

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะ โดยใช้คอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่ต่างกัน กลุ่มตัวอย่าง คือ นักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2549 ในโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษานครราชสีมา เขต 1 จำนวน 924 คน เครื่องมือที่ใช้ในการวิจัย คือ แบบสอบวัดความสามารถทางคณิตศาสตร์ คลังข้อสอบวิชาคณิตศาสตร์ และโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ วิเคราะห์ข้อมูลโดยใช้ค่าสถิติเบื้องต้น ค่าสหสัมพันธ์ และค่าไคสแควร์

ผลการวิจัยพบว่า

1. การเปรียบเทียบอัตราการใช้ข้อสอบซ้ำต่างกัน (10%, 20%, 30%) ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่า ความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ
2. การเปรียบเทียบเกณฑ์ยุติการทดสอบต่างกัน ($SEE \leq 0.30$, $SEE \leq 0.45$) ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่า ความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ
3. การเปรียบเทียบเกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบแตกต่างกัน พบว่า ความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยที่เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ให้ค่าความตรงตามสภาพของการทดสอบสูงที่สุด

The purpose of this research was to compare concurrent validity in the estimation of examinees' ability from computerized adaptive testing using different early stage item selection rules, different item exposure rates and different termination criterion. The research sample consisted of 924 Pathom Saksa six students in the second semester of 2006 academic year from the schools in Nakonratchasima Education Service Area Office 1. The research instruments were Mathematics ability test, Mathematics item pool, and Computer program for computerized adaptive testing. Data were analyzed by using descriptive statistics, correlation coefficient and chi - square.

The research findings were as follows:

1. The comparison of different item exposure rates found that the concurrent validities in the estimation of examinees' ability were not significant.
2. The comparison of different termination criterion found that the concurrent validities in the estimation of examinees' ability were not significant.
3. The comparison of different early stage item selection rules, different item exposure rates and different termination criterion found that the concurrent validities in the estimation of examinees' ability were statistically significant at the .05 level. Kullback-Leibler information, item exposure rate 10 percent and standard error in the estimation of examinees' ability is less than or equal to 0.30 provided the highest concurrent validity.