

การสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน ตามรูปแบบที่มีอัตราส่วนของจำนวนข้อสอบ แบบเลือกตอบต่างกัน กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6

วนิดา รวมธรรม¹

สมนึก กุฬพิทยะ²

สุรทิน นาราริณเย³

บทคัดย่อ

แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนจะมีคุณภาพได้ต้องประกอบด้วยหลายองค์ประกอบ และที่สำคัญยิ่งคือคุณภาพเกี่ยวกับ ความเที่ยงตรง (Validity) ความเชื่อมั่น (Reliability) ความเป็นปรนัย (Objectivity) อำนาจจำแนก (Discrimination) และความยาก (Difficulty) ถ้าแบบทดสอบที่สร้างขึ้นนั้นมีคุณภาพตามหลักวิชาการแล้ว การวัดผลจะมีประสิทธิภาพและเป็นประโยชน์ต่อการจัดการเรียนการสอนยิ่งขึ้น การวิจัยครั้งนี้จึงมีความมุ่งหมายเพื่อสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนตามรูปแบบของข้อสอบแบบเลือกตอบที่มีอัตราส่วนของจำนวนข้อสอบต่างกัน กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ซึ่งเป็นแบบทดสอบแบบเลือกตอบชนิด 4 ตัวเลือก ที่มีอัตราส่วนของประเภทคำถามเดียว ตัวเลือกคงที่ และแบบสถานการณ์ แตกต่างกันจำนวน 4 ฉบับ กลุ่มตัวอย่างประกอบด้วย นักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2552 สังกัดสำนักงานเขตพื้นที่การศึกษาอุดรธานี เขต 3 จำนวน 800 คน จากโรงเรียน 25 โรงเรียน ซึ่งได้มาโดยการสุ่มแบบหลายขั้นตอน (Multi – Stage Random Sampling) เพื่อใช้ทดสอบ และนำคะแนนมาสร้างเกณฑ์ปกติ ในรูปคะแนนที่-ปกติ ส่วนการขยายที่-ปกติ อาศัยการสร้างสมการพหุคูณ เครื่องมือที่ใช้เก็บรวบรวมข้อมูลคือ แบบทดสอบชนิดเลือกตอบ 4 ตัวเลือก จำนวน 4 ฉบับ ฉบับละ 40 ข้อ ที่มีอัตราส่วนของรูปแบบคำถามเดียว : ตัวเลือกคงที่ : แบบสถานการณ์ เท่ากับ 32 : 4 : 4, 24 : 8 : 8, 16 : 12 : 12 และ 16 : 8 : 16 ตามลำดับ ทั้งนี้ได้ปรับมาจากแบบทดสอบฉบับพื้นฐาน ชนิดเลือกตอบ 4 ตัวเลือก ประเภทคำถามเดียว จำนวน 45 ข้อ

ผลการวิจัยปรากฏดังนี้

1. ค่าความยากเฉลี่ยของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนตามรูปแบบของข้อสอบแบบเลือกตอบที่มีอัตราส่วนของจำนวนข้อสอบต่างกัน กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ฉบับที่ 1 มีค่าสูงสุด (0.45) รองลงมาได้แก่ ฉบับที่ 2 (0.44) ฉบับที่ 3 (0.43) และฉบับที่ 4 (0.40) ซึ่งค่าความยากเฉลี่ยของแบบทดสอบทั้ง 4 ฉบับ มีค่าตั้งแต่ 0.40 ถึง 0.45 ถือว่ามีค่าความยากปานกลาง เมื่อทดสอบความแตกต่างของค่าความยากเฉลี่ย พบว่าไม่แตกต่างกัน

¹มหาวิทยาลัยราชภัฏวไลยอลงกรณ์ จ.อยุธยา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

²อาจารย์ ประจำภาควิชาวิจัยและพัฒนาศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

³อาจารย์ ประจำคณะวิทยาศาสตร์ มหาวิทยาลัยมหาสารคาม

2. ค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน ตามรูปแบบของข้อสอบแบบเลือกตอบ ที่มีอัตราส่วนของจำนวนข้อสอบต่างกัน กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ฉบับที่ 2 มีค่าสูงสุด (0.39) รองลงมาได้แก่ ฉบับที่ 1 (0.38) ฉบับที่ 3 (0.37) และฉบับที่ 4 (0.36) ซึ่งค่าอำนาจเฉลี่ยของแบบทดสอบทั้ง 4 ฉบับ มีค่าตั้งแต่ 0.36 ถึง 0.39 ถือว่าสามารถจำแนกได้ เมื่อทดสอบความแตกต่างของค่าอำนาจจำแนกพบว่ามีความแตกต่างไม่แตกต่างกัน

