

บทความวิทยานิพนธ์
ผลของช่วงเวลาการสอบซ้ำและความยาวของแบบทดสอบที่มี
ผลต่อความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบ

The Effect of Testing-Retesting intervals and
Test Lengths on the Test Reliability and validity

วันวิสาข์ ศรีภูมิ¹

มนตรี อนันตรักษ์²

สุรชา อมรพันธุ์³

บทคัดย่อ

การวัดผลการศึกษามีความสำคัญต่อกระบวนการเรียนการสอนทุกระดับที่ต้องอาศัยเครื่องมือที่มีคุณภาพที่จะสามารถวัดพฤติกรรมของผู้เรียนได้อย่างแท้จริง แบบทดสอบจึงเป็นเครื่องมือที่สำคัญที่ใช้กันอย่างแพร่หลายและสามารถนำผลการวัดนั้นมาใช้ให้เป็นประโยชน์ได้ ด้วยเหตุนี้ในการวัดผลแต่ละครั้งจึงควรสร้างเครื่องมือให้มีคุณภาพ การวิจัยครั้งนี้มีความมุ่งหมายเพื่อเปรียบเทียบช่วงเวลาการสอบซ้ำที่มีผลต่อความเชื่อมั่นของแบบทดสอบและเพื่อเปรียบเทียบความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบที่มีความยาวของแบบทดสอบต่างกัน กลุ่มตัวอย่างที่ใช้ในการวิจัยในครั้งนี้คือ นักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2551 สังกัดสำนักงานเขตพื้นที่การศึกษามหาสารคาม เขต 2 จำนวน 360 คน จาก 26 โรงเรียน ซึ่งได้มาโดยวิธีการสุ่มแบบหลายขั้นตอน (Multi – Stage Random Sampling) แล้วนำมาแบ่งเป็น 3 กลุ่ม ๆ ละ 120 คน เครื่องมือที่ใช้ในการวิจัยนี้เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน

¹มหาบัณฑิตสาขาการวัดผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

²อาจารย์ ภาควิชาวิจัยและพัฒนาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

³รองศาสตราจารย์ ภาควิชาวิทยาศาสตร์สุขภาพและการกีฬา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ประจำภาคเรียนที่ 1 ปีการศึกษา 2551 รวม 7 บท เป็นชนิดเลือกตอบ 4 ตัวเลือก จำนวน 3 ฉบับ โดยฉบับที่ 1 มีจำนวน 30 ข้อ ฉบับที่ 2 มีจำนวน 50 ข้อ และฉบับที่ 3 มีจำนวน 70 ข้อ ซึ่งนักเรียนแต่ละกลุ่มทำแบบทดสอบทั้ง 3 ฉบับ ฉบับละ 2 ครั้ง โดยกลุ่มที่ 1 กลุ่มที่ 2 และกลุ่มที่ 3 จะเว้นช่วงทำแบบทดสอบห่างกัน 2 สัปดาห์ 4 สัปดาห์ และ 6 สัปดาห์ตามลำดับ สถิติที่ใช้ ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความคลาดเคลื่อนมาตรฐาน ความเชื่อมั่น ทดสอบความแตกต่างโดยใช้ χ^2 -test และ Z-test

ผลการวิจัยสรุปได้ดังนี้

1. ค่าความเชื่อมั่นของแบบทดสอบที่มีช่วงเวลากการสอบซ้ำที่ต่างกันและความยาวของแบบทดสอบต่างกัน พบว่า แบบทดสอบที่ให้ค่าความเชื่อมั่นดีที่สุดคือ แบบทดสอบฉบับที่ 3 จำนวน 70 ข้อ ที่มีช่วงเวลากการสอบซ้ำ 2 สัปดาห์ มีค่าความเชื่อมั่นเท่ากับ 0.922 รองมาคือแบบทดสอบฉบับที่ 2 จำนวน 50 ข้อ ที่มีช่วงเวลากการสอบซ้ำ 6 สัปดาห์ มีค่าความเชื่อมั่นเท่ากับ 0.920 และแบบทดสอบฉบับที่ 3 จำนวน 70 ข้อ ที่มีช่วงเวลากการสอบซ้ำ 6 สัปดาห์ มีค่าความเชื่อมั่นเท่ากับ 0.917 ตามลำดับ

2. เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่มีความยาวของแบบทดสอบต่างกัน พบว่า แบบทดสอบทั้ง 3 ฉบับ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยแบบทดสอบฉบับที่ 3 จำนวน 70 ข้อ มีค่าความเชื่อมั่นสูงสุดเท่ากับ 0.857

3. เปรียบเทียบค่าความเที่ยงตรงตามสภาพของแบบทดสอบ พบว่า ค่าความเที่ยงตรงตามสภาพของแบบทดสอบทั้ง 3 ฉบับ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยแบบทดสอบฉบับที่ 3 จำนวน 70 ข้อ มีค่าความเที่ยงตรงตามสภาพสูงสุดเท่ากับ 0.674

โดยสรุป ความยาวของแบบทดสอบที่มีจำนวน 70 ข้อโดยเว้นช่วงเวลากการสอบซ้ำ 2 สัปดาห์ เป็นองค์ประกอบหนึ่งที่ทำให้แบบทดสอบมีคุณภาพดีที่สุดในการวิจัยครั้งนี้ จึงเป็นข้อควรคำนึงในการดำเนินการวัดพฤติกรรมของผู้เรียนในครั้งต่อไป

คำสำคัญ : ช่วงเวลากการสอบซ้ำ, ความยาวของแบบทดสอบ, ความเชื่อมั่น, ความเที่ยงตรง

ABSTRACT

Educational measurement is important to the learning-teaching process at every level which must rely on quality instruments that can really measure learners' behaviors. Thus tests are important instruments widely used, and the results of measurement can be utilized. For this reason, in each measurement a quality instrument should be constructed. This study aimed to compare testing-retesting intervals appropriate for determining reliability coefficients of the tests, and to compare reliabilities and validities of the tests with different test lengths. The sample used in this study consisted of 360 Prathomsuksa (grade) 6 students attending schools under the Office of Mahasarakham Educational Service Area Zone 2 in the second semester of the academic year 2008, obtained using the multi-stage random sampling technique. The instruments used in the study were 3 multiple-choice tests of Prathomsuksa 6 mathematics : Tests 1, 2 and 3 had 30, 50 and 70 items respectively. For each test the data were collected twice. They were divided into 3 intervals : testing-retesting intervals of 2, 4 and 6 weeks respectively. The statistics used for analyzing the collected data were mean, standard deviation, standard error, reliability correlation (r_{xy}) ; and Chi-square test and Z-test were employed for testing differences.

The results of the study were as follows :

1. Reliability coefficients of the tests with different testing-retesting intervals and different test lengths were 0.636-0.922. The reliability coefficient of Test 3 with 70 items and with the testing-retesting interval of 2 weeks was the highest (0.922) ; and the reliability coefficient of Test 2 with 30 items and with the testing-retesting interval of 2 weeks was the lowest (0.636).
2. Reliability coefficients of the 3 tests with different test lengths : Tests 1, 2 and 3 with 30, 50 and 70 items respectively were different at the .05 level of significance, in which the test 3 with 70 items had the highest reliability coefficient.
3. Concurrent validities of Tests 1, 2 and 3 with 30, 50 and 70 items respectively were different at the .05 level of significance, in which the tests 3 with 70 items had the highest validity.

In conclusion, testing-retesting intervals and test lengths had influences on the test quality. Two-weeks testing-retesting interval with 70-items test could generate the best quality test which could be actually implemented in measuring learners' behavior.

Key Word : The effect of time intervals , Test – retest , Test lengths , Reliability ,Validity

ภูมิหลัง

การวัดผลการศึกษามีความสำคัญต่อกระบวนการเรียนการสอน ซึ่งเป็นเครื่องมือที่บอกให้ทราบถึงพัฒนาของผู้เรียนว่าได้บรรลุถึงจุดประสงค์เชิงพฤติกรรมที่วางไว้เพียงใด ซึ่งจะเป็นประโยชน์ต่อการจัดการเรียนการสอนให้มีประสิทธิภาพยิ่งขึ้น (อนันต์ ศรีโสภณ. 2524 : 1)

