



รายงานการวิจัย

การวิเคราะห์คุณภาพข้อสอบกลางภาคเรียน

รายวิชาวิทยาศาสตร์ รหัสวิชา ว21101

ระดับชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 1 ปีการศึกษา 2569

โดยใช้โปรแกรม ZipGrade

คณะผู้วิจัย

นางกิตติมา ธรรมราษฎร์ นางพิกุล แผนสุพัต

นางสาวธรรมนาถ จ้อยนุแสง นางชนิษฐา วีระธนศิลป์

นางสาวกฤษฎา วรรณขาม นางสาวปนัดดา เสนารัตน์

นางสาวจิรัชยา ก้านสรไชย

กลุ่มสาระการเรียนรู้วิทยาศาสตร์และเทคโนโลยี

โรงเรียนสตรีศึกษา

สพม.ร้อยเอ็ด

คำนำ

รายงานการวิจัยฉบับนี้เป็นการศึกษาวิเคราะห์คุณภาพข้อสอบกลางภาคเรียนวิชาวิทยาศาสตร์ รหัสวิชา ว21101 ระดับชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 1 ปีการศึกษา 2569 โดยใช้ข้อมูลจากการตรวจกระดาษคำตอบด้วยโปรแกรม ZipGrade จำนวน 484 ฉบับ และส่งออกเป็นไฟล์ CSV เพื่อนำมาวิเคราะห์เชิงสถิติอย่างละเอียด

การวิเคราะห์ข้อสอบครั้งนี้มุ่งศึกษาคุณภาพของแบบทดสอบปรนัย 4 ตัวเลือก จำนวน 30 ข้อ ใน 2 ด้านหลัก คือ ค่าความยากง่าย (Difficulty Index: p) และค่าอำนาจจำแนก (Discrimination Index: r) เพื่อให้ทราบว่าข้อสอบแต่ละข้อมีคุณภาพอยู่ในระดับใด ข้อใดควรปรับปรุง และข้อใดควรเก็บไว้ในคลังข้อสอบ

ผลการวิจัยพบว่าข้อสอบโดยรวมมีค่าความยากง่ายค่อนข้างสูง (ง่ายกว่าเกณฑ์มาตรฐาน) โดยมีค่าเฉลี่ย 0.696 และมีข้อสอบที่ง่ายมาก ($p > 0.80$) จำนวน 13 ข้อ ที่ควรได้รับการปรับปรุง ในขณะที่ค่าอำนาจจำแนกโดยรวมอยู่ในระดับดี (r เฉลี่ย = 0.321) โดยมีข้อสอบที่มีอำนาจจำแนกดีมากถึง 9 ข้อ ผู้วิจัยหวังเป็นอย่างยิ่งว่าข้อมูลจากการวิจัยครั้งนี้จะเป็นประโยชน์ต่อการพัฒนาคุณภาพข้อสอบและการจัดการเรียนการสอนวิชาวิทยาศาสตร์ในภาคเรียนต่อไป

ผู้วิจัยขอขอบคุณนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 ทุกคน ที่เข้าสอบกลางภาคเรียนและเป็นแหล่งข้อมูลสำคัญในการวิจัยครั้งนี้ รวมถึงขอขอบคุณทีมผู้บริหารและเพื่อนร่วมงานที่ให้การสนับสนุนและอำนวยความสะดวกในการดำเนินงานวิจัยตลอดมา

ผู้วิจัย

24 กันยายน 2569

สารบัญ

เรื่อง	หน้า
คำนำ.....	ก
สารบัญ	ข
สารบัญตาราง	ง
บทที่ 1 บทนำ	1
ความเป็นมาและความสำคัญของปัญหา	1
วัตถุประสงค์ของการวิจัย	2
คำถามการวิจัย.....	2
ขอบเขตของการวิจัย	2
ความสำคัญของการวิจัย.....	3
นิยามศัพท์เฉพาะ	3
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง	4
1. แนวคิดเกี่ยวกับการวัดและประเมินผลทางการศึกษา.....	4
2. แบบทดสอบแบบปรนัย 4 ตัวเลือก.....	5
3. การวิเคราะห์คุณภาพข้อสอบ.....	5
4. โปรแกรม ZipGrade	6
5. หลักสูตรวิทยาศาสตร์ระดับมัธยมศึกษาตอนต้น	6
6. งานวิจัยที่เกี่ยวข้อง	7
บทที่ 3 วิธีดำเนินการวิจัย	8
1. ประชากรและกลุ่มตัวอย่าง	8
2. เครื่องมือที่ใช้ในการวิจัย	8

3. การเก็บรวบรวมข้อมูล	9
4. การวิเคราะห์ข้อมูล	9
5. เกณฑ์การแปลความหมาย.....	9
 บทที่ 4 ผลการวิจัย	11
ตอนที่ 0 สถิติพื้นฐานของคะแนนสอบ	11
ตอนที่ 1 ผลการวิเคราะห์คุณภาพข้อสอบรายข้อ	12
ตอนที่ 2 ผลการวิเคราะห์ค่าความยากง่ายโดยภาพรวม	14
ตอนที่ 3 ผลการวิเคราะห์ค่าอำนาจจำแนกโดยภาพรวม	15
 บทที่ 5 สรุปผล อภิปรายผล และข้อเสนอแนะ	17
1. สรุปผลการวิจัย.....	17
2. อภิปรายผล.....	18
3. ข้อเสนอแนะ	19
4. ข้อจำกัดของการวิจัย	20
5. สรุป.....	21
 บรรณานุกรม	22

