ่งเป็น 4 กลุ่มตามแนวคิดของ Pesman & Eryilmaz (2010) ซึ่งผู้วิจัยวิเคราะห์ระดับเหตุผลที่นักเรียนเลือก โดยพิจารณาตามรหัสการตอบที่ได้จาการหัสของตัวลงเชิงวินิจฉัย (Diagnostic Distractors) ทั้ง 2 ตัวเลือกที่มีความสอดคล้องกัน ซึ่งแต่ละตัวเลือกจะสะท้อนความเข้าใจผิดในด้านต่าง ๆ ได้แก่ 1) ด้านการตีความโจทย์ 2) ด้านการใช้ทฤษฎี สูตร กฎ นิยาม 3) ด้านการคิดคำนวณ 4) ด้านการตรวจสอบเมื่อได้รหัสตัวลงเชิงวินิจฉัยที่สอดคล้องกันแล้ว ผู้วิจัยจะพิจารณาการตอบระดับที่ 3 ใน การพิจารณาการจัดกลุ่มนักเรียนว่าอยู่ในกลุ่มมโนทัศน์ถูกต้อง (SK), กลุ่มมโนทัศน์คลาดเคลื่อน

(MC), กลุ่มไม่มีความรู้ (LK) หรือกลุ่มข้อผิดพลาด (ER) แล้วพิจารณาผลการจัดกลุ่มมาวิเคราะห์นักเรียนมาเป็นข้อมูลเพื่อใช้เป็นแนวทาง สำหรับการพัฒนาทักษะที่เหมาะสมแก่ผู้เรียนต่อไป

2. การวิเคราะห์ข้อมูลเชิงปริมาณ โดยใช้สถิติ ซึ่งแบ่งเป็นสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าร้อยละ และใช้สถิติที่ใช้ในการหาคุณภาพเครื่องมือ แบ่งการหาคุณภาพเครื่องมือออกเป็น 2 แนวคิดหลัก ดังนี้ 1) สถิติที่ใช้กับการวิเคราะห์แนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) ใช้สำหรับวิเคราะห์แบบทดสอบในขั้นต้นกับกลุ่มตัวอย่างกลุ่มที่ 1 เพื่อหาค่าความตรงเชิงเนื้อหา (IOC) ค่าความยาก (p) ค่าอำนาจจำแนก (r) 2) สถิติที่ใช้ในการวิเคราะห์ตามแนวคิดทฤษฎีการตอบสนองข้อสอบ (IRT) ในการวิเคราะห์คุณภาพข้อสอบ ผู้วิจัยใช้โมเดลโลจิสติก 2 พารามิเตอร์ (2 PLM) โดยมีเกณฑ์พิจารณาดังนี้ ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าฟังก์ชันสารสนเทศ(TIF) นอกจากนี้ ได้ทำการตรวจสอบข้อตกลงเบื้องต้นของทฤษฎี IRT ประกอบด้วย 1) ความเป็นมิติเดียว (Unidirectionality) โดยใช้การวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) พิจารณาจากค่า Eigenvalue ขององค์ประกอบ 2) ความเป็นอิสระเชิงท้องถิ่น (Local Independence) โดยใช้ดัชนี Q_3 ของ Yen ซึ่งค่าสหสัมพันธ์คะแนนส่วนเหลือของคู่ข้อสอบต้องต่ำกว่า 0.20 และ 3) ความสอดคล้องระหว่างโมเดลกับข้อมูลเชิงประจักษ์ (Model-Data Fit) โดยใช้สถิติทดสอบ $S - \chi^2$ ของ Orlando และ Thissen ซึ่งข้อสอบที่สอดคล้องกับโมเดลต้องไม่มีนัยสำคัญทางสถิติ ($p > 0.05$)

ผลการวิจัย

ผลการสร้างและหาคุณภาพของแบบทดสอบวินิจฉัยสามระดับสำหรับประเมินโน้ตสึนที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6 สามารถแสดงผลได้ ดังนี้

ผลการวิเคราะห์คุณภาพแบบทดสอบตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) พบว่า จากข้อสอบทั้งหมด 30 ข้อ มีข้อสอบผ่านเกณฑ์ดัชนีความสอดคล้อง (IOC) ระหว่าง 0.67 ถึง 1.00 จำนวน 27 ข้อ และเมื่อนำไปทดลองใช้ (Try-out) กับกลุ่มตัวอย่างที่ 1 พบว่ามีข้อสอบที่มีค่าความยาก (p) อยู่ระหว่าง 0.47 ถึง 0.73 ค่าอำนาจจำแนก (r) ระหว่าง 0.40 ถึง 0.60 และมีค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ (KR-20) เท่ากับ 0.85 ซึ่งผ่านเกณฑ์คัดเลือกจำนวน 20 ข้อ จากนั้นผู้วิจัยได้นำข้อสอบที่มีประสิทธิภาพทั้ง 20 ข้อ ไปพัฒนาต่อยอดเป็นแบบทดสอบวินิจฉัยสามระดับ แล้วนำไปให้ผู้เชี่ยวชาญจำนวน 5 ท่าน ประเมินค่าความสอดคล้องอีกครั้ง พบว่ามีข้อสอบจำนวน 18 ข้อที่มีค่า IOC ระหว่าง 0.60 ถึง 0.80 ซึ่งผ่านเกณฑ์ อย่างไรก็ตาม ผู้เชี่ยวชาญมีข้อเสนอแนะให้ปรับลดจำนวนข้อสอบลง เนื่องจากกลุ่มเป้าหมายเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 และมีเวลาจำกัดเพียง 1 ชั่วโมง หากคงข้อสอบไว้ทั้ง 18 ข้อ นักเรียนจะต้องตอบคำถามถึง 54 ครั้ง ซึ่งอาจเกิดความล้าและส่งผลต่อความเที่ยงตรงของการวัด ผู้วิจัยจึงได้ปรับปรุงเครื่องมือตามข้อเสนอแนะของผู้เชี่ยวชาญ โดยคัดเลือกข้อสอบวินิจฉัยสามระดับ เรื่อง โจทย์ปัญหาเศษส่วนที่มีความเหมาะสมที่สุดเหลือจำนวน 15 ข้อ เพื่อนำไปใช้ในการเก็บรวบรวมข้อมูลจริงต่อไป

การตรวจสอบรายงานผลข้อตกลงเบื้องต้น IRT ความเป็นมิติเดียว สามารถแสดงผลได้ ดังตาราง

ตารางที่ 2 ผลการวิเคราะห์องค์ประกอบเพื่อตรวจสอบข้อตกลงเบื้องต้นเรื่องความเป็นมิติเดียวของแบบทดสอบ

ข้อสอบ	ค่าน้ำหนักองค์ประกอบ (Factor Loading: λ)	ค่าความร่วม (Communality: h^2)
ข้อที่ 1	0.420	0.176
ข้อที่ 2	0.630	0.397
ข้อที่ 3	0.716	0.513
ข้อที่ 4	0.578	0.334
ข้อที่ 5	0.679	0.461
ข้อที่ 6	0.607	0.368
ข้อที่ 7	0.680	0.463
ข้อที่ 8	0.702	0.492
ข้อที่ 9	0.536	0.288
ข้อที่ 10	0.720	0.519
ข้อที่ 11	0.516	0.266
ข้อที่ 12	0.592	0.350
ข้อที่ 13	0.493	0.243
ข้อที่ 14	0.488	0.238
ข้อที่ 15	0.595	0.354
ค่า Eigenvalue ขององค์ประกอบที่ 1	5.461	
ค่า Eigenvalue ขององค์ประกอบที่ 2	2.205	
ร้อยละของ ความแปรปรวน	36.40%	

จากตารางที่ 2 ผลการวิเคราะห์องค์ประกอบเพื่อตรวจสอบข้อตกลงเบื้องต้นเรื่องความเป็นมิติเดียวของแบบทดสอบ พบว่า ค่า Eigenvalue ขององค์ประกอบหลักแรกมีค่าเท่ากับ 5.461 และองค์ประกอบหลักที่สองมีค่าเท่ากับ 2.205 โดยองค์ประกอบแรกเพียงองค์ประกอบเดียวสามารถอธิบายร้อยละความแปรปรวน