3. ค่าความเชื่อมั่นของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน ตามรูปแบบของข้อสอบแบบเลือกตอบ ที่มีอัตราส่วนของจำนวนข้อสอบต่างกัน กลุ่มสาระการเรียนรู้วิทยาศาสตร์ชั้นประถมศึกษาปีที่ 6 ฉบับที่ 2 มีค่าสูงสุด (0.77) รองลงมาคือฉบับที่ 1 (0.74) ฉบับที่ 3 (0.73) และฉบับที่ 4 (0.72) ถือว่ามีค่าความเชื่อมั่นสูง เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่น พบว่าไม่แตกต่างกัน

4. ค่าความเที่ยงตรงของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน ตามรูปแบบของข้อสอบแบบเลือกตอบ ที่มีอัตราส่วนของจำนวนข้อสอบต่างกัน กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ฉบับที่ 2 มีค่าสูงสุด (0.99) ส่วนฉบับที่ 1 ฉบับที่ 3 และ ฉบับที่ 4 มีค่าความเที่ยงตรง (0.98)

5. เกณฑ์ปกติของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน ตามรูปแบบของข้อสอบแบบเลือกตอบที่มีอัตราส่วนของจำนวนข้อสอบต่างกัน กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ฉบับที่ 1 มีคะแนน T ปกติ ตั้งแต่ T31 ถึง T70 ฉบับที่ 2 มีคะแนน T ปกติ ตั้งแต่ T28 ถึง T70 ฉบับที่ 3 มีคะแนน T ปกติ ตั้งแต่ T29 ถึง T81 และฉบับที่ 4 มีคะแนน T ปกติ ตั้งแต่ T32 ถึง T72

คำสำคัญ : แบบทดสอบมาตรฐาน

A Construction of Achievement Tests Based on the Multiple-Choice Test Model with Different Ratios of Test Items in the Science Learning Strand for Prathomsueksa 6

ABSTRACT

Achievement tests which will have quality must consist of several factors, and very importantly, the qualities of validity, objectivity, discrimination, and difficulty. If the constructed tests have qualities based on the academic principle, the measurement will be more efficient and more useful to organization of learning and teaching. This study aimed to construct achievement tests based on the multiple choice test model with different ratios of test items in the science learning strand for Prathomsueksa (grade) 6, which were 4-choice tests with the ration of single-question type of constant choices, and different situations with totally 4 tests. The sample consisted of 800 Prathomsueksa 6 students from 25 schools under the Office Udon Thani Educational Service Area Zone 3 in the second semester of the academic year 2009, obtained using the multi-stage random sampling technique. They were tested, and the scores were used to construct norms in the form of T-norm scores. The T-norm extension relied on the construction of predictive equations. The instruments used were 4 four-choice tests, 40 items each, with choices : situations of 32 : 4 : 4, 24 : 8 : 8, 16 : 12 : 12, and 16 : 8 : 16 respectively. All these were adjusted from the 45-item single-question-type 4-choice basic test.

The results of the study were as follows :

1. For mean difficulties of the achievement tests based on the multiple-choice test model with different ratios of the test items in the science learning strand for Prathomsueksa 6, Test 1 had the highest value (0.45), followed by Test 2 (0.44), Test 3 (0.43), and Test 4 (0.40). Mean difficulties of all of the 4 tests ranged 0.40-0.45. They were regarded as having moderate difficulties. When differences of mean difficulties were tested, no differences were found.
2. For mean discriminating powers of the achievement tests based on the multiple-choice test model with different ratios of the test items in the science learning strand for Prathomsueksa 6, Test 2 had the highest value (0.39), followed by Test 1 (0.38), Test 3 (0.37), and Test 4 (0.36). Mean discriminating powers of all of the 4 tests ranged 0.36-0.39. They were regarded as being able to be discriminated. When

differences of discriminating powers were tested, it was found that they did not have different mean powers.