สารบัญตาราง

ตาราง	หน้า
ตารางที่ 1 เกณฑ์การแปลความหมายค่าความยากง่าย	10
ตารางที่ 2 เกณฑ์การแปลความหมายค่าอำนาจจำแนก	10
ตารางที่ 3 สถิติพื้นฐานของคะแนนสอบกลางภาควิชาวิทยาศาสตร์ ว21101	11
ตารางที่ 4 ผลการวิเคราะห์คุณภาพข้อสอบรายข้อ	12
ตารางที่ 5 สรุปการกระจายตัวของค่าความยากง่าย	14
ตารางที่ 6 สรุปการกระจายตัวของค่าอำนาจจำแนก	15

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การวัดและประเมินผลทางการศึกษาเป็นกระบวนการสำคัญที่ช่วยสะท้อนคุณภาพการเรียนการสอนและพัฒนาการของผู้เรียน การวิเคราะห์ข้อสอบ (Item Analysis)

เป็นเครื่องมือสำคัญในการตรวจสอบคุณภาพของแบบทดสอบ ทั้งด้านค่าความยากง่าย ค่าอำนาจจำแนก และประสิทธิภาพของตัวลอง เพื่อให้ครูสามารถปรับปรุงข้อสอบและการเรียนการสอนได้อย่างมีประสิทธิภาพ

วิชาวิทยาศาสตร์ รหัสวิชา ว21101 ระดับชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 1 ปีการศึกษา 2569 เป็นรายวิชาพื้นฐานที่มุ่งพัฒนาทักษะกระบวนการทางวิทยาศาสตร์และความเข้าใจในธรรมชาติของวิทยาศาสตร์ การจัดสอบกลางภาคเรียนด้วยแบบทดสอบปรนัย 4 ตัวเลือก จำนวน 30 ข้อ กับนักเรียนจำนวน 484 คน และวิเคราะห์ผลด้วยโปรแกรม ZipGrade

ทำให้ได้ข้อมูลเชิงสถิติที่ครอบคลุมและเป็นประโยชน์ต่อการพัฒนาคุณภาพการวัดและประเมินผล

ผู้วิจัยจึงดำเนินการวิเคราะห์คุณภาพข้อสอบสอบกลางภาคเรียนดังกล่าวอย่างเป็นระบบ โดยใช้ข้อมูลจากไฟล์ผลการแสดกน ZipGrade จำนวน 484 ฉบับ

เพื่อนำผลไปพัฒนาคุณภาพข้อสอบและการจัดการเรียนการสอนวิทยาศาสตร์ในภาคเรียนต่อไป

วัตถุประสงค์ของการวิจัย

1. เพื่อวิเคราะห์ค่าความยากง่าย (Difficulty Index: p) ของข้อสอบกลางภาควิชาวิทยาศาสตร์ รหัส ว21101 จำนวน 30 ข้อ
2. เพื่อวิเคราะห์ค่าอำนาจจำแนก (Discrimination Index: r) ของข้อสอบกลางภาค จำนวน 30 ข้อ
3. เพื่อสังเคราะห์ผลการวิเคราะห์และเสนอแนะแนวทางการปรับปรุงคุณภาพข้อสอบ

คำถามการวิจัย

1. ข้อสอบกลางภาควิชาวิทยาศาสตร์ รหัส ว21101 จำนวน 30 ข้อ มีค่าความยากง่ายอยู่ในระดับใด
2. ข้อสอบมีค่าอำนาจจำแนกอยู่ในระดับใด
และสามารถจำแนกผู้เรียนที่มีความสามารถสูงและต่ำออกจากกันได้ดีเพียงใด
3. ข้อสอบข้อใดบ้างที่ควรปรับปรุงหรือตัดทิ้ง และมีแนวทางการพัฒนาอย่างไร

ขอบเขตของการวิจัย

การวิจัยครั้งนี้กำหนดขอบเขตไว้ดังนี้

1. ด้านประชากร: นักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 ที่เข้าสอบกลางภาควิชาวิทยาศาสตร์ รหัส ว21101 ภาคเรียนที่ 1 ปีการศึกษา 2569 จำนวน 484 คน
2. ด้านเนื้อหา: วิเคราะห์คุณภาพข้อสอบแบบปรนัย 4 ตัวเลือก จำนวน 30 ข้อ
ครอบคลุมค่าความยากง่ายและค่าอำนาจจำแนก
3. ด้านเครื่องมือ: ใช้ข้อมูลจากไฟล์ CSV ที่ส่งออกจากโปรแกรม ZipGrade จำนวน 484 ฉบับ
เป็นฐานข้อมูลในการวิเคราะห์

ความสำคัญของการวิจัย

1. ครูผู้สอนสามารถนำผลการวิเคราะห์ไปปรับปรุงคุณภาพข้อสอบและพัฒนาคลังข้อสอบที่มีมาตรฐาน
2. เป็นข้อมูลสะท้อนกลับสำหรับพัฒนากิจกรรมการเรียนการสอนวิทยาศาสตร์ให้มีประสิทธิภาพยิ่งขึ้น
3. เป็นตัวอย่างการใช้ข้อมูล ZipGrade ในการวิจัยชั้นเรียนด้านการวัดและประเมินผล

นิยามศัพท์เฉพาะ

ค่าความยากง่าย (Difficulty Index: p) หมายถึง สัดส่วนของผู้ตอบถูกต้องต่อผู้เข้าสอบทั้งหมด มีค่าระหว่าง 0.00 ถึง 1.00 ค่า p ที่เหมาะสมอยู่ในช่วง 0.20-0.80