การวัดผลการศึกษาจะเกิดคุณค่า จึงจำเป็นอย่างยิ่งที่จะต้องอาศัยเครื่องมือในการวัดที่เหมาะสมและมีคุณภาพสูง ทั้งนี้เพราะถ้าได้เครื่องมือที่ไม่ดี ผลของการวัดก็ไม่ถูกต้องแน่นอน ทำให้ไม่สามารถนำผลการวัดมาใช้ประโยชน์ได้ (สวัสดี ประทุมราช . 2531 : 32) ด้วยเหตุนี้ในการวัดผลแต่ละครั้งจึงควรเลือกเครื่องมือที่มีคุณภาพมาใช้ ในบรรดาเครื่องมือวัดผลทั้งหลายนั้น “ แบบทดสอบ ” จัดว่าเป็นเครื่องมือที่ใช้กันอย่างแพร่หลายที่สุด โดยเฉพาะในการวัดผลด้านความรู้และการคิดของบุคคลเพราะจะทำให้ครูผู้สอนทราบถึงสถานภาพของนักเรียนและครูผู้สอนเอง ได้ง่ายกว่าวิธีวัดผลแบบอื่นๆ มาก (ชวาล แพรัตกุล. 2543 : 88) คุณสมบัติที่มีความสำคัญที่สุดของเครื่องมือในการวัดผลการศึกษาคือ ความเชื่อมั่นและความเที่ยงตรง

ด้วยสาระดังกล่าวข้างต้น ทำให้ผู้วิจัยสนใจที่จะทำการวิจัยโดยศึกษาผลของช่วงเวลาการสอบซ้ำที่มีผลต่อความเชื่อมั่นของแบบทดสอบและเปรียบเทียบความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบที่มีความยาวของแบบทดสอบต่างกัน ว่าช่วงเวลาการสอบซ้ำเท่าใดและแบบทดสอบที่มีจำนวนข้อเท่าใดที่จะทำให้แบบทดสอบมีค่าความเชื่อมั่นและความเที่ยงตรงสูง ซึ่งจะได้เป็นแนวทางสำหรับครูผู้สอนที่จะนำไปหาคุณภาพของเครื่องมือในการวัดผลให้มีประสิทธิภาพต่อไป

ความมุ่งหมายของการวิจัย

1. เพื่อศึกษาผลของช่วงเวลาการสอบซ้ำที่เหมาะสมสำหรับการหาค่าความเชื่อมั่นของแบบทดสอบ
2. เพื่อเปรียบเทียบความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบที่มีความยาวของแบบทดสอบต่างกัน

สมมติฐานของการวิจัย

1. ช่วงเวลาการสอบซ้ำในการทำแบบทดสอบที่ต่างกัน ส่งผลต่อค่าความเชื่อมั่นแตกต่างกัน
2. ความยาวของแบบทดสอบที่มีจำนวนข้อมาก มีค่าความเชื่อมั่นและความเที่ยงตรงสูงกว่าความยาวของแบบทดสอบที่มีจำนวนข้อน้อย

วิธีการวิจัย

ประชากรและกลุ่มตัวอย่าง

1. ประชากร

ประชากรที่ใช้ในการวิจัยครั้งนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2551 โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษามหาสารคาม เขต 2 จำนวน 6,507 คน

2. กลุ่มตัวอย่าง

กลุ่มตัวอย่าง ได้แก่ นักเรียนชั้นประถมศึกษาปีที่ 6 ของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษามหาสารคาม เขต 2 ภาคเรียนที่ 2 ปีการศึกษา 2551 จำนวน 360 คน ซึ่งได้มาโดยวิธีการสุ่มแบบหลายขั้นตอน (Multi – Stage Random Sampling) เครื่องมือที่ใช้ในการวิจัย

1. แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ชนิดเลือกตอบ 4 ตัวเลือก จำนวน 3 ฉบับ คือ
 - 1.1 ฉบับที่ 1 มีจำนวน 30 ข้อ
 - 1.2 ฉบับที่ 2 มีจำนวน 50 ข้อ
 - 1.3 ฉบับที่ 3 มีจำนวน 70 ข้อ

การดำเนินการทดลอง

การวิจัยในครั้งนี้เป็นการวิจัยเชิงทดลอง (Experimental Research) โดยผู้วิจัยได้ดำเนินการทดลองในภาคเรียนที่ 2 ปีการศึกษา 2551 ช่วงเดือน กุมภาพันธ์ 2552 – เดือน มีนาคม 2552 ซึ่งมีขั้นตอนการดำเนินการวิจัยดังนี้