ของคะแนนรวมได้ร้อยละ 36.40 ซึ่งสูงกว่าเกณฑ์ขั้นต่ำที่ยอมรับได้ตามแนวคิดของ Carmines และ Zeller (1979) เมื่อพิจารณาค่าน้ำหนักองค์ประกอบ (Factor Loading) รายข้อ พบว่าข้อสอบทั้ง 15 ข้อ มีค่าน้ำหนักองค์ประกอบผ่านเกณฑ์มาตรฐานรายข้อ โดยมีค่าอยู่ระหว่าง 0.420 ถึง 0.720 และมีค่าความร่วม (Communality) อยู่ระหว่าง 0.176 ถึง 0.519 ข้อสอบที่มีค่าน้ำหนักองค์ประกอบสูงที่สุดคือ ข้อที่ 10 มีค่า $\lambda = 0.720$ รองลงมาคือ ข้อที่ 3 มีค่า $\lambda = 0.716$

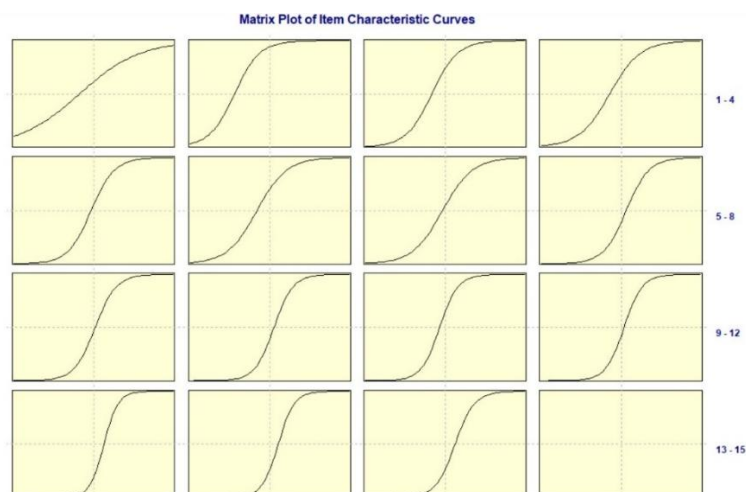
เมื่อพิจารณาดัชนีความกลมกลืนเชิงโครงสร้าง M_2 ของโมเดลการวัดมิติเดียว พบว่าค่า Comparative Fit Index (CFI) เท่ากับ 0.89 และ Tucker-Lewis Index (TLI) เท่ากับ 0.87 ซึ่งมีค่าต่ำกว่าเกณฑ์สากลขั้นสูงเล็กน้อย และมีค่า RMSEA เท่ากับ 0.15 ซึ่งค่อนข้างสูง เครื่องมือวัดชุดนี้ได้ผ่านกระบวนการพัฒนาและตรวจสอบความเที่ยงตรงเชิงเนื้อหาตามกรอบแนวคิดมิติเดียวอย่างเป็นระบบ (AERA, 2014) และมีค่าอำนาจจำแนกรายข้อ (a) ทุกข้อตามเกณฑ์ของ Baker (2001) นอกจากนี้ เมื่อทำการตรวจสอบข้อตกลงเบื้องต้นเรื่องความเป็นอิสระ เชิงท้องถิ่น (Local Independence) ด้วยค่าสหสัมพันธ์ส่วนเหลือ Yen's Q_3 ระหว่างคู่ข้อสอบ พบว่า มีค่าสูงสุดไม่เกินเกณฑ์ 0.20 และฟังก์ชันการตอบสนองข้อสอบมีลักษณะเป็นฟังก์ชันทางเดียว ซึ่งคุณสมบัติเหล่านี้ช่วยสนับสนุนมิติหลักในการวัด ผู้วิจัยจึงพิจารณาว่าข้อมูลเชิงประจักษ์มีความเป็นมิติเดียวในระดับที่ยอมรับได้ และมีความเหมาะสมเพียงพอในการนำข้อมูลคะแนนจากแบบทดสอบชุดนี้ไปวิเคราะห์ค่าพารามิเตอร์รายข้อ ตามทฤษฎีการตอบสนองข้อสอบ (IRT) ในขั้นตอนต่อไป

ต่อจากนั้นผู้วิจัยได้ทำการวิเคราะห์คุณภาพของข้อสอบเป็นรายข้อ ได้ผลดังตารางต่อไปนี้

ตารางที่ 3 ค่าพารามิเตอร์อำนาจจำแนก (a) และค่าความยาก (b) ของแบบทดสอบวินิจัยรายข้อ (N = 306)

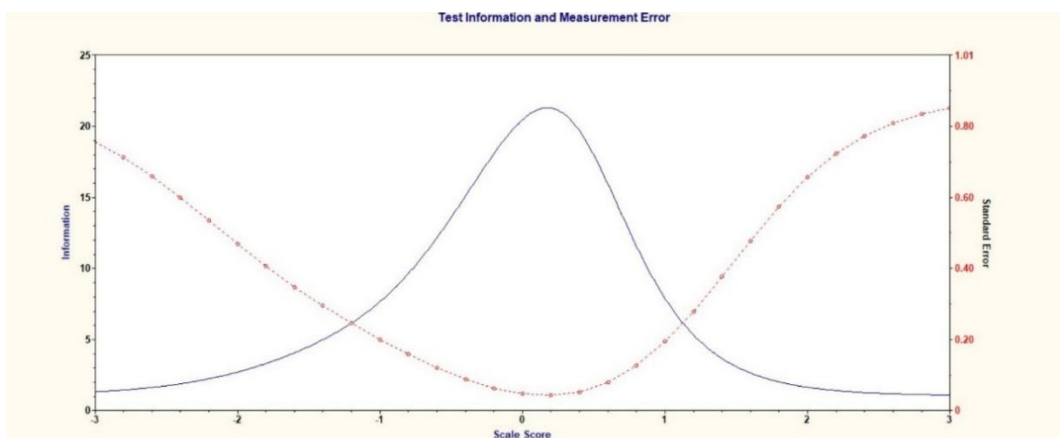
ข้อสอบ (Item)	ค่าอำนาจจำแนก (a)	ค่าความยาก (b)	ข้อสอบ (Item)	ค่าอำนาจจำแนก (a)	ค่าความยาก (b)
item1	2.53	0.50	item9	2.69	0.02
item2	2.19	-1.30	item10	2.99	0.13
item3	2.19	-0.51	item11	3.14	-0.23
item4	1.81	-0.43	item12	3.12	0.08
item5	2.39	-0.12	item13	3.94	0.34
item6	1.87	-0.48	item14	3.48	0.28
item7	1.84	-0.12	item15	2.85	0.33
item8	2.46	0.17			
ภาพรวม	1.81 - 3.94	-1.30 - 0.50			

จากตารางที่ 3 ผลการวิเคราะห์ค่าพารามิเตอร์ของแบบทดสอบวินิจัยจำนวน 15 ข้อ ภายใต้โมเดล 2PL พบว่า มีค่าอำนาจจำแนก (a) อยู่ระหว่าง 1.81 ถึง 3.94 ซึ่งสูงกว่าเกณฑ์สากล 1.70 ทุกข้อ และมีค่าความยาก (b) อยู่ระหว่าง -1.30 ถึง 0.50 จากผลวิเคราะห์ดังกล่าวสามารถเขียนแสดงเป็นกราฟ ได้ผลแสดงดังภาพต่อไปนี้



ภาพที่ 2 กราฟแสดงผลค่าฟังก์ชันการตอบสนองข้อสอบ

จากภาพที่ 2 เส้นโค้งลักษณะข้อสอบ (Item Characteristic Curve: ICC) ของข้อสอบทั้ง 15 ข้อ มีลักษณะทอดตัวสูงขึ้นจากซ้ายไปขวาอย่างเป็นระบบ ส่งผลให้ข้อมูลเชิงประจักษ์มีความพร้อมสำหรับการประมาณค่าพารามิเตอร์ภายใต้โมเดลโลจิสติกส์แบบ 2 พารามิเตอร์ (2PL) และจากผลการวิเคราะห์สามารถแสดงผลฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Function : TIF) และเส้นความคลาดเคลื่อนมาตรฐานของการวัด (Standard Error) ได้ดังกราฟต่อไปนี้