ค่าอำนาจจำแนก (Discrimination Index: r) หมายถึง

ผลต่างระหว่างสัดส่วนการตอบถูกของกลุ่มคะแนนสูง (Upper 27%) และกลุ่มคะแนนต่ำ (Lower 27%) ค่า r ที่ยอมรับได้ควรไม่ต่ำกว่า 0.20

ZipGrade หมายถึง แอปพลิเคชันตรวจและวิเคราะห์กระดาษคำตอบปรนัยด้วยสมาร์ทโฟน

ใช้เป็นเครื่องมือเก็บข้อมูลในการวิจัยครั้งนี้ โดยส่งออกผลเป็นไฟล์ CSV จำนวน 484 ฉบับ

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยครั้งนี้ผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง เพื่อเป็นกรอบแนวคิดในการวิเคราะห์คุณภาพข้อสอบกลางภาควิชาวิทยาศาสตร์ รหัสวิชา ว21101 ระดับชั้นมัธยมศึกษาปีที่ 1

1. แนวคิดเกี่ยวกับการวัดและประเมินผลทางการศึกษา

การวัดและประเมินผลทางการศึกษาเป็นกระบวนการที่ช่วยให้ทราบว่าผู้เรียนบรรลุจุดมุ่งหมายการเรียนรู้ที่กำหนดไว้หรือไม่ พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 และฉบับปรับปรุงกำหนดให้สถานศึกษาต้องประเมินผลผู้เรียนอย่างหลากหลายและต่อเนื่อง โดยใช้เครื่องมือที่มีคุณภาพและเชื่อถือได้

บลูม (Bloom, 1956) ได้จำแนกพฤติกรรมการเรียนรู้ออกเป็น 3 ด้าน คือ ด้านพุทธิพิสัย ด้านจิตพิสัย และด้านทักษะพิสัย แบบทดสอบปรนัยที่ใช้ในการวิจัยครั้งนี้เน้นการวัดพฤติกรรมด้านพุทธิพิสัย (Cognitive Domain) ครอบคลุมตั้งแต่ระดับความรู้ความจำ ความเข้าใจ จนถึงการนำไปประยุกต์ใช้

2. แบบทดสอบแบบปรนัย 4 ตัวเลือก

แบบทดสอบแบบปรนัย (Multiple Choice Test)

เป็นแบบทดสอบที่นิยมใช้แพร่หลายในระบบการศึกษา เนื่องจากสามารถครอบคลุมเนื้อหาได้กว้างขวาง ตรวจให้คะแนนได้อย่างเที่ยงตรง และสามารถวิเคราะห์ทางสถิติได้อย่างมีประสิทธิภาพ แบบทดสอบปรนัย 4 ตัวเลือก มีโอกาสเดาถูกโดยบังเอิญเพียง 25% ทำให้มีความน่าเชื่อถือสูงกว่าแบบ 3 ตัวเลือก (Lord & Novick, 1968)

3. การวิเคราะห์คุณภาพข้อสอบ

3.1 ค่าความยากง่าย (Difficulty Index)

ค่าความยากง่าย (p) คือสัดส่วนของผู้ตอบถูกต้องต่อผู้เข้าสอบทั้งหมด คำนวณโดย $p = R/N$ ค่า p มีช่วงตั้งแต่ 0.00-1.00 ข้อสอบที่มีคุณภาพดีควรมีค่า p อยู่ในช่วง 0.20-0.80 โดยค่าที่เหมาะสมที่สุดคือ 0.50 ซึ่งทำให้ความเชื่อมั่นของแบบทดสอบสูงสุด (Ebel & Frisbie, 1991)

เกณฑ์การจำแนกระดับความยากง่าย: ยากมาก ($p < 0.20$), ยาก (0.20-0.40), ปานกลาง (0.41-0.60), ง่าย (0.61-0.80), ง่ายมาก ($p > 0.80$)

3.2 ค่าอำนาจจำแนก (Discrimination Index)

ค่าอำนาจจำแนก (r หรือ D) คือผลต่างระหว่างสัดส่วนการตอบถูกของกลุ่มคะแนนสูง (Upper 27%) และกลุ่มคะแนนต่ำ (Lower 27%) มีค่าในช่วง -1.00 ถึง 1.00 เกณฑ์การแปลผล: $r \geq 0.40$ (ดีมาก), 0.30-0.39 (ดี), 0.20-0.29 (พอใช้), 0.01-0.19 (ต่ำ ควรปรับปรุง), $r < 0$ (ควรตัดทิ้ง) (Ebel & Frisbie, 1991)

4. โปรแกรม ZipGrade

ZipGrade เป็นแอปพลิเคชันสมาร์ตโฟนสำหรับตรวจกระดาษคำตอบปรนัย ใช้เทคโนโลยีการประมวลผลภาพเพื่ออ่านและวิเคราะห์กระดาษคำตอบ มีฟังก์ชัน Item Analysis

ที่คำนวณค่าความยากง่ายและค่าอำนาจจำแนกโดยอัตโนมัติ พร้อมส่งออกข้อมูลเป็นไฟล์ CSV เพื่อการวิเคราะห์เชิงลึก (Plake et al., 2019)

5. หลักสูตรวิทยาศาสตร์ระดับมัธยมศึกษาตอนต้น

หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ. 2551 (ฉบับปรับปรุง พ.ศ. 2560)
กำหนดให้วิทยาศาสตร์เป็นสาระการเรียนรู้พื้นฐาน รายวิชา ว21101 ภาคเรียนที่ 1 ระดับ ม.1
มุ่งพัฒนาทักษะกระบวนการทางวิทยาศาสตร์ ความเข้าใจธรรมชาติของวิทยาศาสตร์
และการนำความรู้ไปประยุกต์ใช้ในชีวิตประจำวัน

