การเปรียบเทียบความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ที่ใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่ต่างกัน

นางสาวสิริลักษณ์ เกษรปทุมานันท์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรมหาบัณฑิต
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2549
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A COMPARISON OF CONCURRENT VALIDITY IN THE ESTIMATE OF EXAMINEES'ABILITY FROM COMPUTERIZED ADAPTIVE TESTING USING DIFFERENT EARLY STAGE ITEM SELECTION RULES, DIFFERENT ITEM EXPOSURE RATES AND DIFFERENT TERMINATION CRITERION

Miss Sirilux Kesornpathumanunt

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Education Program in Educational Measurement and Evaluation
Department of Educational Research and Psychology
Faculty of Education
Chulalongkorn University
Academic Year 2006

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การเปรียบเทียบความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ที่ใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่ต่างกัน
โดย	นางสาวสิริลักษณ์ เกษรปทุมานันท์
สาขาวิชา	การวัดและประเมินผลการศึกษา
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.ณัฏฐภรณ์ หลาวทอง

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

..... คณบดีคณะครุศาสตร์
(รองศาสตราจารย์ ดร.พฤทธิ์ ศิริบรรณพิทักษ์)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี)

..... อาจารย์ที่ปรึกษา
(ผู้ช่วยศาสตราจารย์ ดร.ณัฏฐภรณ์ หลาวทอง)

..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.อวยพร เรืองตระกูล)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สิริลักษณ์ เกษรปทุมานันท์: การเปรียบเทียบความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ที่ใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่ต่างกัน(A COMPARISON OF CONCURRENT VALIDITY IN THE ESTIMATE OF EXAMINEES'ABILITY FROM COMPUTERIZED ADAPTIVE TESTING USING DIFFERENT EARLY STAGE ITEM SELECTION RULES, DIFFERENT ITEM EXPOSURE RATES AND DIFFERENT TERMINATION CRITERION) อ.ที่ปรึกษา: ผู้ช่วยศาสตราจารย์ ดร. ณัฏฐภรณ์ หลาวทอง, 210 หน้า.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่ต่างกัน กลุ่มตัวอย่าง คือ นักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2549 ในโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษานครราชสีมา เขต 1 จำนวน 924 คน เครื่องที่ใช้ในการวิจัย 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

ภาควิชา วิจัยและจิตวิทยาการศึกษา ลายมือชื่อนิสิต สิริลักษณ์ เกษรปทุมานันท์

สาขาวิชา การวัดและประเมินผลการศึกษา ลายมือชื่ออาจารย์ที่ปรึกษา

ปีการศึกษา 2549

478 37482 27: MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION
KEY WORDS: CONCURRENTE VALIDITY / COMPUTERIZED ADAPTIVE TESTING / FISHER INFORMATION / KULLBACK-LEIBLER INFORMATION

SIRILUX KESORNPATHUMANUNT : A COMPARISON OF CONCURRENT VALIDITY IN THE ESTIMATE OF EXAMINEES' ABILITY FROM COMPUTERIZED ADAPTIVE TESTING USING DIFFERENT EARLY STAGE ITEM SELECTION RULES, DIFFERENT ITEM EXPOSURE RATES AND DIFFERENT TERMINATION CRITERION. THESIS ADVISOR: ASST. PROF. NUTTAPORN LAWTHONG, Ph.D., 210 pp.

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 comparision of different item exposure rates found that the concurrent validities in the estimation of examinees' ability were not significant.

2. The comparision of different termination criterion found that the concurrent validities in the estimation of examinees' ability were not significant.

3. The comparision 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.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Department : Educational Research and Psychology Student's Signature : Sirilux Kesornpathumanunt
Field of Study : Educational Measurement and Evaluation Advisor's Signature : Nuttaporn Lawthong
Academic Year : 2006

กิตติกรรมประกาศ

วิทยานิพนธ์เล่มนี้สามารถสำเร็จได้ด้วยดี เนื่องจากได้รับความกรุณาช่วยเหลือเป็นอย่างดีจากบุคคลหลายท่าน ตั้งแต่ผู้ช่วยศาสตราจารย์ ดร. ณัฏฐภรณ์ หลาวทอง อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้สละเวลาอันมีค่าให้คำปรึกษา พร้อมทั้งให้ข้อเสนอแนะเพื่อแก้ไข ปรับปรุงข้อบกพร่องต่างๆ ด้วยความเอาใจใส่ และให้กำลังใจแก่ผู้วิจัยมาโดยตลอด