ภาพที่ 3 ฟังก์ชันสารสนเทศของแบบทดสอบและเส้นความคลาดเคลื่อนมาตรฐานของการวัด

เมื่อพิจารณาฟังก์ชันสารสนเทศของแบบทดสอบในภาพรวม (Test Information Function: TIF) ดังแสดงในภาพที่ 3 พบว่า เส้นโค้งสารสนเทศของแบบทดสอบ (เส้นทึบสีน้ำเงิน) มีลักษณะเป็นรูปประฆังคว่ำ โดยให้ค่าสารสนเทศสูงสุด (Maximum Information) อยู่ที่ประมาณ 21.00 ณ ระดับความสามารถ (θ) ระหว่าง 0.00 ถึง 0.50 ในขณะที่เส้นความคลาดเคลื่อนมาตรฐานของการวัด (Standard Error: SEM) (เส้นประสีแดง) มีค่าต่ำที่สุดในช่วงระดับความสามารถเดียวกัน ซึ่งสอดคล้องกับพารามิเตอร์ความยากรายข้อ ที่กระจายตัวอยู่รอบค่า 0 เป็นหลัก จากผลการวิเคราะห์นี้ ทำให้ทราบว่าแบบทดสอบวินิจฉัยสามระดับจำนวน 15 ข้อ ชุดนี้มีความแม่นยำในการวัดและให้สารสนเทศในระดับสูงเมื่อนำไปใช้ประเมินผู้เรียนที่มีระดับความสามารถ (θ) ปานกลางถึงค่อนข้างยาก และสามารถทำหน้าที่เป็นเครื่องมือวัดผลที่มีความเที่ยงตรง เชิงโครงสร้างในช่วงระดับความสามารถ (θ) ตั้งแต่ -1.0 ถึง +1.0 ได้

จากนั้นผู้วิจัยนำผลการศึกษานโน้ตทัศน์ที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วนของนักเรียน ชั้นประถมศึกษาปีที่ 6 มาจากกลุ่มข้อมูล ดังแสดงในตาราง ต่อไปนี้

ตารางที่ 4 ผลการจัดกลุ่มโน้ตทัศน์และพฤติกรรมการเรียนรู้ของนักเรียนในการแก้โจทย์ปัญหาเศษส่วน จากการวิเคราะห์ด้วยแบบทดสอบวินิจฉัยสามระดับ (N = 306, จำนวนรวม คือ 306 คน * 45 ตัวเลือก = 4,590 ข้อ)

กลุ่มมโนทัศน์และพฤติกรรม	อักษรย่อ	จำนวนครั้งที่ถูกเลือก	ร้อยละ (%)
1. กลุ่มมโนทัศน์ถูกต้อง (Sound Knowledge)	SK	2,326	50.68
2. กลุ่มมโนทัศน์คลาดเคลื่อน (Misconception)	MC	1,018	22.18
3. กลุ่มไม่มีความรู้ (Lack of Knowledge)	LK	1,148	25.01
4. กลุ่มข้อผิดพลาดทั่วไป (Error)	ER	98	2.14
รวม		4,590	100.00

จากตารางที่ 4 ผลการวิเคราะห์และจัดกลุ่มมโนทัศน์ในการแก้โจทย์ปัญหาเศษส่วนของนักเรียน ชั้นประถมศึกษาปีที่ 6 โดยใช้แบบทดสอบวินิจฉัยสามระดับ พบว่า จากจำนวนคำตอบทั้งหมด 4,590 ข้อ นักเรียนส่วนใหญ่จัดอยู่ในกลุ่มมโนทัศน์ถูกต้อง (SK) พบว่ามีจำนวนข้อที่ถูกเลือกเป็น 2,326 ข้อ คิดเป็นร้อยละ 50.68 เมื่อพิจารณาในส่วนของกลุ่มที่ตอบคำตอบไม่ถูกต้อง พบว่า มีนักเรียนจัดอยู่ในกลุ่มไม่มีความรู้ (LK) สูงที่สุดเป็นอันดับสอง มีจำนวนข้อที่ถูกเลือกเป็น 1,148 ข้อ คิดเป็นร้อยละ 25.01 รองลงมาคือกลุ่มมโนทัศน์คลาดเคลื่อน (MC) มีจำนวนข้อที่ถูกเลือกเป็น 1,018 ข้อ คิดเป็นร้อยละ 22.18 และพบกลุ่มข้อผิดพลาดหรือกลุ่มนักเรียนผู้มีความสับสนในการคำนวณ (ER) น้อยที่สุดมีจำนวนข้อที่ถูกเลือกเป็น 98 ข้อ คิดเป็นร้อยละ 2.14

จากผลการวิเคราะห์ในภาพรวมนี้สะท้อนให้เห็นประสิทธิภาพของแบบทดสอบวินิจัยสามระดับที่สามารถจำแนกความแตกต่างระหว่างนักเรียนที่มีความเข้าใจคลาดเคลื่อนเชิงมโนทัศน์อย่างเป็นระบบ หรือกลุ่มที่มีมโนทัศน์คลาดเคลื่อน (MC) ออกจากนักเรียนที่ตอบผิดเนื่องจากไม่มีความรู้เชิงเนื้อหาหรือใช้วิธีการเดาคำตอบ (LK) ได้ ทั้งนี้ ผู้วิจัยได้ทำการวิเคราะห์เจาะลึกกลุ่มมโนทัศน์ที่คลาดเคลื่อน (MC) จำแนกตามลักษณะกระบวนการแก้โจทย์ปัญหาคณิตศาสตร์ 4 ด้าน ซึ่งปรากฏผลดังแสดงในตารางถัดไป

ตารางที่ 5 จำนวนและร้อยละของนักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วน จำแนกตามรายด้าน (N = 306)

ด้านมโนทัศน์ที่คลาดเคลื่อน	จำนวน ตัวเลือก	จำนวนครั้งของ ตัวเลือกที่ถูกเลือก	ร้อยละ
1. ด้านการตีความโจทย์ปัญหา	9	316	31.04
2. ด้านการใช้ทฤษฎีบทหรือนิยาม	7	295	28.98
3. ด้านการคิดคำนวณ	7	102	10.02
4. ด้านการตรวจสอบการแก้ปัญห	7	305	29.96
รวม	30	1,018	22.18

จากตารางที่ 5 ผลการศึกษามโนทัศน์ที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6 จากกลุ่มตัวอย่าง 306 คน โดยใช้แบบทดสอบวินิจัยสามระดับ พบว่า ในภาพรวมนักเรียนมีมโนทัศน์ที่คลาดเคลื่อน (MC) คิดเป็นร้อยละ 22.18 ของการตอบทั้งหมด เมื่อพิจารณาจำแนกรายด้าน ตามกระบวนการแก้โจทย์ปัญหาคณิตศาสตร์ สามารถเรียงลำดับมโนทัศน์ที่คลาดเคลื่อนจากมากไปน้อยได้ ดังนี้ 1) ด้านการตีความโจทย์ปัญหา เป็นด้านที่นักเรียนมีมโนทัศน์คลาดเคลื่อนสูงที่สุด ถูกเลือกจำนวน 316 ครั้ง คิดเป็นร้อยละ 31.04 ด้านการตรวจสอบการแก้ปัญห นักเรียนมีมโนทัศน์คลาดเคลื่อนสูงเป็นอันดับสองถูกเลือกจำนวน 305 ครั้ง คิดเป็นร้อยละ 29.96 ด้านการใช้ทฤษฎีบทหรือนิยาม มีมโนทัศน์คลาดเคลื่อน ถูกเลือกจำนวน 295 ครั้ง คิดเป็นร้อยละ 28.98 และ ด้านการคิดคำนวณ เป็นด้านที่นักเรียนมีมโนทัศน์คลาดเคลื่อนน้อยที่สุด ถูกเลือกจำนวน 102 ครั้ง คิดเป็นร้อยละ 10.02