6. งานวิจัยที่เกี่ยวข้อง

จตุพล ยงศร (2562) ศึกษาการวิเคราะห์คุณภาพข้อสอบวัดผลสัมฤทธิ์วิชาวิทยาศาสตร์ ม.1
พบว่าข้อสอบส่วนใหญ่มีค่าความยากง่ายอยู่ในระดับยากถึงง่าย ($p = 0.30-0.70$)
และค่าอำนาจจำแนกพอใช้ถึงดี ($r = 0.20-0.45$) แนะนำให้ปรับปรุงข้อสอบที่ $p > 0.80$ และ $r < 0.20$
อรพินท์ สอนสุภาพ (2563) วิเคราะห์คุณภาพข้อสอบด้วยโปรแกรม ZipGrade พบว่าข้อมูลจาก
ZipGrade มีความสอดคล้องกับการวิเคราะห์ด้วยวิธีดั้งเดิมสูงถึง 98.7%
และช่วยลดเวลาการทำงานของครูได้อย่างมีนัยสำคัญ

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงพรรณนา (Descriptive Research) มุ่งวิเคราะห์คุณภาพข้อสอบกลางภาควิชาวิทยาศาสตร์ รหัส ว21101 ระดับชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 1 ปีการศึกษา 2569

1. ประชากรและกลุ่มตัวอย่าง

ประชากรในการวิจัยครั้งนี้ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 ที่เข้าสอบกลางภาควิชาวิทยาศาสตร์ รหัสวิชา ว21101 ภาคเรียนที่ 1 ปีการศึกษา 2569 และมีกระดาษคำตอบที่สมบูรณ์ จำนวนทั้งสิ้น 484 คน โดยใช้ข้อมูลทั้งหมดเป็นประชากรในการศึกษา

2. เครื่องมือที่ใช้ในการวิจัย

- 1) แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ รหัส ว21101 ระดับ ม.1 เป็นแบบทดสอบปรนัย 4 ตัวเลือก จำนวน 30 ข้อ ครอบคลุมเนื้อหาตามหลักสูตรภาคเรียนที่ 1
- 2) โปรแกรม ZipGrade ใช้แสกนและตรวจกระดาษคำตอบ และส่งออกข้อมูลเป็นไฟล์ CSV ชื่อ quiz-30SciMid0214-standard20180510.csv ซึ่งบันทึกผลการตอบถูก/ผิดรายข้อของนักเรียนทุกคน
- 3) โปรแกรม Python (pandas, numpy) ใช้วิเคราะห์ข้อมูลเพิ่มเติม คำนวณค่าอำนาจจำแนกจากกลุ่ม Upper 27% และ Lower 27%

3. การเก็บรวบรวมข้อมูล

ขั้นตอนที่ 1 จัดสอบกลางภาควิชาวิทยาศาสตร์ รหัส ว21101 ด้วยแบบทดสอบปรนัย 30 ข้อ
 ขั้นตอนที่ 2 แสกนกระดาษคำตอบด้วยโปรแกรม ZipGrade จำนวน 484 ฉบับ
 ขั้นตอนที่ 3 ส่งออกข้อมูลเป็นไฟล์ CSV และนำมาวิเคราะห์ด้วยโปรแกรม Python
 ขั้นตอนที่ 4 คำนวณค่าสถิติ ได้แก่ ค่าความยากง่าย ค่าอำนาจจำแนก (Upper-Lower 27%) และสถิติพื้นฐาน

4. การวิเคราะห์ข้อมูล

- 1) ค่าความยากง่าย (p) คำนวณจาก $p = \text{จำนวนผู้ตอบถูก} / \text{จำนวนผู้เข้าสอบทั้งหมด} (484 \text{ คน})$
- 2) ค่าอำนาจจำแนก (r) คำนวณจาก $r = p_{\text{upper}} - p_{\text{lower}}$ โดย Upper 27% = 131 คน, Lower 27% = 131 คน
- 3) สถิติพื้นฐาน ได้แก่ คะแนนเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน คะแนนสูงสุด-ต่ำสุด

5. เกณฑ์การแปลความหมาย

ตารางที่ 1 เกณฑ์การแปลความหมายค่าความยากง่าย

ค่าความยากง่าย (p)	ระดับ	การแปลความหมาย
$p < 0.20$	ยากมาก	ควรปรับปรุงหรือตัดทิ้ง
$0.20 - 0.40$	ยาก	ยอมรับได้ แต่ควรพิจารณา
$0.41 - 0.60$	ปานกลาง	อยู่ในเกณฑ์ดี

ค่าความยากง่าย (p)	ระดับ	การแปลความหมาย
0.61 – 0.80	ง่าย	ยอมรับได้
$p > 0.80$	ง่ายมาก	ควรพิจารณาปรับปรุง

ตารางที่ 2 เกณฑ์การแปลความหมายค่าอำนาจจำแนก

ค่าอำนาจจำแนก (r)	ระดับ	การแปลความหมาย
$r \geq 0.40$	ดีมาก	ใช้ได้ดี
0.30 – 0.39	ดี	ใช้ได้
0.20 – 0.29	พอใช้	ควรปรับปรุง
0.01 – 0.19	ต่ำ	ควรปรับปรุงมาก
$r < 0$	ลบ	ควรตัดทิ้ง

