

ชื่อเรื่องวิทยานิพนธ์ การศึกษาประสิทธิผลของแบบทดสอบแบบเทเลอร์บางรูปแบบในการตัดสินความรอบรู้ในการเรียนวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6

ชื่อผู้เขียน นายสิทธิชัย หาญสมบัติ

ศึกษาศาสตรมหาบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา

คณะกรรมการตรวจสอบวิทยานิพนธ์

รศ. นิโบล	นันทกัมภีร์	ประธานกรรมการ
รศ. ดร. สุธรรม	จันทร์หอม	กรรมการ
ผศ. สวรรค์	อ่อนนาค	กรรมการ

บทคัดย่อ

การวิจัยครั้งนี้เป็นการศึกษาแบบทดสอบแบบเทเลอร์ 3 รูปแบบ ที่ใช้ในการตัดสินความรอบรู้ คือ แบบปิรามิดที่ใช้ขนาดชั้นคงที่จำนวน 6 ชั้นและ 10 ชั้น แบบปิรามิดที่มีข้อสอบหลายข้อในแต่ละชั้นจำนวน 3 ชั้นและ 5 ชั้น และแบบเปลี่ยนระดับจำนวน 6 ชั้นและ 10 ชั้น ดังนั้นแต่ละรูปแบบจึงมีจำนวนข้อที่นักเรียนต้องตอบ 2 ขนาด คือ 6 ข้อกับ 10 ข้อ โดยมีจุดมุ่งหมายเพื่อศึกษาความเที่ยงตรงในการตัดสินความรอบรู้ และ เปรียบเทียบผลการตัดสินของแบบทดสอบดังกล่าวที่มีจำนวนข้อที่ต้องตอบเท่ากับ การตัดสินตามปกติของครูและการตัดสินด้วยผลการสอบจากแบบทดสอบความรอบรู้แบบเทเลอร์ตามยุทธวิธีของเบย์ (AMT : Adaptive Mastery Testing) และเพื่อเปรียบเทียบผลการตัดสินความรอบรู้ระหว่างแบบทดสอบรูปแบบเดียวกันแต่มีจำนวนข้อที่นักเรียนต้องตอบต่างกัน

กลุ่มข้อสอบที่ใช้ในการวิจัย เป็นข้อสอบวิชาคณิตศาสตร์ที่วัดในจุดประสงค์การเรียนรู้ที่ 4 "เขียนเลขในรูปยกกำลัง ทา ห.ร.ม. และ ค.ร.น. ได้" ของชั้นประถมศึกษาปีที่ 6 จำนวน 119 ข้อ ใช้สำหรับการทดสอบความรอบรู้แบบเทเลอร์และจัดเป็นแบบทดสอบแบบเทเลอร์ 3 รูปแบบที่ต้องการศึกษา ข้อสอบเหล่านี้ได้รับการวิเคราะห์แบบโลจิสติกที่มีพารามิเตอร์ 3 ตัวโดยมีค่าอำนาจจำแนกตั้งแต่ 0.60 ขึ้นไปและมีค่าสัมประสิทธิ์การเดาไม่เกิน 0.35

กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนอนุบาลอุดรดิตถ์ ที่ได้จากการสุ่มแบบเป็นกลุ่ม (Cluster Sampling) จำนวน 33 คน ทดสอบความรอบรู้แบบเทเลอร์ตามยุทธวิธีของเบย์โดยใช้คอมพิวเตอร์ช่วยในการดำเนินการสอบ และทดสอบด้วยแบบทดสอบแบบเทเลอร์ 3 รูปแบบที่ต้องการศึกษาเป็นรายบุคคล ประมาณค่าความสามารถโดยใช้สูตรของเบกท์และไวส์ แล้วตัดสินความรอบรู้โดยเปรียบเทียบกับเกณฑ์ค่าความสามารถเฉลี่ย = 0.6329 ที่ได้จากความเห็นของครูในการพิจารณาข้อสอบ 30 ข้อ ที่สุ่มจากกลุ่มข้อสอบ ถ้าค่าความสามารถของนักเรียนเท่ากับหรือมากกว่าเกณฑ์แสดงว่าเป็นผู้รอบรู้ ถ้าต่ำกว่าแสดงว่ายังไม่เป็นผู้รอบรู้

วิเคราะห์ข้อมูลโดยนำผลการตัดสินความรอบรู้จากแบบทดสอบแบบเทเลอร์ที่ศึกษาทุกรูปแบบและขนาดความยาวของจำนวนข้อที่นักเรียนต้องตอบมาตรวจสอบความเที่ยงตรงในการตัดสินโดยหาความสัมพันธ์แบบพาย (ϕ) กับ ผลการตัดสินจากการทดสอบความรอบรู้แบบเทเลอร์ (AMT) เปรียบเทียบผลการตัดสินระหว่างแบบทดสอบแบบเทเลอร์ 3 รูปแบบ ที่มีจำนวนข้อที่นักเรียนต้องตอบเท่ากัน กับ ผลการตัดสินตามปรกติของครู และ ผลการตัดสินจากการทดสอบความรอบรู้แบบเทเลอร์ โดยใช้สถิติทดสอบครอกแควน (Cochran Test) และทดสอบภายหลัง (Post - hoc Comparison) ด้วยวิธีของ ดัน - บอนเฟอโรนี (Dunn - Bonferroni) และ เปรียบเทียบผลการตัดสินความรอบรู้จากแบบทดสอบแบบเทเลอร์รูปแบบเดียวกันแต่มีจำนวนข้อที่นักเรียนต้องตอบต่างกันโดยใช้การทดสอบทีสำหรับสัดส่วนแบบกลุ่มสัมพันธ์