อภิปรายผล

จากการพัฒนาและประยุกต์ใช้แบบทดสอบวินิจัยสามระดับในการศึกษามโนทัศน์ที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6 ผู้วิจัยมีประเด็นสำคัญนำมาอภิปรายผลได้ ดังนี้ ผลการพัฒนาแบบทดสอบวินิจัยสามระดับ เรื่อง โจทย์ปัญหาเศษส่วน พบว่าได้แบบทดสอบฉบับสมบูรณ์จำนวน 15 ข้อ เป็นเครื่องมือวัดที่มีคุณภาพเพียงพอและมีความเหมาะสมในการนำไปใช้จริง ทั้งนี้เนื่องมาจาก

เครื่องมือผ่านกระบวนการตรวจสอบอย่างเป็นระบบตามแนวคิดทฤษฎีการทดสอบแบบมาตรฐานเดิม (CTT) โดยมีค่าความยากและค่าอำนาจจำแนกรายข้ออยู่ในเกณฑ์ที่เหมาะสม สามารถจำแนกผู้สอบที่มีระดับความรู้ต่างกันได้ดี และได้ค่าความเชื่อมั่น (KR-20) 0.85 ซึ่งแสดงถึงความเสถียรและความคงเส้นคงวาในการวัด นอกจากนี้ ได้ทำการปรับลดจำนวนข้อสอบจาก 18 ข้อ เหลือ 15 ข้อตามข้อเสนอแนะของผู้เชี่ยวชาญ เพื่อให้สอดคล้องกับข้อจำกัดด้านเวลาและระดับพัฒนาการของนักเรียนชั้นประถมศึกษาปีที่ 6 นั้น ถือเป็นสิ่งที่ช่วยลดความล้าในการทำข้อสอบ (Test Fatigue) เพื่อให้ผลการวัดสะท้อนความรู้และโน้ตทัศน์ที่คลาดเคลื่อนที่แท้จริงของนักเรียนได้อย่างเหมาะสม และช่วยคงไว้ซึ่งความเที่ยงตรงในการวัด (Validity) ของเครื่องมือวิจัยฉบับนี้ แล้วจากผลการหาคุณภาพเครื่องมือ พบว่า คุณภาพและความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบภายใต้ทฤษฎีการตอบสนองข้อสอบ (IRT) ผลการวิเคราะห์เพื่อตรวจสอบข้อตกลงเบื้องต้น (Assumptions) ของทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) พบว่า แบบทดสอบชุดนี้มีความเป็นมิติเดียวและความเป็นอิสระเชิงท้องถิ่น ในระดับที่ยอมรับได้และมีความเหมาะสมเพียงพอ ซึ่งมีประเด็นสำคัญที่สามารถนำมาอภิปรายผลได้ ดังนี้ เกี่ยวกับ ความเป็นมิติเดียว

ผลการวิเคราะห์องค์ประกอบเชิงสำรวจ พบว่า องค์ประกอบแรกเพียงองค์ประกอบเดียวสามารถอธิบายความแปรปรวนของคะแนนรวมได้ร้อยละ 36.40 ซึ่งสูงกว่าเกณฑ์ขั้นต่ำร้อยละ 20.00 ตามแนวคิดของ Carmines and Zeller (1979) และ Reckase (1979) ม้องค์ประกอบแรกที่เด่นชัดเพียงพอ ย่อมยืนยันได้ว่า เครื่องมือวัดชุดนี้มุ่งวัดคุณลักษณะหรือความสามารถหลักเพียงมิติเดียวเป็นสำคัญสอดคล้องกับผลการพิจารณาค่าน้ำหนักองค์ประกอบ (Factor Loading) รายข้อที่มีค่าอยู่ระหว่าง 0.420 ถึง 0.720 ซึ่งผ่านเกณฑ์มาตรฐานรายข้อ และมีค่าความร่วมกัน (Communality) อยู่ในระดับที่เหมาะสม แสดงให้เห็นว่าข้อสอบทั้ง 15 ข้อมีความเป็นเอกพันธ์ (Homogeneity) และร่วมกันวัดในสิ่งเดียวกันอย่างเป็นระบบบนส่วนของดัชนีความกลมกลืนเชิงโครงสร้าง (M_2) ของโมเดลการวัดมิติเดียว แม้จะพบว่าค่า Comparative Fit Index (CFI = 0.89) และ Tucker-Lewis Index (TLI = 0.87) มีค่าต่ำกว่าเกณฑ์สากลขั้นสูงเล็กน้อย และมีค่า RMSEA เท่ากับ 0.15 ซึ่งค่อนข้างสูง

นอกจากนี้ เครื่องมือวัดชุดนี้ได้ผ่านกระบวนการพัฒนาและตรวจสอบความเที่ยงตรงเชิงเนื้อหา (Content Validity) ตามกรอบแนวคิดมิติเดียวอย่างเป็นระบบตามมาตรฐานของ American Educational Research Association (AERA, 2014) ประกอบกับข้อสอบทุกข้อมีค่าอำนาจจำแนกรายข้อ (a) อยู่ในเกณฑ์ดีตามเกณฑ์ของ Baker (2001) จึงเป็นหลักฐานเชิงประจักษ์ที่ช่วยลดทอนความกังวลเกี่ยวกับดัชนีความกลมกลืนที่เบี่ยงเบนไปเล็กน้อยดังกล่าว และผลการตรวจสอบข้อตกลงเบื้องต้นเรื่องความเป็นอิสระเชิงท้องถิ่นด้วยค่าสหสัมพันธ์ส่วนเหลือ Yen's Q_3 ระหว่างคู่ข้อสอบ พบว่ามีค่าสูงสุดไม่เกินเกณฑ์ 0.20 และฟังก์ชันการตอบสนองข้อสอบมีลักษณะเป็นฟังก์ชันทางเดียว (Monotonicity) ตามทฤษฎี IRT ของ Lord (1980) นั้น ข้อตกลงเบื้องต้นเรื่องความเป็นมิติเดียวและความเป็นอิสระเชิงท้องถิ่นเป็นเงื่อนไขที่เกื้อหนุนและเป็นเหตุเป็น

ผลซึ่งกันและกัน (Condition of Local Independence) หมายความว่า เมื่อควบคุมความสามารถของผู้สอบ แล้วคะแนนจากการตอบสนองในข้อสอบแต่ละข้อต้องเป็นอิสระต่อกัน การที่ค่า Yen's Q_3 ผ่านเกณฑ์อย่างสมบูรณ์ จึงเป็นสิ่งยืนยันว่าไม่มีอิทธิพลปนเปื้อนจากปัจจัยภายนอกอื่นใดที่ส่งผลต่อการตอบข้อสอบของผู้สอบ นอกเหนือจากความสามารถในมิติหลักที่ต้องการวัด