3. For reliabilities of the achievement tests based on the multiple-choice test model with different ratios of the test items in the science learning strand for Prathomsueksa 6, Test 2 had the highest value (0.77), followed by Test 1 (0.74), Test 3 (0.73), and Test 4 (0.72). They were regarded as having high reliabilities. When differences of reliabilities were tested, no differences were found.

4. For validities of the achievement tests based on the multiple-choice test model with different ratios of the test items in the science learning strand for Prathomsueksa 6, Test 2 had the highest value (0.99) which Test 1, Test 3, and Test 4 each had a validity of 0.98.

5. For norms of the achievement tests based on the multiple-choice test model with different ratios of the test items in the science learning strand for Prathomsueksa 6, Test 1 had T-scores ranging T31-T70 ; Test 2 had T-scores ranging T28-T70 ; Test 3 had T-scores ranging T29-T81 ; and Test 4 had T-scores ranging T32-T72.

Keyword : Standardized Tests

บทนำ

การประเมินผลที่ดีนั้นต้องมาจากการวัดผลที่ดี คือจะต้องเป็นการวัดผล ที่มีความถูกต้อง (Validity) และมีความเชื่อมั่น (Reliability) การวัดผลต้องมีการวัดด้วยวิธีการต่างๆ ที่หลากหลายตามสภาพผู้เรียน และผู้สอน จะต้องวัดให้ต่อเนื่องครอบคลุมและทั่วถึง เมื่อนำผลการวัดมารวมสรุปก็จะทำให้การประเมินผลนั้นถูกต้องใกล้เคียงตามสภาพจริง แบบทดสอบเป็นเครื่องมือวัดผล ที่มีความสำคัญทำให้ครูได้ทราบถึงพฤติกรรมการเรียนรู้และประสิทธิภาพในการจัดการเรียนการสอน แบบทดสอบแบบเลือกตอบ เป็นแบบทดสอบแบบปรนัยที่นิยมใช้กันมาก เพราะสามารถวัดความสามารถในการเรียนรู้ตั้งแต่ระดับความรู้ความจำจนถึงระดับประเมินผลได้ วิธีตรวจให้คะแนน ได้ผลคงที่ ประหยัดเวลา สามารถวิเคราะห์คุณภาพของข้อสอบตลอดจนตัวเลือกต่างๆได้ และยังสามารถจำแนกเด็กออกเป็นกลุ่ม ที่มีความสามารถแตกต่างกันได้อีกด้วย

ดังนั้นผู้วิจัยจึงมีความสนใจสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนตามรูปแบบที่มีอัตราส่วนของจำนวนข้อสอบแบบเลือกตอบต่างกัน เพื่อเป็นแนวทางในการสร้างและปรับปรุงแบบทดสอบชนิดเลือกตอบ อันจะเป็นประโยชน์สำหรับครู อาจารย์ ผู้ปกครองและผู้ที่เกี่ยวข้องในการนำแบบทดสอบที่มีคุณภาพที่สุดไปใช้ในการวัดและประเมินนักเรียน เพื่อแก้ไขปรับปรุงเพิ่มประสิทธิภาพการเรียนการสอน และพัฒนาคุณภาพทางการศึกษาต่อไป

ความมุ่งหมายของการวิจัย

1. เพื่อสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนตามรูปแบบที่มีอัตราส่วนของจำนวนข้อสอบแบบเลือกตอบต่างกัน กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6
2. เพื่อสร้างเกณฑ์ปกติ (Norms) ระดับประถมศึกษาชั้นประถมศึกษาปีที่ 6 สังกัดสำนักงานเขตพื้นที่การศึกษามุกดาหาร เขต 3
3. เพื่อเปรียบเทียบค่าความยาก ค่าอำนาจจำแนก และค่าความเชื่อมั่น ของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนตามรูปแบบที่มีอัตราส่วนของจำนวนข้อสอบแบบเลือกตอบต่างกัน กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6