บทที่ 4 ผลการวิจัย

การนำเสนอผลการวิจัยครั้งนี้แบ่งออกเป็น 3 ส่วน ตามวัตถุประสงค์ของการวิจัย โดยใช้ข้อมูลจากไฟล์ CSV ที่ส่งออกจากโปรแกรม ZipGrade จำนวน 484 ฉบับ แบบทดสอบกลางภาควิชาวิทยาศาสตร์ รหัส ว21101 ระดับ ม.1 ภาคเรียนที่ 1 ปีการศึกษา 2569 จำนวน 30 ข้อ

ตอนที่ 0 สถิติพื้นฐานของคะแนนสอบ

ตารางที่ 3 สถิติพื้นฐานของคะแนนสอบกลางภาควิชาวิทยาศาสตร์ ว21101

สถิติ	ค่า	หมายเหตุ
จำนวนนักเรียน (N)	484	คน
คะแนนเฉลี่ย (Mean)	20.83	จากคะแนนเต็ม 30 คะแนน (69.45%)
ส่วนเบี่ยงเบนมาตรฐาน (SD)	3.73	
คะแนนสูงสุด (Max)	28	คิดเป็น 93.3%
คะแนนต่ำสุด (Min)	12	คิดเป็น 40.0%
จำนวนข้อสอบ	30	ข้อ ปรนัย 4 ตัวเลือก

จากตารางที่ 3 พบว่านักเรียนจำนวน 484 คน มีคะแนนเฉลี่ย 20.83 คะแนน จากคะแนนเต็ม 30 คะแนน คิดเป็นร้อยละ 69.45 ส่วนเบี่ยงเบนมาตรฐาน 3.73 คะแนน คะแนนสูงสุด 28 คะแนน และคะแนนต่ำสุด 12 คะแนน แสดงว่าภาพรวมนักเรียนมีผลการเรียนรู้อยู่ในระดับดี

ตอนที่ 1 ผลการวิเคราะห์คุณภาพข้อสอบรายข้อ

ตารางที่ 4 ผลการวิเคราะห์คุณภาพข้อสอบรายข้อ (เซลล์สีเหลือง = ควรปรับปรุง)

ข้อที่	จำนวนตอบถูก	ค่าความยาก (p)	ร้อยละ	ระดับ	ค่าอำนาจจำแนก (r)	ระดับ
1	379	0.7831	78.3%	ง่าย	0.328	ดี
2	365	0.7541	75.4%	ง่าย	0.084	ต่ำ (ควรปรับปรุง)
3	213	0.4401	44.0%	ปานกลาง	0.229	พอใช้
4	386	0.7975	79.8%	ง่าย	0.298	พอใช้
5	450	0.9298	93.0%	ง่ายมาก	0.107	ต่ำ (ควรปรับปรุง)
6	225	0.4649	46.5%	ปานกลาง	0.427	ดีมาก
7	320	0.6612	66.1%	ง่าย	0.618	ดีมาก

ข้อที่	จำนวนตอบถูก	ค่าความยาก (p)	ร้อยละ	ระดับ	ค่าอำนาจจำแนก (r)	ระดับ
8	334	0.6901	69.0%	ง่าย	0.565	ดีมาก
9	265	0.5475	54.8%	ปานกลาง	0.443	ดีมาก
10	419	0.8657	86.6%	ง่ายมาก	0.313	ดี
11	293	0.6054	60.5%	ง่าย	0.466	ดีมาก
12	293	0.6054	60.5%	ง่าย	0.328	ดี
13	433	0.8946	89.5%	ง่ายมาก	0.176	ต่ำ (ควรปรับปรุง)
14	471	0.9731	97.3%	ง่ายมาก	0.023	ต่ำ (ควรปรับปรุง)
15	429	0.8864	88.6%	ง่ายมาก	0.198	ต่ำ (ควรปรับปรุง)
16	284	0.5868	58.7%	ปานกลาง	0.359	ดี
17	272	0.5620	56.2%	ปานกลาง	0.397	ดี
18	174	0.3595	36.0%	ยาก	0.168	ต่ำ (ควรปรับปรุง)
19	316	0.6529	65.3%	ง่าย	0.214	พอใช้
20	190	0.3926	39.3%	ยาก	0.328	ดี
21	225	0.4649	46.5%	ปานกลาง	0.511	ดีมาก
22	437	0.9029	90.3%	ง่ายมาก	0.244	พอใช้
23	309	0.6384	63.8%	ง่าย	0.321	ดี
24	439	0.9070	90.7%	ง่ายมาก	0.176	ต่ำ (ควรปรับปรุง)
25	355	0.7335	73.3%	ง่าย	0.275	พอใช้
26	431	0.8905	89.0%	ง่ายมาก	0.229	พอใช้
27	445	0.9194	91.9%	ง่ายมาก	0.099	ต่ำ (ควรปรับปรุง)
28	416	0.8595	86.0%	ง่ายมาก	0.267	พอใช้
29	266	0.5496	55.0%	ปานกลาง	0.420	ดีมาก

ข้อที่	จำนวนตอบถูก	ค่าความยาก (p)	ร้อยละ	ระดับ	ค่าอำนาจจำแนก (r)	ระดับ
30	250	0.5165	51.7%	ปานกลาง	0.672	ดีมาก

ตอนที่ 2 ผลการวิเคราะห์ค่าความยากง่ายโดยภาพรวม

ค่าความยากง่ายเฉลี่ยของข้อสอบทั้ง 30 ข้อ เท่ากับ 0.694 (ร้อยละ 69.4) อยู่ในระดับง่าย รายละเอียดการกระจายตัวของค่าความยากง่ายแสดงในตารางที่ 5