จากเหตุผลเชิงทฤษฎีและหลักฐานเชิงประจักษ์ข้างต้น ทั้งในด้านสัดส่วนความแปรปรวนที่อธิบายได้ กระบวนการพัฒนาตามมาตรฐาน AERA (2014) คำอำนาจจำแนกตามเกณฑ์ Baker (2001) และค่า Yen's Q_3 จึงมีความสมเหตุสมผลเพียงพอที่จะสรุปได้ว่า ข้อมูลเชิงประจักษ์มีความเป็นมิติเดียวในระดับที่ยอมรับได้ และมีความเหมาะสมในการนำข้อมูลคะแนนจากแบบทดสอบชุดนี้ไปวิเคราะห์เพื่อประมาณค่าพารามิเตอร์รายข้อตามทฤษฎีการตอบสนองข้อสอบ (IRT) ในขั้นต่อไป ผลการวิเคราะห์ค่าพารามิเตอร์ของแบบทดสอบวินิจัย จำนวน 15 ข้อ ภายใต้โมเดล 2 พารามิเตอร์ (2PLM) พบว่า มีค่าอำนาจจำแนก (a) อยู่ระหว่าง 1.81 ถึง 3.94 ซึ่งสูงกว่าเกณฑ์สากล 1.70 ทุกข้อ แสดงว่าแบบทดสอบวินิจัยชุดนี้ทำหน้าที่วัดตัวแปรที่ต้องการศึกษาได้อย่างมีเอกภาพ และเป็นเครื่องมือที่มีความสอดคล้องภายใน และมีค่าความยาก (b) อยู่ระหว่าง -1.30 ถึง 0.50 อยู่ในช่วงเกณฑ์มาตรฐานสากลที่ยอมรับได้ เนื่องจากเครื่องมือที่ใช้เป็นแบบทดสอบวินิจัย ซึ่งมีวัตถุประสงค์เพื่อค้นหาข้อบกพร่องในส่วนที่ผู้เรียนยังไม่เข้าใจ ดังนั้น ข้อสอบจึงไม่ควรยากเกินไปการที่ค่าความยากเกาะกลุ่มอยู่ที่ระดับง่ายถึงปานกลาง จึงมีความเหมาะสมในการนำไปใช้ ซึ่งสอดคล้องกับ เยาวดี วิบูลย์ศรี (2548) ที่กล่าวว่า ธรรมชาติของแบบทดสอบวินิจัยมุ่งเน้นการค้นหาจุดบกพร่องเฉพาะจุดของนักเรียน ดังนั้น ข้อสอบวินิจัยจึงถูกออกแบบมาเพื่อวัดทักษะย่อย ๆ ระดับความยากจึงต้องสอดคล้องกับจุดประสงค์การเรียนรู้พื้นฐาน ไม่ควรยากเกินไป

นอกจากนี้ยังสอดคล้องกับงานวิจัยของ ขนาธิป หุ้ยแป (2553) ที่พบว่า ค่าความยากที่เกาะกลุ่มในระดับปานกลางจะช่วยลดความคลาดเคลื่อนและช่วยให้สามารถวินิจฉัยแนวคิดที่คลาดเคลื่อนของผู้เรียนได้อย่างแม่นยำยิ่งขึ้น แล้วเมื่อพิจารณาฟังก์ชันสารสนเทศของแบบทดสอบในภาพรวม (Test Information Function : TIF) พบว่า เส้นโค้งสารสนเทศให้ค่าสารสนเทศสูงสุดอยู่ที่ประมาณ 21.00 ณ ระดับความสามารถ (θ) ระหว่าง 0.0 ถึง 0.5 ในขณะที่เส้นความคลาดเคลื่อนมาตรฐานของการวัด (SEM) มีค่าต่ำที่สุดในช่วงระดับความสามารถเดียวกัน ลักษณะของกราฟที่ปรากฏนี้มีความสมเหตุสมผลเชิงทฤษฎีการตอบสนองข้อสอบ (IRT) สอดคล้องกับแนวคิดของ Lord (1980) และ Hambleton et al. (1991) ที่ระบุว่า ค่าสารสนเทศของแบบทดสอบในภาพรวมเกิดจากการรวมกันของค่าสารสนเทศรายข้อ ซึ่งสัมพันธ์โดยตรงกับค่าพารามิเตอร์ความยาก (b) และอำนาจจำแนก (a) การที่แบบทดสอบชุดนี้ให้ค่าสารสนเทศสูงในช่วงระดับความสามารถ (θ) ตั้งแต่ -1.0 ถึง +1.0 แสดงให้เห็นว่าแบบทดสอบวินิจัยชุดนี้มีความเที่ยงตรงเชิงโครงสร้างและมีความแม่นยำในการวัดสูงที่สุดเมื่อนำไปใช้ประเมินผู้เรียนที่มีระดับความสามารถปานกลาง ซึ่งลักษณะดังกล่าวถือเป็นคุณลักษณะที่เหมาะสมอย่างยิ่งสำหรับแบบทดสอบวินิจัย สอดคล้องกับงานวิจัยของ ศิริชัย กาญจนวาสี

(2556) ในเรื่องทฤษฎีการทดสอบแนวใหม่ ที่อธิบายว่าเครื่องมือวัดผลที่ดีไม่จำเป็นต้องให้สารสนเทศสูงครอบคลุมผู้สอบทุกระดับความสามารถ แต่ควรมุ่งให้สารสนเทศสูงสุดสอดคล้องกับกลุ่มเป้าหมายที่ต้องการวัด นอกจากนี้ยังสอดคล้องกับงานวิจัยของ กมลวรรณ ดังธณกานนท์ (2557) ที่พบว่า แบบทดสอบวินิจฉัยที่สร้างขึ้นตามกรอบพารามิเตอร์ที่เหมาะสมจะให้ค่าสารสนเทศสูงสุดรอบ ๆ ระดับความสามารถปานกลาง ซึ่งช่วยให้ผู้สอนสามารถจำแนกและระบุข้อบกพร่องทางการเรียนรู้ของผู้เรียนกลุ่มใหญ่ในชั้นเรียนได้อย่างถูกต้องและแม่นยำ นำไปสู่การวางแผนจัดการเรียนรู้ได้อย่างมีประสิทธิภาพ

จากความสมเหตุสมผลเชิงทฤษฎีและข้อตกลงเบื้องต้นของเครื่องมือวัดผลและให้สารสนเทศสูงในการประเมินผู้เรียนกลุ่มดังกล่าว ผู้วิจัยจึงได้นำแบบทดสอบวินิจฉัยสามระดับที่ผ่านการตรวจสอบคุณภาพเชิงโครงสร้างอย่างเป็นระบบนี้ ไปดำเนินการจัดกระทำข้อมูลเชิงประจักษ์เพื่อจำแนกพฤติกรรมการเรียนรู้และระดับโน้ตทัศน์ในการแก้โจทย์ปัญหาเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 306 คน ซึ่งจากการวิเคราะห์ข้อมูลผลการตอบจากคำตอบทั้งหมด 4,590 คำตอบ พบประเด็นสำคัญที่น่าสนใจและสามารถนำมาอภิปรายผลเชิงพฤติกรรมการเรียนรู้ได้ดังต่อไปนี้

ผลการวิเคราะห์และจัดกลุ่มโมทัศน์และพฤติกรรมการเรียนรู้ของนักเรียนชั้นประถมศึกษาปีที่ 6 ในการแก้โจทย์ปัญหาเศษส่วน ดังนี้ ประเด็นแรก นักเรียนกลุ่มที่มีความเข้าใจที่ถูกต้อง (Sound Knowledge : SK) คิดเป็นร้อยละ 50.68 ซึ่งสะท้อนให้เห็นว่า มีนักเรียนประมาณครึ่งหนึ่งที่มีความรู้ความเข้าใจเชิงมโนทัศน์ และสามารถนำไปประยุกต์ใช้ในการแก้โจทย์ปัญหาเศษส่วนได้อย่างถูกต้องสมบูรณ์ ในขณะที่นักเรียนอีกเกือบครึ่งหนึ่งยังคงประสบปัญหาในการเรียนรู้ ซึ่งสอดคล้องกับแนวคิดของ Vamvakoussi and Vosniadou (2004) ที่ระบุว่า มโนทัศน์เรื่องเศษส่วนและโจทย์ปัญหาเป็นเนื้อหาที่มีความเป็นนามธรรมสูงและมีความซับซ้อน นักเรียนจึงมักประสบความยากลำบากในการสร้างความเข้าใจให้ถูกต้องสมบูรณ์ ประเด็นที่สอง เมื่อพิจารณาจากกลุ่มนักเรียนที่ตอบไม่ถูกต้อง พบว่ากลุ่มที่ไม่มีความรู้ (Lack of Knowledge : LK) มีสัดส่วนสูงที่สุดเป็นอันดับสอง คิดเป็นร้อยละ 25.01 ผลดังกล่าวชี้ให้เห็นว่า สาเหตุหลักที่ทำให้นักเรียนไม่สามารถแก้โจทย์ปัญหาเศษส่วนได้ เกิดจากการขาดความรู้พื้นฐานหรือทักษะที่จำเป็นในการคำนวณ ซึ่งสอดคล้องกับการจัดประเภทข้อผิดพลาดทางคณิตศาสตร์ของ Radatz (1979) ที่ระบุว่า ความบกพร่องในความรู้พื้นฐานเป็นเรื่องสำคัญที่ขัดขวางกระบวนการแก้โจทย์ปัญหาในขั้นต่อไป ประเด็นที่สาม นักเรียนที่มีมโนทัศน์คลาดเคลื่อน (Misconception : MC) มีสัดส่วนถึงร้อยละ 22.18 ซึ่งเป็นข้อมูลที่มีความสำคัญอย่างยิ่งในการวิจัยเชิงวินิจฉัยครั้งนี้ เนื่องจากนักเรียนกลุ่มนี้มีได้ขาดความรู้ แต่มีความเข้าใจในกฎเกณฑ์หรือหลักการที่ผิดพลาดและฝังลึกในระบบความคิด เช่น การนำหลักการของจำนวนนับไปใช้กับเศษส่วนโดยตรง ซึ่งสอดคล้องกับงานวิจัยของ ชนาธิป พุ้ยแป (2553) ที่พบว่า เรื่องเศษส่วนเป็นเนื้อหาที่นักเรียนเกิดมโนทัศน์คลาดเคลื่อนสูงที่สุด โดยเฉพาะอย่างยิ่งการสับสนระหว่างกระบวนการบวก ลบ คูณ และหารเศษส่วน การที่แบบทดสอบวินิจฉัยสามระดับสามารถแยกแยะกลุ่มที่มีมโนทัศน์คลาดเคลื่อน (MC) ออกจากกลุ่มที่ไม่มีความรู้