ระเบียบวิธีวิจัย

1. ประชากรและกลุ่มตัวอย่าง
ประชากรที่ใช้ในการศึกษาค้นคว้านี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2552 สังกัดสำนักงานเขตพื้นที่การศึกษามุกดาหาร เขต 3 จำนวน 5,432 คน จากโรงเรียน 228 โรง กลุ่มตัวอย่างที่ใช้ในการวิจัยเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2552 สังกัดสำนักงานเขตพื้นที่การศึกษามุกดาหาร เขต 3 จำนวน 800 คน จาก โรงเรียน 25 โรง ซึ่งได้มาโดยการสุ่มแบบหลายขั้นตอน (Multi - Stage Random Sampling)
2. เครื่องมือที่ใช้ในการวิจัย
เครื่องมือที่ใช้ในการวิจัย เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนกลุ่มสาระการเรียนรู้วิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ตามรูปแบบที่มีอัตราส่วนจำนวนข้อสอบต่างกัน แบบเลือกตอบชนิด 4 ตัวเลือก จำนวน 40 ข้อ โดยมี อัตราส่วนของจำนวนข้อสอบประเภทคำถามเดียว คำถามตัวเลือกคงที่ และคำถามแบบสถานการณ์ที่ต่างกัน แยกเป็น 4 ฉบับ ดังนี้

ฉบับที่ 1 อัตราส่วนของจำนวนข้อสอบคิดเป็นร้อยละ 80 : 10 : 10 หมายถึง รูปแบบคำถามเดียว จำนวน 32 ข้อ ตัวเลือกคงที่ จำนวน 4 ข้อ และแบบสถานการณ์ จำนวน 4 ข้อ

ฉบับที่ 2 อัตราส่วนของจำนวนข้อสอบคิดเป็นร้อยละ 60 : 20 : 20 หมายถึง รูปแบบคำถามเดียว จำนวน 24 ข้อ ตัวเลือกคงที่ จำนวน 8 ข้อ และแบบสถานการณ์ จำนวน 8 ข้อ

ฉบับที่ 3 อัตราส่วนของจำนวนข้อสอบคิดเป็นร้อยละ 40 : 30 : 30 หมายถึง รูปแบบคำถามเดียว จำนวน 16 ข้อ ตัวเลือกคงที่ จำนวน 12 ข้อ และแบบสถานการณ์ จำนวน 12 ข้อ

ฉบับที่ 4 อัตราส่วนของจำนวนข้อสอบคิดเป็นร้อยละ 40 : 20 : 40 หมายถึง รูปแบบคำถามเดียว จำนวน 16 ข้อ ตัวเลือกคงที่ จำนวน 8 ข้อ และแบบสถานการณ์ จำนวน 16 ข้อ

3. วิธีดำเนินการเก็บรวบรวมข้อมูล

ติดต่อขอความร่วมมือกับโรงเรียนที่เป็นกลุ่มตัวอย่าง จำนวน 9 โรงเรียน จำนวนนักเรียนที่เป็นกลุ่มตัวอย่าง จำนวน 250 คน เพื่อหาคุณภาพของแบบทดสอบ และดำเนินการสอบอีก 2 ครั้ง โดยนำผลการสอบในแต่ละครั้ง มาตรวจให้คะแนน นำผลมาวิเคราะห์ทางสถิติเพื่อหาค่าพารามิเตอร์ของข้อสอบ ได้แก่ค่าความยากและค่าอำนาจจำแนก และปรับปรุงข้อที่มีคุณภาพไม่เข้าเกณฑ์

4. การวิเคราะห์ข้อมูล

4.1 หาค่าความยากรายข้อ ค่าอำนาจจำแนกรายข้อ และค่าความเชื่อมั่นทั้งฉบับของแบบทดสอบ ทั้ง 4 ฉบับ

4.2 หาค่าสถิติพื้นฐาน ค่าความยากเฉลี่ย ค่าอำนาจจำแนกเฉลี่ย ค่าความเชื่อมั่นทั้งฉบับ ค่าความเที่ยงตรง และค่าสหสัมพันธ์ของคะแนนสอบกับคะแนน T ปกติของแบบทดสอบทั้ง 4 ฉบับ จากการทดสอบครั้งที่ 2

4.3 ทดสอบความแตกต่างของค่าความยากเฉลี่ย ค่าอำนาจจำแนก และค่าความเชื่อมั่นของแบบทดสอบแต่ละฉบับ