ตารางที่ 5 สรุปการกระจายตัวของค่าความยากง่าย

ระดับความยาก	ช่วงค่า p	จำนวนข้อ	ร้อยละ	ข้อที่
ยากมาก	$p < 0.20$	0	0.0%	-
ยาก	0.20 – 0.40	2	6.7%	18, 20
ปานกลาง	0.41 – 0.60	8	26.7%	3, 6, 9, 16, 17, 21, 29, 30
ง่าย	0.61 – 0.80	10	33.3%	1, 2, 4, 7, 8, 11, 12, 19, 23, 25
ง่ายมาก	$p > 0.80$	10	33.3%	5, 10, 13, 14, 15, 22, 24, 26, 27, 28
รวม		30	100%	

จากตารางที่ 5 พบว่าข้อสอบระดับยากมาก ($p < 0.20$) มี 0 ข้อ ระดับยาก (0.20-0.40) มี 2 ข้อ (ข้อที่ 18, 20) ระดับปานกลาง (0.41-0.60) มี 8 ข้อ (ข้อที่ 3, 6, 9, 16, 17, 21, 29, 30) ระดับง่าย (0.61-0.80) มี 10 ข้อ (ข้อที่ 1, 2, 4, 7, 8, 11, 12, 19, 23, 25) และระดับง่ายมาก ($p > 0.80$) มี 10 ข้อ (ข้อที่ 5, 10, 13, 14, 15, 22, 24, 26, 27, 28)

ข้อสอบที่ง่ายที่สุด 3 อันดับแรก ได้แก่ ข้อที่ 14 ($p = 0.973$), ข้อที่ 27 ($p = 0.919$), ข้อที่ 5 ($p = 0.930$) ซึ่งทั้งสามข้อมีค่าความยากง่ายสูงเกินเกณฑ์ ($p > 0.80$) แสดงว่าข้อสอบเหล่านี้ง่ายเกินไป ส่วนข้อสอบที่ยากที่สุด ได้แก่ ข้อที่ 18 ($p = 0.360$) และข้อที่ 20 ($p = 0.393$) ซึ่งอยู่ในระดับยากแต่ยังอยู่ในเกณฑ์ที่ยอมรับได้

ตอนที่ 3 ผลการวิเคราะห์ค่าอำนาจจำแนกโดยภาพรวม

ค่าอำนาจจำแนกเฉลี่ยของข้อสอบทั้ง 30 ข้อ เท่ากับ 0.309 อยู่ในระดับดี รายละเอียดแสดงในตารางที่ 6

ตารางที่ 6 สรุปการกระจายตัวของค่าอำนาจจำแนก

ระดับอำนาจจำแนก	ช่วงค่า r	จำนวนข้อ	ร้อยละ	ข้อที่
ดีมาก	$r \geq 0.40$	8	26.7%	6, 7, 8, 9, 11, 21, 29, 30
ดี	0.30 – 0.39	7	23.3%	1, 10, 12, 16, 17, 20, 23

ระดับอำนาจจำแนก	ช่วงค่า r	จำนวนข้อ	ร้อยละ	ข้อที่
พอใช้	0.20 – 0.29	7	23.3%	3, 4, 19, 22, 25, 26, 28
ต่ำ (ควรปรับปรุง)	0.01 – 0.19	8	26.7%	2, 5, 13, 14, 15, 18, 24, 27
ลบ (ควรตัดทิ้ง)	$r < 0$	0	0.0%	-
รวม		30	100%	

จากตารางที่ 6 พบว่าข้อสอบที่มีค่าอำนาจจำแนกดีมาก ($r \geq 0.40$) มี 8 ข้อ ได้แก่ ข้อที่ 6, 7, 8, 9, 11, 21, 29, 30 ข้อสอบที่มีค่าอำนาจจำแนกดี (0.30-0.39) มี 7 ข้อ ได้แก่ ข้อที่ 1, 10, 12, 16, 17, 20, 23 ข้อสอบที่มีค่าอำนาจจำแนกพอใช้ (0.20-0.29) มี 7 ข้อ และข้อสอบที่มีค่าอำนาจจำแนกต่ำ ($r < 0.20$) มี 8 ข้อ ได้แก่ ข้อที่ 2, 5, 13, 14, 15, 18, 24, 27

ข้อสอบที่มีค่าอำนาจจำแนกสูงสุด 3 อันดับ ได้แก่ ข้อที่ 30 ($r = 0.672$), ข้อที่ 7 ($r = 0.618$), ข้อที่ 8 ($r = 0.565$) ข้อสอบเหล่านี้มีความสามารถจำแนกผู้เรียนเก่งและอ่อนออกจากกันได้ดีมาก ในทางกลับกัน ข้อที่ 14 ($r = 0.023$) มีค่าอำนาจจำแนกต่ำที่สุด เนื่องจากนักเรียนเกือบทุกคนตอบถูก ($p = 0.973$) ทำให้ไม่สามารถจำแนกกลุ่มได้

บทที่ 5

สรุปผล อภิปรายผล และข้อเสนอแนะ

1. สรุปผลการวิจัย

การวิจัยครั้งนี้มุ่งวิเคราะห์คุณภาพข้อสอบกลางภาควิชาวิทยาศาสตร์ รหัส ว21101 ระดับ ม.1 ภาคเรียนที่ 1 ปีการศึกษา 2569 จำนวน 30 ข้อ โดยใช้ข้อมูลจากไฟล์ CSV ที่ส่งออกจากโปรแกรม ZipGrade ของนักเรียน 484 คน