(LK) ได้อย่างชัดเจน ย่อมช่วยให้ผู้สอนสามารถเลือกกลยุทธ์การจัดการเรียนรู้เพื่อปรับเปลี่ยนมโนทัศน์ให้นักเรียนได้ตรงตามประเด็นข้อบกพร่องที่แท้จริง สอดคล้องกับประโยชน์ของเครื่องมือวัดผลเชิงวินิจฉัยตามแนวคิดของ เยาวดี วิบูลย์ศรี (2556)

สำหรับกลุ่มข้อผิดพลาดหรือนักเรียนที่มีความสับสนในการคำนวณ (ER) มีสัดส่วนน้อยที่สุดร้อยละ 2.14 แสดงให้เห็นว่าความผิดพลาดของนักเรียนส่วนใหญ่ในงานวิจัยนี้ เกิดจากความบกพร่องทางปัญญาและมโนทัศน์เป็นหลัก มิใช่เกิดจากความไม่ใส่ใจหรือความสับสนในขณะทำแบบทดสอบ เมื่อพิจารณาไปที่กลุ่มผู้มีมโนทัศน์คลาดเคลื่อน (MC) ซึ่งพิจารณาจำแนกรายด้านตามกระบวนการแก้โจทย์ปัญหาคณิตศาสตร์ พบประเด็นสำคัญที่สะท้อนถึงลักษณะมโนทัศน์ที่คลาดเคลื่อนของนักเรียน ดังนี้ ด้านการตีความโจทย์ปัญหา เป็นด้านที่นักเรียนมีมโนทัศน์คลาดเคลื่อนสูงที่สุด คิดเป็นร้อยละ 31.04 จากผลดังกล่าวสะท้อนให้เห็นว่า อุปสรรคสำคัญที่สุดในการแก้โจทย์ปัญหาเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6 คือการขาดความสามารถในการแปลความหมายจากภาษาตัวอักษรไปสู่แนวคิดทางคณิตศาสตร์ นักเรียนมักสับสนกับคำสำคัญในโจทย์ เช่น คำว่าของ หรือของที่เหลืออยู่ ซึ่งต้องใช้การคิดขั้นสูงในการหาส่วนสัมพันธ์ ลักษณะข้อผิดพลาดนี้สอดคล้องกับขั้นตอนการแก้โจทย์ปัญหาของ Polya (1957) ที่ระบุว่า ขั้นทำความเข้าใจโจทย์ (Understanding the Problem) เป็นโครงสร้างพื้นฐานที่สำคัญที่สุด อีกทั้งยังสอดคล้องกับงานวิจัยของ สิริพร ทิพย์คง (2545) ที่พบว่า ข้อบกพร่องส่วนใหญ่ของเด็กไทยในการทำโจทย์ปัญหาคณิตศาสตร์เกิดจากการตีความโจทย์ไม่ถูกต้อง และสอดคล้องกับงานวิจัยของ มะนิวรรณ แก้วดวงดี (2567) ที่ทำการศึกษาและชี้ให้เห็นว่า นักเรียนในระดับชั้นประถมศึกษา มักมีมโนทัศน์คลาดเคลื่อนในขั้นตอนการอ่านและการวิเคราะห์โจทย์ปัญหามากที่สุด เนื่องจากไม่สามารถเชื่อมโยงสถานการณ์ในโจทย์เข้ากับสัญลักษณ์ทางคณิตศาสตร์ได้อย่างถูกต้อง

นอกจากนี้ยังสอดคล้องกับงานวิจัยของ Verschaffel, Greer, and De Corte (2000) ที่ระบุว่า นักเรียนส่วนใหญ่มีแนวโน้มที่จะละเลยบริบทจริงของโจทย์ปัญหา และมุ่งหาเฉพาะตัวเลขเพื่อนำมาคำนวณ ด้วยวิธีการที่ตื้นเขินเคยโดยปราศจากการตีความที่แท้จริง สำหรับด้านการตรวจสอบการแก้ปัญหา เป็นด้านที่มีมโนทัศน์คลาดเคลื่อนสูงเป็นอันดับที่สอง คิดเป็นร้อยละ 29.96 ตัวเลขนี้ชี้ให้เห็นว่า นักเรียนไทยส่วนใหญ่มีพฤติกรรมละเลยการตรวจทานคำตอบ เมื่อคำนวณเสร็จสิ้นแล้วมักจะยอมรับ และใช้คำตอบนั้นทันทีโดยไม่พิจารณาความสมเหตุสมผลของผลลัพธ์ ซึ่งสอดคล้องกับงานวิจัยของ กมลวรรณ ดังชนกานนท์ (2557) ที่พบว่า พฤติกรรมการเรียนรู้ของนักเรียนมักมองข้ามขั้นตอนการประเมินและตรวจสอบความถูกต้อง ส่งผลให้เกิดมโนทัศน์คลาดเคลื่อนแบบถาวรเนื่องจากเข้าใจว่าคำตอบที่ตนหาได้นั้นถูกต้องแล้ว ในด้านการใช้ทฤษฎีบทหรือนิยาม คิดเป็นร้อยละ 28.98 และด้านการคิดคำนวณ คิดเป็นร้อยละ 10.02

ผลการวิเคราะห์แสดงให้เห็นว่า นักเรียนมีมโนทัศน์คลาดเคลื่อนในด้านการคิดคำนวณน้อยที่สุด ซึ่งเป็นประเด็นที่น่าสนใจมากกว่า นักเรียนมิได้มีปัญหาที่ทักษะการคำนวณ แต่ปัญหากลับไปอยู่ที่หลักการเชิงนิยามหรือทฤษฎี กล่าวคือ นักเรียนไม่เข้าใจกฎเกณฑ์เฉพาะของเศษส่วน เช่น มโนทัศน์คลาดเคลื่อนว่าการ