ผลการวิจัย

1. ค่าความยากเฉลี่ย ค่าอำนาจจำแนกเฉลี่ย ค่าความเชื่อมั่นทั้งฉบับ และค่าสหสัมพันธ์ ของคะแนนสอบกับคะแนน T ปกติ ของแบบทดสอบ จำนวน 4 ฉบับ

แบบทดสอบ	N	$\bar{P}$	$\bar{r}$	r_{tt}	r_{xy}
ฉบับที่ 1	75	0.45	0.38	0.74	0.98
ฉบับที่ 2	75	0.44	0.39	0.77	0.99
ฉบับที่ 3	75	0.43	0.37	0.73	0.98
ฉบับที่ 4	75	0.40	0.36	0.72	0.98

** มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง พบว่า แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนตามรูปแบบข้อสอบแบบเลือกตอบที่มีอัตราส่วนของจำนวนข้อสอบต่างกัน กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 ปรากฏดังนี้

1. ค่าความยากเฉลี่ย ฉบับที่ 1 มีค่าต่ำสุด รองลงมาได้แก่ ฉบับที่ 2 และฉบับที่ 3 ตามลำดับ ส่วนฉบับที่ 4 มีค่าความยากเฉลี่ยสูงสุด
2. ค่าอำนาจจำแนกเฉลี่ย ฉบับที่ 2 มีค่าสูงสุด รองลงมาได้แก่ ฉบับที่ 1 ฉบับที่ 3 ตามลำดับ และฉบับที่ 4 มีค่าอำนาจจำแนกเฉลี่ยต่ำสุด
3. ค่าความเชื่อมั่น ฉบับที่ 2 มีค่าสูงสุด รองลงมาคือฉบับที่ 3 และฉบับที่ 4 ตามลำดับ ส่วนฉบับที่ 1 มีค่าความเชื่อมั่นต่ำสุด
4. ค่าสหสัมพันธ์ของคะแนนสอบกับคะแนน T ปกติ มีค่าตั้งแต่ 0.98-0.99 ถือว่า มีความเที่ยงตรงสูง เมื่อทดสอบความแตกต่าง พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01
2. การทดสอบความแตกต่างของค่าความยากเฉลี่ย ค่าอำนาจจำแนก และค่าความเชื่อมั่น ของแบบทดสอบแต่ละฉบับ

แหล่งของการแปรปรวน	ผลรวมของกำลังสอง (SS)	df	ค่าประมาณของความแปรปรวน(MS)	F	P
ระหว่างกลุ่ม	0.043	3	0.009	0.250	0.710
ภายในกลุ่ม	8.870	156	0.016		
ทั้งหมด	8.913	159			

จากตาราง พบว่าค่าความยากเฉลี่ยของแบบทดสอบทั้ง 4 ฉบับ ไม่แตกต่างกัน

ตารางการทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบทั้ง 4 ฉบับ

แบบทดสอบสอบ	N	N-3	r_{tt}	Zr	Zr(N-3)	$Zr^2(N-3)$	P
ฉบับที่ 1	75	72	.74	0.94	68.11	64.434	0.34
ฉบับที่ 2	75	72	.77	1.02	73.65	75.350	
ฉบับที่ 3	75	72	.73	0.94	67.82	63.890	
ฉบับที่ 4	75	72	.72	0.90	65.08	58.840	
รวม	300	288	2.96	3.80	274.68	262.52	

จากตาราง พบว่าค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบทั้ง 4 ฉบับ มีค่าอำนาจจำแนกไม่แตกต่างกัน

อภิปรายผล

1. ค่าความยากเฉลี่ยของทดสอบ ถ้าให้เวลาสอบเท่ากันทุกฉบับ จะทำให้ค่าความยากของแบบแตกต่างกัน องค์ประกอบที่มีอิทธิพลต่อความยากของแบบทดสอบ คือธรรมชาติของเนื้อหาวิชา ภาษาที่ใช้ รูปแบบคำถามจากแบบทดสอบ ที่ผู้วิจัยสร้างขึ้นพบว่าค่าเฉลี่ยของแบบทดสอบ ทั้ง 4 ฉบับ มีค่าใกล้เคียงกัน อยู่ในเกณฑ์ที่ใช้ได้ทุกฉบับ และเมื่อทดสอบความแตกต่างของค่าเฉลี่ยทั้ง 4 ฉบับ พบว่าค่าเฉลี่ยไม่แตกต่างกัน