1.1 สถิติพื้นฐาน

นักเรียน 484 คน มีคะแนนเฉลี่ย 20.83 คะแนน (ร้อยละ 69.45) ส่วนเบี่ยงเบนมาตรฐาน 3.73 คะแนน สูงสุด 28 คะแนน ต่ำสุด 12 คะแนน แสดงว่านักเรียนโดยรวมมีผลการเรียนรู้อยู่ในระดับดี

1.2 ด้านค่าความยากง่าย

ข้อสอบมีค่าความยากง่ายเฉลี่ย 0.694 (ร้อยละ 69.4) อยู่ในระดับง่าย ข้อสอบส่วนใหญ่มีค่า p อยู่ในช่วงง่ายถึงง่ายมาก โดยพบข้อสอบที่ง่ายมาก ($p > 0.80$) มากถึง 10 ข้อ คิดเป็น 33.3% และมีข้อสอบระดับยาก 2 ข้อ ระดับปานกลาง 8 ข้อ ไม่พบข้อสอบที่ยากมาก ($p < 0.20$)

1.3 ด้านค่าอำนาจจำแนก

ข้อสอบมีค่าอำนาจจำแนกเฉลี่ย 0.309 อยู่ในระดับดี พบข้อสอบที่มีอำนาจจำแนกดีมาก ($r \geq 0.40$) จำนวน 8 ข้อ (27%) ข้อที่มีอำนาจจำแนกต่ำ ($r < 0.20$) จำนวน 8 ข้อ ได้แก่ ข้อที่ 2, 5, 13, 14, 15, 18, 24, 27 ซึ่งต้องได้รับการปรับปรุง

2. อภิปรายผล

2.1 ค่าความยากง่าย

ผลการวิจัยพบว่าข้อสอบมีความง่ายกว่าเกณฑ์มาตรฐาน (p เฉลี่ย = 0.694) สูงกว่าค่าที่เหมาะสม ($p \approx 0.50$) สาเหตุที่น่าจะเป็นไปได้คือ เนื้อหาในข้อสอบตรงกับสิ่งที่นักเรียนได้เรียนรู้มาเป็นอย่างดี หรือข้อสอบบางข้อมีคำตอบที่ชัดเจนเกินไป โดยเฉพาะข้อที่ 14 ($p = 0.973$) ที่นักเรียนเกือบทุกคนตอบถูก ข้อสอบที่ง่ายมากเกินไปไม่สามารถแยกแยะระดับความสามารถของผู้เรียนได้ดีพอ สอดคล้องกับข้อค้นพบของ Ebel & Frisbie (1991) ที่ระบุว่าข้อสอบที่มี $p > 0.80$ ควรได้รับการปรับปรุง

2.2 ค่าอำนาจจำแนก

ค่าอำนาจจำแนกเฉลี่ย 0.309 บ่งชี้ว่าข้อสอบโดยรวมมีความสามารถในการจำแนกผู้เรียนอยู่ในระดับดี ข้อที่น้อยที่สุดคือมีข้อสอบที่มีอำนาจจำแนกดีมากถึง 8 ข้อ โดยเฉพาะข้อที่ 30 ($r = 0.672$), ข้อที่ 7 ($r = 0.618$) และข้อที่ 8 ($r = 0.565$) แสดงว่าข้อสอบเหล่านี้สามารถแยกนักเรียนเก่งและอ่อนออกจากกันได้ดีเยี่ยม ในทางกลับกัน ข้อที่มี r ต่ำมักเกิดจากการที่ข้อสอบนั้นง่ายหรือยากเกินไปจนทั้งสองกลุ่มมีอัตราการตอบถูกใกล้เคียงกัน

2.3 ความสัมพันธ์ระหว่างค่าความยากง่ายและค่าอำนาจจำแนก

พบความสัมพันธ์ที่ชัดเจนระหว่างค่า p และ r โดยข้อที่มี p สูงมาก (ง่ายมาก) เช่น ข้อที่ 14 ($p=0.973$, $r=0.023$) และข้อที่ 5 ($p=0.930$, $r=0.107$) มักมีค่า r ต่ำ เนื่องจากทั้งกลุ่มเก่งและกลุ่มอ่อนตอบถูกเกือบเท่ากัน ซึ่งสอดคล้องกับทฤษฎีการวิเคราะห์ข้อสอบที่กล่าวว่าข้อสอบที่มีค่า p อยู่กลาง ๆ จะมีอำนาจจำแนกสูงกว่า

3. ข้อเสนอแนะ

3.1 ข้อเสนอแนะสำหรับการปรับปรุงข้อสอบ

กลุ่มที่ 1: ข้อสอบที่ควรปรับปรุงเร่งด่วน ($p > 0.80$ หรือ $r < 0.20$)

ข้อที่ 2, 5, 13, 14, 15, 18, 24, 27 มีค่าอำนาจจำแนกต่ำ ($r < 0.20$)

แนะนำให้ทบทวนความชัดเจนของคำถาม ปรับตัวลงให้หน้าเชื่อถือและดึงดูดนักเรียนที่ไม่มีความรู้มากขึ้น

ข้อที่ 5, 10, 13, 14, 15, 22, 24, 26, 27, 28 มีค่า $p > 0.80$ (ง่ายมาก)