ผลการวิจัยพบว่าแบบทดสอบแบบเทเลอร์ทั้ง 3 รูปแบบคือ แบบปิรามิดที่ใช้ขนาดชั้น
คงที่ แบบปิรามิดที่มีสองชั้นในแต่ละชั้นและแบบเปลี่ยนระดับ ไม่ว่าจะมียังจำนวนข้อที่นักเรียนต้อง
ตอบเป็น 6 ข้อหรือ 10 ข้อ ต่างก็ให้ผลการตัดสินที่มีความเที่ยงตรงอย่างมีนัยสำคัญทางสถิติ
ที่ระดับ .01 ผลการตัดสินจากแบบทดสอบแบบเทเลอร์ทั้ง 3 รูปแบบในแต่ละรูปแบบมีจำนวน
ข้อที่นักเรียนต้องตอบ 6 ข้อ ให้ผลการตัดสินแตกต่างจากการตัดสินตามปกติของครู แต่ไม่
แตกต่างจากการทดสอบความรอบรู้แบบเทเลอร์ตามยุทธวิธีของเบย์อย่างมีนัยสำคัญทางสถิติที่ระดับ
.05 ส่วนผลการตัดสินจากแบบทดสอบแบบเทเลอร์ทั้ง 3 รูปแบบในแต่ละรูปแบบมีจำนวนข้อที่
นักเรียนต้องตอบ 10 ข้อ มีความแตกต่างจากผลการตัดสินตามปกติของครู เฉพาะในรูปแบบ
เปลี่ยนระดับ ส่วนระหว่างแบบทดสอบแบบเทเลอร์ที่มีรูปแบบเดียวกันแต่มีจำนวนข้อที่นักเรียนต้อง
ตอบต่างกันให้ผลการตัดสินที่ไม่แตกต่างกัน

Thesis Title	A Study of Effectiveness of Some Tailored Testing Models in Judging Prathomsuksa 6 Students' Mastery of Mathematics		
Author	Mr. Sittichai Harnsombat		
M.Ed.	Educational Measurement and Evaluation		
Examining Committee	Assoc.Prof.Nilobol	Nimkingratana	Chairman
	Assoc.Prof.Dr.Suthan	Janhom	Member
	Assist.Prof.Sawan	Onnark	Member

Abstract

The purpose of this research is to study three models of tailored testing for judging students' mastery of mathematics. The three models are : The Constant Step Size Pyramid Model; 6-stage and 10-stage, The Multiple Item Pyramid Model; 3-stage and 5-stage, and The Flexilevel Test Model; 6-stage and 10 stage. So, each model contains two categories of test items, consisting of six test items and ten test items. The validity of judging the mastery of mathematics is focussed in this study. A comparison of the test judgement among equal test items, ordinary judging by teachers and Baysian Strategy Adaptive Mastery Testing (AMT) within the tailored testing models, is used. As well, the comparison of students' mastery of mathematics within the same test between two categories with more test items and less test items is considered.

A pool of 119 test items was used in the study are from the mathematic examination paper for the prathomsuksa 6th , measuring the 4th learning objective namely transforming the integer to the

indicies, and finding the Greatest Common Division (GCD) and the Least Common Multiple (LCM). Tailored testing of the three models will be employed. The 119 test items were analyzed by a 3-parameter logistic model with discrimination power of over 0.60 and under 0.35 of guessing coefficient.

The sample for the study is 33 students, a cluster sampling from Anuban Uttaradit School. Two tests, AMT and the tailored testing for three models for individual, will be used. The estimation of students' ability is calculated by using Betz and Weiss's formula. The judging of students' mastery of mathematics can be done by the comparison of the average criterion of 0.6329, obtained by the opinion of 14 mathematician teachers who do a study of a 30 test item sampling from the 119 test items pool. Any student whose final ability is equal to or above the average criterion is considered to have the mastery of mathematics. On the other hand, any student whose final ability is below the average criterion has not mastered mathematics.

Data analysis will be done by validating the test results of the three models of tailored testing to the results of AMT. Phi coefficient (ϕ) is used as an index of the validity for judging of the three models of tailored test.

The comparison of the test results on students' mastery of mathematics among tailored testing of three models, with the equal test items, AMT, and the ordinary judging by teachers will be constructed

ผล

using Crochran-Test with post-hoc comparison by Dunn-Bonferroni methodology.

The proportion of the test results on mastery and nonmastery of mathematics by the same tailored testing models, with 6 and 10 test items, will be compared using Z-test for proportion.

The study revealed that test results from all three models of tailored testing, namely, the Constant Step Size Pyramid Model, the Multiple Item Pyramid Model and the Flexilevel Test Model yielded the statistical significant of validity at .01 level of confidence.

The study also showed that the test results of the mastery of mathematics by tailored testing three models, with 6 test items in each test, were statistically significantly different from the ordinary judging by teachers, but were not statistically significantly different from the test results obtained from AMT at .05 level of confidence.

The study also indicated that the test results from all three models of tailored testing, with 10 test items in each test, were statistically different only within the Multiple Item Pyramid Model. All three models of tailored testing, with different test items, yielded no statistically significant differences.