ขอกราบขอบพระคุณ ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี ที่ได้กรุณาให้ความรู้ที่มีคุณค่าซึ่งเป็นประโยชน์อย่างมากในการทำวิจัยครั้งนี้และผู้ช่วยศาสตราจารย์ ดร.อวยพร เรืองตระกูล ที่ได้กรุณาให้คำแนะนำในการแก้ไข ปรับปรุงให้วิทยานิพนธ์ฉบับนี้มีความสมบูรณ์มากยิ่งขึ้น นอกจากนี้ผู้วิจัยขอกราบขอบพระคุณอาจารย์ประภาส และคณาจารย์ทุกท่าน ที่ได้ประสิทธิ์ประสาทวิชาความรู้อันเป็นประโยชน์ให้แก่ผู้วิจัยตั้งแต่อดีตจนถึงปัจจุบัน

ขอกราบขอบพระคุณ คุณพ่อเอกชัยและคุณแม่ดวงรัตน์ เกษรปทุมานันท์ ที่รักและเคารพอย่างสูง ผู้ที่คอยเป็นกำลังใจและให้พลังในการดำเนินชีวิตแก่ผู้วิจัย รวมทั้งให้ความช่วยเหลือในทุกด้านรวมถึงการเก็บรวบรวมข้อมูล ให้ทุนทรัพย์ในการทำวิจัยครั้งนี้ และเห็นความสำคัญของการศึกษาให้การสนับสนุนส่งเสริมการศึกษาของผู้วิจัยเป็นอย่างดียิ่งมาโดยตลอด

ขอกราบขอบพระคุณ ดร. รังสรรค์ มณีเล็ก และดร. เกียรติศักดิ์ ส่องแสงที่อนุญาตให้ใช้คลังข้อสอบวิชาคณิตศาสตร์ รวมทั้ง ดร. กิตติพงศ์ บุญโล่ง อาจารย์กนกวรรณ รัตนธน คุณโสพส สุขานนท์สวัสดิ์ ที่ให้ข้อเสนอแนะที่เป็นประโยชน์อย่างยิ่งต่อการทำวิจัยครั้งนี้

ขอกราบขอบพระคุณ ผู้เชี่ยวชาญทุกท่านที่ได้กรุณาสละเวลาในการตรวจสอบและให้ข้อเสนอแนะที่เป็นประโยชน์ในการพัฒนาเครื่องมือที่ใช้ในการวิจัย ขอขอบพระคุณผู้บริหารโรงเรียน คุณครูและนักเรียนในโรงเรียนที่เป็นกลุ่มตัวอย่างซึ่งได้สละเวลาอันมีค่าให้ความร่วมมือเป็นอย่างดียิ่งจนการเก็บรวบรวมข้อมูลสามารถสำเร็จได้ด้วยดี

ขอขอบคุณเป็นพิเศษสำหรับ คุณอธิษฐ์ ลิมมโนนุกูลและครอบครัวที่คอยให้ความช่วยเหลือ ห่วงใยและเป็นกำลังใจที่ดีแก่ผู้วิจัยเสมอมา รวมทั้งขอขอบคุณ คุณวรรณแสงชัย แซ่เฮง ผู้เชี่ยวชาญและให้คำปรึกษาในการเขียนโปรแกรม

ขอขอบคุณผู้เป็นมหามิตรทุกท่านโดยเฉพาะ คุณปาจารีย์ คุณมารุต คุณกมลชนก คุณฤทัยรัตน์ คุณมนต์ชัย คุณวิสุทธิ ที่ให้ความช่วยเหลือในการเก็บรวบรวมข้อมูลและคีย์ข้อมูล รวมถึงคุณปฏิภาณี คุณเฉลิมฤทธิ์ คุณพิริดา คุณศิริวรรณและเพื่อนในภาควิชาวิจัยและจิตวิทยาการศึกษาทุกท่านที่เป็นกำลังใจและแลกเปลี่ยนความรู้อันมีค่ากับผู้วิจัย สุดท้ายนี้คุณค่าและประโยชน์ที่อาจมีต่อผู้อื่นของวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแด่ ผู้มีพระคุณทุกท่าน

ต้นฉบับไม่มีหน้าสารบัญ

NO CONTENTS PAGE IN ORIGINAL

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การปฏิรูปการศึกษาตามแนวพระราชบัญญัติการศึกษาแห่งชาติ พุทธศักราช 2542 มุ่งพัฒนาสังคมไทยให้เป็นสังคมแห่งการเรียนรู้ เพื่อเป็นฐานรองรับการปรับตัวเข้าสู่ระบบเศรษฐกิจองค์ความรู้ (Society and Economy Knowledge - based) จึงจำเป็นต้องใช้กระบวนการศึกษา เป็นเครื่องมือพัฒนาคุณภาพของคนไทย การจัดการศึกษาจำเป็นต้องจัดสาระการเรียนรู้ จัดกระบวนการเรียนรู้และการประเมินการเรียนรู้ให้เหมาะสมและสอดคล้องกัน โดยมีจุดมุ่งหมายสำคัญเพื่อให้ผู้เรียนเกิดการเรียนรู้และพัฒนาสูงสุด เป็นคนเก่ง คนดี และมีความสุข (บุญเชิด ภิญโญอนันตพงษ์, 2544; เอกรินทร์ สี่มหาศาล, 2546)