2. ค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบ องค์ประกอบที่มีอิทธิพลต่อค่าอำนาจจำแนกของแบบทดสอบ คือ ค่าอำนาจจำแนกของแบบทดสอบจะมากขึ้น โดยส่วนที่มากขึ้นจะมีลักษณะเป็นเอกพันธ์กับข้อเดิมจะทำให้ค่าอำนาจจำแนกสูงขึ้น จากแบบทดสอบที่ใช้ในการวิจัยทั้ง 4 ฉบับ เป็นแบบทดสอบเท่ากัน ซึ่งจำนวนข้อสอบไม่ได้เพิ่มขึ้น จึงทำให้ค่าอำนาจจำแนกของแบบทดสอบใกล้เคียงกัน และมีค่าค่อนข้างสูง

3. ค่าความเชื่อมั่นของแบบทดสอบ องค์ประกอบที่มีอิทธิพลต่อค่าความเชื่อมั่นของแบบทดสอบคือ ค่าความเชื่อมั่นของแบบทดสอบจะมากขึ้น โดยส่วนที่มากขึ้นมีลักษณะเป็นเอกพันธ์กับข้อเดิม จะทำให้ค่าความเชื่อมั่นสูงขึ้น จากแบบทดสอบที่ผู้วิจัยสร้างขึ้น พบว่ามีค่าความเชื่อมั่นใกล้เคียงกัน เมื่อพิจารณาแต่ละฉบับพบว่า มีค่าความเชื่อมั่นค่อนข้างสูง โดยเฉพาะ ฉบับที่ 2 มีค่าความเชื่อมั่นสูงสุด และเมื่อทดสอบความแตกต่างของค่าความเชื่อมั่น ทั้ง 4 ฉบับ พบว่าค่าความเชื่อมั่นไม่แตกต่างกัน

4. ค่าสหสัมพันธ์ของการสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน ระหว่าง คะแนนดิบ กับคะแนน T ปกติ มีความสัมพันธ์กันทางบวก ถือว่ามีความเที่ยงตรงสูง เมื่อทดสอบความแตกต่าง ระหว่างค่าความเที่ยงตรง พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .01

5. เกณฑ์ปกติ ผู้วิจัยสร้างเกณฑ์ปกติ โดยนำคะแนนจากการหาคุณภาพขอแบบทดสอบ ทั้ง 4 ฉบับ มาแปลงเป็นคะแนน ทิปกติ พบว่าเกณฑ์ปกติของแบบทดสอบ ฉบับที่ 1 มีคะแนนดิบ ตั้งแต่ 8 ถึง 28 คะแนนปกติ ตั้งแต่ T31 ถึง T70 ฉบับที่ 2 มีคะแนนดิบ ตั้งแต่ 6 ถึง 28 คะแนนปกติ ตั้งแต่ T28 ถึง T70 ฉบับที่ 3 มีคะแนนดิบตั้งแต่ 5 ถึง 37 คะแนนปกติ ตั้งแต่ T29 ถึง T81 และฉบับที่ 4 มีคะแนนดิบ ตั้งแต่ 8 ถึง 32 คะแนน T ตั้งแต่ T32 ถึง T72 เนื่องจากการกระจายของคะแนนไม่ครอบคลุมทุกช่วง การสร้างเกณฑ์ปกติ จึงต้องขยายขอบเขตของคะแนนดิบในบางช่วงคะแนน เพื่อแสดงความสัมพันธ์จากคะแนนดิบของแบบทดสอบ ทั้ง 4 ฉบับ

ข้อเสนอแนะ

การนำแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนตามรูปแบบของข้อสอบแบบเลือกตอบที่มีอัตราส่วนของจำนวนข้อสอบต่างกัน กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 6 แบบเลือกตอบ 4 ตัวเลือก ไปใช้กับนักเรียนชั้นประถมศึกษาปีที่ 6 เพื่อเป็นแนวทางในการปรับปรุงการพัฒนาพฤติกรรมการวิเคราะห์ของนักเรียนให้ดียิ่งขึ้น ควรนำแบบทดสอบฉบับที่ 2 ไปใช้ เนื่องจากมีค่าอำนาจจำแนกและความเชื่อมั่นสูง และวัดได้เที่ยงตรงตามสภาพ