1. ทำหนังสือติดต่อและขออนุญาตผู้บริหารโรงเรียนและโรงเรียนที่เป็นกลุ่มตัวอย่างเพื่อนำเครื่องมือไปทดลอง
2. นำแบบทดสอบทั้ง 3 ฉบับไปทดสอบกับกลุ่มตัวอย่างตามแผนการดำเนินการสอบจำนวน 3 กลุ่ม กลุ่มละ 120 คน โดยดำเนินการทดสอบ 2 ครั้ง โดยเว้นช่วงเวลาการสอบซ้ำ 2 สัปดาห์ 4 สัปดาห์ และ 6 สัปดาห์ โดยให้เวลาในการสอบดังนี้

ฉบับที่ 1 จำนวน 30 ข้อ กำหนดเวลาสอบ 30 นาที

ฉบับที่ 2 จำนวน 50 ข้อ กำหนดเวลาสอบ 50 นาที

ฉบับที่ 3 จำนวน 70 ข้อ กำหนดเวลาสอบ 70 นาที

3. นำผลที่ได้มาทำการวิเคราะห์หาคุณภาพของแบบทดสอบต่อไป

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. สถิติพื้นฐาน ได้แก่ $\bar{X}$, S.D. , SE_{meas}

2. สถิติที่ใช้วิเคราะห์ข้อมูล ได้แก่ r_{xy} , χ^2 -test และ Z-test

ผลการวิจัย

1. ค่าความเชื่อมั่นของแบบทดสอบที่มีช่วงเวลาการสอบซ้ำที่ต่างกันและความยาวของแบบทดสอบต่างกัน พบว่า แบบทดสอบที่ให้ค่าความเชื่อมั่นดีที่สุดคือ แบบทดสอบฉบับที่ 3 จำนวน 70 ข้อ ที่มีช่วงเวลาการสอบซ้ำ 2 สัปดาห์ มีค่าความเชื่อมั่นเท่ากับ 0.922 รองมาคือแบบทดสอบฉบับที่ 2 จำนวน 50 ข้อ ที่มีช่วงเวลาการสอบซ้ำ 6 สัปดาห์ มีค่าความเชื่อมั่นเท่ากับ 0.920 และแบบทดสอบฉบับที่ 3 จำนวน 70 ข้อ ที่มีช่วงเวลาการสอบซ้ำ 6 สัปดาห์ มีค่าความเชื่อมั่นเท่ากับ 0.917 ตามลำดับ

2. เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่มีความยาวของแบบทดสอบต่างกัน พบว่า แบบทดสอบทั้ง 3 ฉบับ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยแบบทดสอบฉบับที่ 3 จำนวน 70 ข้อ มีค่าความเชื่อมั่นสูงสุดเท่ากับ 0.857

3. เปรียบเทียบค่าความเที่ยงตรงตามสภาพของแบบทดสอบ พบว่า ค่าความเที่ยงตรงตามสภาพของแบบทดสอบทั้ง 3 ฉบับ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยแบบทดสอบฉบับที่ 3 จำนวน 70 ข้อ มีค่าความเที่ยงตรงตามสภาพสูงสุดเท่ากับ 0.674

โดยสรุป ความยาวของแบบทดสอบที่มีจำนวน 70 ข้อโดยเว้นช่วงเวลาการสอบซ้ำ 2 สัปดาห์ เป็นองค์ประกอบหนึ่งที่ทำให้แบบทดสอบมีคุณภาพดีที่สุดในการวิจัยครั้งนี้ จึงเป็นข้อควรคำนึงในการดำเนินการวัดพฤติกรรมของผู้เรียนในครั้งต่อไป

อภิปรายผล

จากผลการวิจัย สามารถอภิปรายผล ได้ดังนี้

1. ค่าความเชื่อมั่นของแบบทดสอบที่ให้ค่าความเชื่อมั่นดีที่สุดคือ แบบทดสอบฉบับที่ 3 จำนวน 70 ข้อ ที่มีช่วงเวลาการสอบซ้ำ 2 สัปดาห์ มีค่าความเชื่อมั่นเท่ากับ 0.922 รองมาคือแบบทดสอบฉบับที่ 2 จำนวน 50 ข้อ ที่มีช่วงเวลาการสอบซ้ำ 6 สัปดาห์ มีค่าความเชื่อมั่นเท่ากับ 0.920 และแบบทดสอบฉบับที่ 3 จำนวน 70 ข้อ ที่มีช่วงเวลาการสอบซ้ำ 6 สัปดาห์ มีค่าความเชื่อมั่นเท่ากับ 0.917 ทั้งนี้เนื่องมาจาก