บวกเศษส่วนสามารถนำตัวเศษบวกตัวเศษ และตัวส่วนบวกตัวส่วนได้เลย หรือจำสับสนระหว่างเกณฑ์การบวก และการคูณเศษส่วน ซึ่งจุดนี้สอดคล้องกับงานวิจัยของ อนันริกา เนียมแก้ว (2565) ที่พบว่า นักเรียนมักมีความเข้าใจคลาดเคลื่อนเกี่ยวกับข้อตกลงและหลักการพื้นฐานของเศษส่วน โดยเฉพาะเรื่องการเปรียบเทียบ และการดำเนินชีวิตประจำวันที่ต้องอาศัยนิยามทางคณิตศาสตร์ที่ถูกต้อง และสอดคล้องกับแนวคิดของ Ni and Zhou (2005) ที่อธิบายปรากฏการณ์นี้ว่าเกิดจากการนำกฎเกณฑ์ของระบบจำนวนนับ (Whole number bias) ที่ตนเองคุ้นเคยมาใช้ตัดสินและแก้โจทย์ปัญหาในระบบเศษส่วน จากผลการวิเคราะห์ข้างต้น ทำให้ค้นพบข้อเท็จจริงที่ว่า ปัญหาที่แท้จริงในการเรียนเรื่องโจทย์ปัญหาเศษส่วน เกิดจากความบกพร่องที่เริ่มมาจากการตีความภาษาและการยึดติดกับนิยามเดิม การที่แบบทดสอบวินิจฉัยสามระดับสามารถระบุข้อบกพร่องตามกระบวนการแก้ปัญหาเหล่านี้ได้อย่างชัดเจน ย่อมสอดคล้องกับแนวคิดของ ศิริเดช สุชีวะ (2538) ที่ระบุว่า เครื่องมือวัดผลเชิงวินิจฉัยที่ดีจะต้องมีความสามารถในการจำแนกรูปพรรณของข้อผิดพลาด หรือจุดบกพร่องเฉพาะตัวของนักเรียนได้อย่างเป็นระบบ เพื่อให้ผู้สอนใช้เป็นสารสนเทศเชิงลึกในการช่วยเหลือ และออกแบบการเรียนการสอนซ่อมเสริมเพื่อปรับรื้อระบบความคิดของผู้เรียนรายบุคคลได้อย่างตรงจุดและมีประสิทธิภาพเหมาะสม สอดคล้องกับประโยชน์ของเครื่องมือวัดผลเชิงวินิจฉัยตามแนวคิดของ เยาวดี วิบูลย์ศรี (2556)

สรุปและข้อเสนอแนะ

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างและหาคุณภาพของแบบทดสอบวินิจฉัยสามระดับสำหรับประเมินโน้ตทัศน์ที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6 และเพื่อศึกษามโนทัศน์ที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6 ใน 4 ด้าน ได้แก่ ด้านการตีความ ด้านการใช้ทฤษฎีบทหรือนิยามด้านการคิดคำนวณ และด้านการตรวจสอบการแก้ปัญหา โดยใช้แบบวินิจฉัยสามระดับที่พัฒนาขึ้น

ผลการวิจัยสรุปได้ว่า แบบทดสอบวินิจฉัยสามระดับสำหรับประเมินโน้ตทัศน์ที่คลาดเคลื่อนในการแก้โจทย์ปัญหาเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6 เป็นเครื่องมือวัดที่มีคุณภาพเพียงพอและมีความเหมาะสมในการนำไปใช้จริง ทั้งยังสามารถใช้ในการจัดกลุ่มลักษณะมโนทัศน์ที่คลาดเคลื่อนรายด้าน อันประกอบไปด้วย 1) ด้านการตีความโจทย์ปัญหา 2) ด้านการตรวจสอบการแก้ปัญหา 3) ด้านการใช้ทฤษฎีบทหรือนิยาม และ 4) ด้านการคิดคำนวณ ได้ด้วย

สำหรับข้อค้นพบสำคัญของงานวิจัย จากการวิจัยครั้งนี้สะท้อนให้เห็นว่าการที่ผู้เรียนทำแบบทดสอบไม่ถูกต้อง หรือตอบคำถามเกี่ยวกับเรื่องโจทย์ปัญหาเศษส่วนไม่ถูกต้อง มิได้แสดงว่าผู้เรียนขาดความรู้ในเรื่องโจทย์ปัญหาเศษส่วนกรณีเดียว เนื่องจากการที่ผู้เรียนอาจมีปัจจัยอื่น ๆ ที่ส่งผลกับผู้เรียนที่ทำให้เกิดการเข้าใจผิด หรือเกิดมโนทัศน์ความคลาดเคลื่อนขึ้นได้ตามที่เป็นผลในการวิจัยครั้งนี้ ทั้งยังได้ค้นพบแบบทดสอบที่แตกต่าง

จากที่เคยใช้มา นั่นคือแบบทดสอบวินิจฉัยสามระดับที่สามารถวัดได้มากกว่าแค่การตอบคำถามของโจทย์ข้อนั้น ๆ แต่ยังสามารถมีการเพิ่มระดับเหตุผลและระดับความมั่นใจ ที่สะท้อนผลที่ระเียด และเป็นผลดีต่อการวิเคราะห์ผู้เรียนรายบุคคลที่ครอบคลุมยิ่งขึ้นได้

โดยสรุป งานวิจัยในครั้งนี้มีส่วนเติมเต็มองค์ความรู้ด้านการสร้างแบบทดสอบที่ครอบคลุมการวิเคราะห์ที่ละเอียด และตรงจุดเหมาะสมต่อการนำผลไว้วิเคราะห์เพื่อใช้ในการพัฒนาผู้เรียนได้ตรงประเด็นยิ่งขึ้น ซึ่งเป็นผลดีต่อผู้สอนในการจัดการเรียนรู้ที่ตอบสนองต่อประเด็นที่ผู้เรียนควรพัฒนาได้เป็นอย่างดี

ข้อเสนอแนะ

1. ด้านการจัดการเรียนการสอนเพื่อแก้ไขข้อบกพร่อง จากผลการวิจัยพบว่า นักเรียนมีมีโนทัศน์คลาดเคลื่อนในด้านการตีความโจทย์ปัญหา และด้านการตรวจสอบการแก้ปัญหาลูกที่สุด ครูผู้สอนวิชาคณิตศาสตร์ในระดับชั้นประถมศึกษาจึงควรเปลี่ยนจุดเน้นจากการฝึกฝนทักษะการคิดคำนวณเพียงอย่างเดียว ไปสู่การจัดการกิจกรรมการเรียนรู้ที่เน้นการถอดรหัสภาษาคณิตศาสตร์ (Mathematical Literacy) เช่น การใช้วิธีการวาดรูปบาร์โมเดล (Bar Model) หรือการฝึกใช้ขั้นตอนของโพลยา (Polya) ให้มากขึ้นโดยเฉพาะขั้นตอนการวิเคราะห์สิ่งที่โจทย์กำหนดและการตรวจทานความสมเหตุสมผลของคำตอบก่อนส่งงาน

2. ด้านการนำเครื่องมือไปใช้จำแนกผู้เรียน สำหรับสถานศึกษา สามารถนำแบบทดสอบวินิจฉัยสามระดับจำนวน 15 ข้อชุดนี้ ซึ่งผ่านการตรวจสอบข้อตกลงเบื้องต้นเรื่องความเป็นมิติเดียว และมีฟังก์ชันสารสนเทศ (TIF) สูงในช่วงความสามารถปานกลางไปใช้เป็นเครื่องมือคัดกรอง (Screening Tool) นักเรียนชั้นประถมศึกษาปีที่ 6 ในช่วงต้นปีการศึกษา หรือก่อนการทดสอบระดับชาติ (O-NET) เพื่อระบุและจำแนกรูปแบบของข้อผิดพลาดของนักเรียนเป็นรายบุคคล เพื่อให้โรงเรียนสามารถจัดกลุ่มเด็กเพื่อทำสอนซ่อมเสริมได้ตรงตามความไม่เข้าใจอย่างตรงจุด

3. จากผลการวิจัยชี้ให้เห็นสัดส่วนของกลุ่มโนทัศน์คลาดเคลื่อน (Misconception) ที่สูงถึงร้อยละ 22.18 ซึ่งเด็กกลุ่มนี้เข้าใจว่าตัวเองทำถูกแต่ใช้หลักการที่ผิด ครูผู้สอนจึงไม่ควรใช้วิธีสอนซ่อมเสริมแบบดั้งเดิมที่เน้นการให้ทำโจทย์ซ้ำ ๆ แต่ควรใช้กลยุทธ์การเผชิญหน้าทางปัญญา (Cognitive Conflict) หรือการตั้งคำถามย้อนกลับ เพื่อให้นักเรียนตระหนักถึงความขัดแย้งในความคิดของตนเองและนำไปสู่การปรับเปลี่ยนโนทัศน์ที่ถูกต้อง