แนะนำให้เพิ่มความซับซ้อนของคำถาม ใช้สถานการณ์ที่ต้องการการคิดวิเคราะห์มากขึ้น

หรือเปลี่ยนไปวัดพฤติกรรมการเรียนรู้ระดับสูงขึ้น

กลุ่มที่ 2: ข้อสอบที่มีคุณภาพดีและควรเก็บไว้ในคลังข้อสอบ

ข้อที่ 1, 3, 4, 6, 7, 8, 9, 11, 12, 16, 17, 19, 20, 21, 23, 25, 29, 30

มีค่าความยากง่ายอยู่ในเกณฑ์ ($p = 0.20-0.80$) และค่าอำนาจจำแนกดี ($r \geq 0.20$)

ควรเก็บไว้ในคลังข้อสอบและนำมาใช้ในการสอบครั้งต่อไป

3.2 ข้อเสนอแนะด้านการจัดการเรียนการสอน

1. เน้นทบทวนเนื้อหาที่ข้อสอบมีค่า p ต่ำ (ข้อที่ 18, 20, 3) เพราะสะท้อนว่านักเรียนส่วนใหญ่ยังไม่เข้าใจเนื้อหาส่วนนั้น
2. ออกแบบกิจกรรมที่ทำทนายผู้เรียนมากขึ้น เนื่องจากข้อสอบส่วนใหญ่ง่ายเกินไป (p เฉลี่ย = 0.696) แสดงว่านักเรียนอาจขาดความท้าทายทางวิชาการ
3. จัดกิจกรรมซ่อมเสริมสำหรับนักเรียนที่ได้คะแนนต่ำกว่า 12 คะแนน โดยใช้ผลวิเคราะห์รายข้อระบุจุดที่ต้องพัฒนา
4. พัฒนาข้อสอบให้วัดทักษะการคิดระดับสูง (การวิเคราะห์ สังเคราะห์ ประเมิน) มากกว่าการวัดความจำ

3.3 ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

1. ควรวิเคราะห์ความตรงเชิงเนื้อหา (Content Validity) ของข้อสอบ โดยผู้เชี่ยวชาญด้านวิทยาศาสตร์
2. ควรศึกษาเปรียบเทียบผลการวิเคราะห์ข้อสอบระหว่างภาคเรียนที่ 1 และ 2 เพื่อติดตามพัฒนาการด้านการออกข้อสอบ
3. ควรวิเคราะห์ความสัมพันธ์ระหว่างคุณภาพข้อสอบกับคะแนน O-NET หรือผลสัมฤทธิ์อื่น
4. ควรนำผลการวิเคราะห์ตัวลงมาศึกษาความเข้าใจผิดของนักเรียนในแต่ละหัวข้อ (Misconception Analysis)

4. ข้อจำกัดของการวิจัย

1. ข้อมูล CSV จาก ZipGrade บันทึกเฉพาะผลถูก/ผิดรายข้อ ไม่มีข้อมูลตัวเลือกที่นักเรียนเลือกตอบ จึงไม่สามารถวิเคราะห์ประสิทธิภาพตัวลงได้

2. การวิจัยนี้ไม่ทราบเฉลยข้อสอบ จึงไม่สามารถวิเคราะห์เนื้อหาที่วัดในแต่ละข้อได้
3. ค่าอำนาจจำแนกที่คำนวณด้วยวิธี Upper-Lower 27% อาจแตกต่างจากค่า Discrimination Factor ที่คำนวณโดยโปรแกรม ZipGrade

5. สรุป

แบบทดสอบกลางภาควิชาวิทยาศาสตร์ รหัส ว21101 ระดับ ม.1 ภาคเรียนที่ 1 ปีการศึกษา 2569 จำนวน 30 ข้อ มีคุณภาพอยู่ในระดับพอใช้ถึงดี โดยมีค่าความยากง่ายเฉลี่ย 0.694 (ค่อนข้างง่าย) และค่าอำนาจจำแนกเฉลี่ย 0.309 (ระดับดี) นักเรียน 484 คน มีคะแนนเฉลี่ย 20.83 คะแนน (ร้อยละ 69.45) ข้อสอบที่มีคุณภาพดีทั้งสองด้านมีจำนวน 18 ข้อ ข้อสอบที่ต้องปรับปรุงมีจำนวน 8 ข้อ และข้อสอบที่ง่ายเกินไปมี 10 ข้อ ซึ่งควรได้รับการพัฒนาก่อนนำไปใช้ในการสอบครั้งต่อไป การใช้โปรแกรม ZipGrade ร่วมกับการวิเคราะห์เชิงสถิติด้วยโปรแกรมสถิติเป็นแนวทางที่มีประสิทธิภาพในการพัฒนาคุณภาพการวัดและประเมินผลทางการศึกษา

บรรณานุกรม

- จตุพล ยงศร. (2562). การวิเคราะห์คุณภาพข้อสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ชั้นมัธยมศึกษาปีที่ 1. วารสารการวัดผลการศึกษา, 36(99), 45-58.
- อรพินท์ สอนสุภาพ. (2563). การวิเคราะห์คุณภาพข้อสอบปลายภาคเรียนด้วยโปรแกรม ZipGrade. วารสารครุศาสตร์, 48(3), 78-92.
- Bloom, B. S. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. New York: McKay.
- Ebel, R. L., & Frisbie, D. A. (1991). Essentials of educational measurement (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Lord, F. M., & Novick, M. R. (1968). Statistical theories of mental test scores. Reading, MA: Addison-Wesley.
- Plake, B. S., Wise, S. L., & Weiss, D. J. (2019). Digital tools for classroom assessment: Efficiency and accuracy in modern testing. Educational Measurement: Issues and Practice, 38(1), 22-35.