1.1 แบบทดสอบฉบับที่ 3 เป็นแบบทดสอบที่มีจำนวน 70 ข้อ ที่ผู้วิจัยได้สร้างขึ้นนั้น ผ่านการประเมินจากผู้เชี่ยวชาญพิจารณาความสอดคล้องระหว่างข้อสอบกับผล การเรียนรู้ที่คาดหวัง จึงทำให้แบบทดสอบฉบับที่ 3 ที่มีจำนวนข้อสอบมากกว่าฉบับอื่นๆ นั้น สามารถวัดพฤติกรรมได้ครอบคลุมเนื้อหาได้มากกว่า จึงทำให้มีค่าความเชื่อมั่นสูงกว่าฉบับอื่น

1.2 ช่วงเวลาการสอบซ้ำ 2 สัปดาห์ เป็นช่วงเวลาที่เหมาะสมที่สุดสำหรับสัมประสิทธิ์ค่าความเชื่อมั่นของแบบทดสอบ เพราะการเว้นช่วงเวลาการสอบซ้ำน้อยไปอาจส่งผลให้เด็กยังจดจำข้อสอบได้อยู่ แต่ถ้านานเกินไปเด็กอาจลืมหรือเกิดการเรียนรู้อะหว่างการเว้นช่วงการสอบซ้ำนั้น ดังนั้นช่วงเวลาการสอบซ้ำ 2 สัปดาห์จึงเป็นช่วงเวลาการสอบซ้ำที่ดีที่สุดที่ทำให้แบบทดสอบมีคุณภาพ

2. เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบที่มีความยาวของแบบทดสอบต่างกัน พบว่า แบบทดสอบทั้ง 3 ฉบับ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยแบบทดสอบฉบับที่ 3 จำนวน 70 ข้อ มีค่าความเชื่อมั่นสูงสุดเท่ากับ 0.857 ทั้งนี้เนื่องจาก

2.1 แบบทดสอบที่ผู้วิจัยสร้างขึ้นได้ผ่านการประเมินจากผู้เชี่ยวชาญด้านคณิตศาสตร์และการวัดผลการศึกษา

2.2 แบบทดสอบที่เพิ่มจำนวนข้อเพิ่มมากขึ้นในฉบับที่ 3 นั้นได้คำนึงถึงเกณฑ์ที่ใช้ในการคัดเลือกข้อสอบแต่ละฉบับ คุณภาพของแบบทดสอบ และเวลาที่ใช้ในการทดสอบ จึงทำให้แบบทดสอบฉบับที่ 3 ที่มีจำนวน 70 ข้อ มีค่าความเชื่อมั่นสูงกว่าฉบับอื่นๆ

3. เปรียบเทียบค่าความเที่ยงตรงตามสภาพของแบบทดสอบ พบว่า ค่าความเที่ยงตรงตามสภาพของแบบทดสอบทั้ง 3 ฉบับ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยแบบทดสอบฉบับที่ 3 จำนวน 70 ข้อ มีค่าความเที่ยงตรงตามสภาพสูงสุดเท่ากับ 0.674 ทั้งนี้เนื่องจาก

3.1 การหาค่าความเที่ยงตรงตามสภาพได้ใช้คะแนนจากแบบทดสอบที่เป็นแบบทดสอบมาตรฐานจากสำนักทดสอบทางการศึกษา (O – Net) ปีการศึกษา 2551 เป็นเกณฑ์มาหาค่าสหสัมพันธ์กับผลการทดสอบครั้งที่ 1 ของนักเรียนที่เป็นกลุ่มตัวอย่างที่เป็นเกณฑ์ปัจจุบัน จึงทำให้แบบทดสอบฉบับที่ 3 ที่ผู้วิจัยสร้างขึ้นมีค่าความเที่ยงตรงตามสภาพทำให้สามารถประมาณสภาพอันแท้จริงของนักเรียนในปัจจุบันได้เป็นอย่างดี

เอกสารอ้างอิง

ชวาล แพรัตกุล . เทคนิคการวัดผล . ไทยวัฒนาพานิช , 2543 .

สวัสดี ประทุมราช . แนวคิดเชิงทฤษฎีการวิจัยและประเมินผล . กรุงเทพฯ : ม.ป.พ. , 2531 .อนันต์

ศรีโสภณ . การวัดและการประเมินผลการศึกษา . ไทยวัฒนาพานิช 2524 , 251 หน้า