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

1. เนื่องจากการวิจัยครั้งนี้เสร็จสิ้นในขั้นของการวินิจฉัย และจำแนกประเภทความเข้าใจคลาดเคลื่อนรายด้านตามกระบวนการแก้โจทย์ปัญหาเศษส่วนแล้ว ในการวิจัยครั้งต่อไปจึงควรมีการพัฒนาแผนการจัดการเรียนรู้ นวัตกรรม หรือชุดกิจกรรมซ่อมเสริมเฉพาะทาง เพื่อนำมาทดลองใช้แก้ปัญหากับกลุ่มนักเรียนที่ถูกระบุว่าเป็นกลุ่ม MC หรือ LK จากงานวิจัยชิ้นนี้ แล้วประเมินผลสัมฤทธิ์ในลักษณะการวิจัยเชิงทดลอง (Experimental Research)

2. ในการวิจัยครั้งนี้ใช้โมเดลพารามิเตอร์โลจิสติก 2 พารามิเตอร์ (2PL Model) ในทฤษฎีการตอบสนองข้อสอบ (IRT) มิติเดียว ในการวิจัยครั้งต่อไปอาจขยายผลโดยการประยุกต์ใช้โมเดลทฤษฎีการตอบสนองข้อสอบหลายมิติ (Multidimensional IRT) หรือการวิเคราะห์ด้วยโมเดลการวินิจฉัยทางปัญญา (Cognitive Diagnosis Models: CDMs) เพื่อให้สามารถวิเคราะห์ทักษะย่อยที่ซ่อนอยู่ในการแก้โจทย์ปัญหาเศษส่วนได้อย่างละเอียดและจำเพาะเจาะจงมากยิ่งขึ้น

3. สำหรับกลุ่มนักเรียนที่มีมีโน้ตคนที่คลาดเคลื่อน (MC) การวิจัยครั้งต่อไปควรใช้รูปแบบการวิจัยแบบผสมผสาน (Mixed Methods) โดยเพิ่มการสัมภาษณ์เชิงลึก (In-depth Interview) หรือการให้นักเรียนใช้วิธีการคิดออกเสียงขณะแก้โจทย์ปัญหา (Think-Aloud Protocol) กับนักเรียนกลุ่มนี้ เพื่อถอดรหัสกระบวนการคิดและหาสาเหตุที่แท้จริงว่าเหตุใดนักเรียนจึงเกิดความเข้าใจที่ผิดพลาดเหล่านั้น ซึ่งจะทำได้ ข้อมูลเชิงคุณภาพมาสนับสนุนข้อมูลเชิงปริมาณให้สมบูรณ์ยิ่งขึ้น

บรรณานุกรม

กมลวรรณ ตังธกานนท์. (2557). *การวัดและประเมินทักษะการปฏิบัติ*. สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ชนาธิป ห้วยแป. (2553). *คู่มือการประเมินคุณภาพผู้เรียนและเครื่องมือวัดผลสัมฤทธิ์ทางการเรียน*.

สำนักทดสอบทางการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน.

_____. (2553). *แนวทางการจัดกิจกรรมการเรียนรู้และการประเมินผลตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551*. สำนักทดสอบทางการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน.

ทิพย์เกสร พรหมภักดี. (2560). *การสร้างระบบสอบวัดมาตรฐานความรู้ทางด้านเทคโนโลยีและสารสนเทศ ในมหาวิทยาลัยรังสิตด้วยทฤษฎีทดสอบแบบมาตรฐานเดิมและทฤษฎีตอบสนองข้อสอบ* (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ, มหาวิทยาลัยรังสิต). คลังปัญญา มหาวิทยาลัยรังสิต.

<https://rsuir-library.rsu.ac.th/handle/123456789/1056>

มะนิวรรณ แก้วดวงดี. (2567). *การศึกษามโนทัศน์ที่คลาดเคลื่อนทางคณิตศาสตร์ เรื่อง เศษส่วน ของนักศึกษาครูประถมศึกษาในเขตภาคใต้ของสาธารณรัฐประชาธิปไตยประชาชนลาว* (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ, มหาวิทยาลัยศรีนครินทรวิโรฒ). iThesis Srinakharinwirot University.

<https://ir-ithesis.swu.ac.th/server/api/core/bitstreams/ea8dcc00-e2b9-4321-ad23-13a38fd11387/content>

- เวชฤทธิ์ อังชนะภัทรขจร. (2546). *การสังเคราะห์งานวิจัยที่เกี่ยวข้องกับมโนทัศน์ที่คลาดเคลื่อนในวิชาคณิตศาสตร์* (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ, จุฬาลงกรณ์มหาวิทยาลัย). คลังปัญญาจุฬาเพื่อประเทศไทย (Chula ETD). <https://digital.car.chula.ac.th/chulaetd/33212/>
- ศิริเดช สุชีวะ. (2538). *การวินิจฉัยมโนทัศน์ที่คลาดเคลื่อน*. ศูนย์ตำราและเอกสารทางวิชาการ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2555). *ทฤษฎีการทดสอบแบบดั้งเดิม* (พิมพ์ครั้งที่ 7). โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- _____. (2555). *ทฤษฎีการตอบสนองข้อสอบ*. โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- _____. (2556). *วิธีการวิทยาการประเมิน: ทางเลือกที่หลากหลาย* (พิมพ์ครั้งที่ 8). โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สิริพร ทิพย์คง. (2545). *การแก้ปัญหาคณิตศาสตร์*. พัฒนาคุณภาพวิชาการ (พว.).
- _____. (2545). *หลักสูตรและการสอนคณิตศาสตร์*. พัฒนาคุณภาพวิชาการ (พว.).
- เยาวดี วิบูลย์ศรี. (2548). *การวัดผลและการสร้างแบบสอบผลสัมฤทธิ์* (พิมพ์ครั้งที่ 4). สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- _____. (2556). *การวัดผลและการสร้างแบบสอบผลสัมฤทธิ์* (พิมพ์ครั้งที่ 10). สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- อนันตริกา เนียมแก้ว. (2565). *การศึกษามโนทัศน์ทางคณิตศาสตร์และยุทธวิธีในการแก้ปัญหาเกี่ยวกับทศนิยมและเศษส่วนของนักเรียนชั้นประถมศึกษาปีที่ 6* (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจไม่ได้ตีพิมพ์, มหาวิทยาลัยราชภัฏพิบูลสงคราม). คลังสารสนเทศดิจิทัลพิบูลสงคราม.
<https://psruir.psru.ac.th/handle/123456789/384>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Baker, F. B. (2001). *The basics of item response theory* (2nd ed.). Heinemann.
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. SAGE Publications.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. SAGE Publications.

- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Routledge.
- Ni, Y., & Zhou, Y.-D. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist*, 40(1), 27–52. https://doi.org/10.1207/s15326985ep4001_3
- Pesman, H., & Eryilmaz, A. (2010). Development of a three-tier test to assess misconceptions about introductory mechanics. *The Journal of Educational Research*, 103(3), 208–222. <https://doi.org/10.1080/00220670903383126>
- Radatz, H. (1979). Error analysis in mathematics education. *Journal for Research in Mathematics Education*, 10(3), 163–172. <https://doi.org/10.2307/748804>
- Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics*, 4(3), 207–230. <https://doi.org/10.2307/1164671>
- Vamvakoussi, X., & Vosniadou, S. (2004). Understanding the structure of the set of rational numbers: A conceptual change approach. *Learning and Instruction*, 14(5), 453–467. <https://doi.org/10.1016/j.learninstruc.2004.06.013>
- Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making sense of word problems*. Swets & Zeitlinger.