

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) สร้างแบบสอบมาตรฐานทางคณิตศาสตร์ สำหรับนักเรียนระดับประถมศึกษา 2) ตรวจสอบคุณภาพของแบบสอบมาตรฐานทางคณิตศาสตร์ในด้านความเที่ยง (Reliability) และความตรง (Validity) โดยมีวิธีดำเนินการรวม 11 ขั้นตอน คือ 1) วิเคราะห์หลักสูตรคณิตศาสตร์ของกระทรวงศึกษาธิการรวมทั้งสมรรถภาพที่นักเรียนแต่ละระดับควรรู้ในระดับมาตรฐานสากล 2) สร้างผังแบบสอบ (Table of Test Specification) 3) กำหนดลักษณะเฉพาะของข้อสอบ (Item Specification) 4) ร่างแบบสอบตามลักษณะเฉพาะของข้อสอบ 5) ให้ผู้เชี่ยวชาญตรวจสอบความตรงตามเนื้อหา และตรวจสอบความตรงของข้อสอบรายข้อกับจุดมุ่งหมาย 6) ปรับปรุงและนำข้อสอบไปทดลองใช้กับกลุ่มตัวอย่าง 7) วิเคราะห์ข้อสอบ เพื่อคำนวณหาค่าความยากและอำนาจการจำแนกของข้อสอบ 8) คัดเลือกแบบสอบที่มีความยากพอเหมาะและอำนาจจำแนกสูงมาจัดชุดเป็นแบบสอบคู่ขนาน 9) นำแบบสอบคู่ขนานไปทดลองใช้กับนักเรียนกลุ่มตัวอย่าง 10) วิเคราะห์แบบสอบหาความเที่ยง (Reliability) และ 11) สร้างเกณฑ์แปลผล

ผลการวิจัยโดยสรุปพบว่าแบบสอบมาตรฐานทางคณิตศาสตร์ ทั้ง 3 ชุดมีคุณภาพดังนี้

1. แบบสอบแต่ละฉบับ มีความตรงตามเนื้อหา (content validity) มีความพืดเทียบกัน (equivalent forms) ตามหลักสูตรคณิตศาสตร์ระดับประถมศึกษา พ.ศ.2521 (ฉบับปรับปรุง 2533)

2. แบบสอบชุดที่ 1 ชุดที่ 2 และชุดที่ 3 มีระดับความยากเฉลี่ยเท่ากับ 0.47, 0.42 และ 0.45 ตามลำดับ ซึ่งไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 มีค่าสัมประสิทธิ์ความเที่ยงเท่ากับ .82, .80, และ .78 ตามลำดับ และมีค่าความคลาดเคลื่อนมาตรฐานของการวัดเท่ากับ 2.81, 2.77, 2.88 ตามลำดับ

This research purposes are to 1) construct the Mathematics Standardized Tests for elementary students 2) examine the Mathematics Standardized Tests in two dimensions, reliability and validity by using 11 procedures as follows

a) analyze the elementary mathematics curriculum based on the requirement of the Ministry of Education and the international efficiency requirement in mathematics for elementary students

b) construct the table of test specifications

c) specify the items specifications.

d) write the first draft of the test items

e) examine the content validity and the item objective congruence.

f) revise the items and try out with samples

g) determine the difficulty and discriminant power of the item

h) select the items which are determined as proper difficulty and discriminant power to make equivalent tests

i) pilot those tests with sample of elementary students

j) analyze the reliability of the tests

k) create the criteria for interpretation

The research results are as follows

1. each of the tests has equivalent content validity as in the elementary mathematics curriculum.
2. mean difficulty of the tests series 1, 2 and 3 are 0.47 , 0.42 and 0.45, respectively, the reliability coefficient are .82, .80 and .78, respectively. Standard Errors are 2.81, 2.77 and 2.88 respectively. Those 3 series have no significant different in item difficulty  $\alpha$  level of 0.05.