การศึกษาความเที่ยง ความตรงของแบบทดสอบ และความสัมพันธ์ของคะแนนสอบระหว่างการให้คะแนนตามทฤษฎีการทดสอบดั้งเดิม กับการให้คะแนนตามทฤษฎีการตอบสนองข้อสอบ

A Study of Reliability, Validity of Test and Correlation of Test Scores between the Scoring of Classical Test Theory and the Scoring of Item Response Theory

นภดล ยิ่งยงสกุล

Napadon Yingyongsakul

เลขหมู่ LB3060.75 N43 2540

เลขทะเบียน.....

-1/ก.ย. 2540

Order Key.....

BIB Key.....

วิทยานิพนธ์ศึกษาศาสตรมหาบัณฑิต สาขาวิชาการวัดผลและวิจัยการศึกษา

มหาวิทยาลัยสงขลานครินทร์

Master of Education Thesis in Educational Measurement and Research

Prince of Songkla University

2540

ชื่อวิทยานิพนธ์ การศึกษาความเที่ยง ความตรงของแบบทดสอบ และความสัมพันธ์ของคะแนนสอบระหว่างการให้คะแนนตามทฤษฎีการทดสอบดั้งเดิม กับการให้คะแนนตามทฤษฎีการตอบสนองข้อสอบ

ผู้เขียน นายนภดล ยิ่งยงสกุล

สาขาวิชา การวัดผลและวิจัยการศึกษา

ปีการศึกษา 2539

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาค่าความเที่ยง ความตรงของแบบทดสอบ และความสัมพันธ์ของคะแนนสอบระหว่างการให้น้ำหนักคะแนนรายข้อเท่ากัน ตามทฤษฎีการทดสอบดั้งเดิม การให้น้ำหนักคะแนนรายข้อต่างกันตามค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ และการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าความสามารถของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ

กลุ่มตัวอย่างที่ใช้ในการวิจัยนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2539 ในเขตจังหวัดยะลา จำนวน 231 คน ซึ่งเลือกมาโดยวิธีการสุ่มอย่างง่าย เครื่องมือที่ใช้ในการรวบรวมข้อมูลเป็นแบบทดสอบวิชาคณิตศาสตร์ ชนิด 4 ตัวเลือก จำนวน 28 ข้อ ที่ผู้วิจัยสร้างขึ้น การวิเคราะห์ข้อมูลโดยคำนวณหา 1) ค่าความเที่ยงของแบบทดสอบ โดยสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's Alpha Coefficients) และเปรียบเทียบความแตกต่างโดยแปลงเป็นคะแนนมาตรฐานของฟิชเชอร์ (Zr) แล้วเปรียบเทียบความแตกต่างด้วยสถิติทดสอบซี (Z - test) 2) ค่าความตรงของแบบทดสอบ โดยสูตรอันดับที่ของสเปียร์แมน (Spearman's Rank Correlation Coefficient) และเปรียบเทียบความแตกต่างโดยแปลงเป็นคะแนนมาตรฐานของฟิชเชอร์ (Zr) แล้วเปรียบเทียบความแตกต่างด้วยสถิติทดสอบซี (Z - test) 3) ค่าความสัมพันธ์ของคะแนนสอบ โดยสูตรเพียร์สัน โพรดัก โมเมนต์ (Pearson Product Moment Correlation) และทดสอบนัยสำคัญด้วยสถิติทดสอบที (t-test)

ผลการวิจัยสรุปได้ดังนี้

1. ค่าความเที่ยงของแบบทดสอบจากวิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามความสามารถของผู้สอบ สูงกว่าวิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าพารามิเตอร์ของข้อสอบ และวิธีการให้น้ำหนักคะแนนรายข้อเท่ากัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนวิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าพารามิเตอร์ของข้อสอบ มีค่าสูงกว่าวิธีการให้น้ำหนักคะแนนรายข้อเท่ากัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2. ค่าความตรงของแบบทดสอบจากวิธีการให้น้ำหนักคะแนนรายข้อเท่ากัน วิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามค่าพารามิเตอร์ของข้อสอบ และวิธีการให้น้ำหนักคะแนนรายข้อต่างกันตามความสามารถของผู้สอบ มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

3. ค่าความสัมพันธ์ของคะแนนสอบระหว่างวิธีการให้คะแนนทั้ง 3 วิธี มีค่าตั้งแต่ .975 ถึง .993 และจากการทดสอบนัยสำคัญทางสถิติ พบว่ามีนัยสำคัญทางสถิติที่ระดับ .01 ทุกคู่

Thesis Title A Study of Reliability, Validity of Test and Correlation of Test Scores between the Scoring of Classical Test Theory and the Scoring of Item Response Theory

Author Mr. Napadon Yingyongsakul

Major Program Educational Measurement and Research

Academic Year 1996

Abstract

The purposes of this research were to determine the reliability, the validity of Test and the correlation of test scores between the raw scoring method of Classical Test Theory, the optimal scoring weight by parameter of item method of Item Response Theory and the optimal scoring weight by ability of examinee parameter method of Item Response Theory.

The subject of this research were 231 students obtained by simple random sampling from Prathomsuksa VI students in the academic years 1996 in Changwat Yala. The 28-item mathematics test of 4 multiple choices was used in data collection. Data analyses were carried out 1) for test reliability using Cronbach's Alpha coefficients, Fisher's standard scores and Z-test, 2) for test validity using Spearman's Rank correlation coefficients, Fisher's standard scores and Z-test, 3) for the correlation of test scores using Pearson Product - Moment correlation and t - test.

The finding were as follows :

1. The test reliability determined by ability of examinee (θ) parameter method was significantly higher than both of the optimal scoring weight by parameter of item method and the raw scoring method at .01 level, while that of the optimal scoring weight by parameter of item method was significantly higer than that of the raw scoring method at .01 level.

2. There was no significant difference in the test validity determined by the raw scoring method, the optimal scoring weight by parameter of item method and the optimal scoring weight by ability of examinee (θ) parameter method.

3. The correlation of tests scores of the three scoring methods ranged from .975 to .993 and was found to be significant at .01 level.