สำหรับแนวทางการประเมินการเรียนรู้ในพระราชบัญญัติการศึกษาแห่งชาติ พุทธศักราช 2542 หมวด 4 มาตรา 26 มีใจความว่า ให้สถานศึกษาจัดการประเมินผู้เรียนโดยพิจารณาจากพัฒนาการของผู้เรียน ความประพฤติ การสังเกตพฤติกรรมการเรียน การร่วมกิจกรรม และการทดสอบควบคู่ไปในกระบวนการเรียนการสอนตามความเหมาะสมของแต่ละระดับและรูปแบบการศึกษา และนำผลการประเมินดังกล่าวมาใช้ประกอบการพิจารณาในการจัดสรรโอกาสเข้าศึกษาต่อด้วยวิธีการหลากหลาย (สำนักงานคณะกรรมการการศึกษาแห่งชาติ, 2542; บุญเชิด ภิญโญอนันตพงษ์, 2544) ดังนั้นการวัดและประเมินผลจึงเป็นองค์ประกอบที่สำคัญอย่างหนึ่งในกระบวนการจัดการศึกษาเพื่อให้ทราบผลของการจัดการศึกษาว่าบรรลุเป้าหมายตามที่ได้กำหนดไว้หรือไม่ มากน้อยเพียงไร ในการวัดผลทางการศึกษาเป็นการวัดทางพฤติกรรมศาสตร์ หรือลักษณะทางจิตวิทยา ซึ่งไม่สามารถวัดคุณลักษณะหรือความสามารถของผู้สอบได้โดยตรงเหมือนกับการวัดทางกายภาพ ดังนั้นนักวัดผลจึงได้พยายามสร้างเครื่องมือและพัฒนาทฤษฎีการวัดต่าง ๆ ขึ้นมาเพื่อที่จะอธิบายหรือทำนายคุณลักษณะที่แท้จริงของบุคคล ให้มีความถูกต้องแม่นยำมากที่สุดหรือมีความคลาดเคลื่อนน้อยที่สุด เครื่องมือที่นิยมใช้กันอยู่ทั่วไปคือ แบบสอบประเพณีนิยม (Convention test) ซึ่งถือว่าเป็นเครื่องมือตรวจสอบทางการศึกษาที่ให้บุคคลได้แสดงพฤติกรรมการเรียนรู้ของตนออกมา (อุทุมพร จามรมาน, 2535; วสันต์ ทองไทย, 2539; รังสรรค์ มณีเล็ก, 2540)

การจัดชุดของแบบทดสอบแบบประเพณีนิยม (Convention test) ให้แก่กลุ่มผู้สอบที่มีจำนวนข้อสอบแน่นอนและใช้สำหรับสอบกับผู้สอบทุกคน โดยผู้สอบทุกคนต้องทำข้อสอบทุกข้อเหมือนกันในเวลาเดียวกัน แบบสอบประกอบด้วยข้อสอบที่มีความยากและง่ายแตกต่างกัน ผลที่ได้

จากการทดสอบซึ่งกำหนดให้อยู่ในรูปของคะแนนที่ตัดสินความสามารถของผู้สอบ ซึ่งการดำเนินการสอบตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical test theory) นี้มีข้อจำกัดหลายประการ ได้แก่ แบบสอบมีความเหมาะสมกับผู้สอบที่มีความสามารถระดับปานกลาง ไม่เหมาะกับผู้สอบที่มีความสามารถสูงหรือต่ำ เพราะให้สารสนเทศเกี่ยวกับผู้สอบน้อย (Vale and Weiss, 1975 อ้างถึงใน รังสรรค์ มณีเล็ก, 2540) และมีข้อสอบบางส่วนที่ง่ายเกินไปสำหรับผู้สอบที่มีความสามารถสูง และบางส่วนยากเกินไปสำหรับผู้สอบที่มีความสามารถต่ำ ข้อสอบเหล่านี้อาจทำให้ผู้สอบเสียเวลาในการตอบข้อสอบที่ไม่เหมาะสมกับความสามารถของตนเอง ซึ่งจะทำให้ผู้สอบเกิดความเมื่อยล้าเบื่อหน่ายก่อให้เกิดการเดาในการตอบข้อที่ยากเกินไปสำหรับผู้ที่มีความสามารถต่ำ และเกิดความสะเพร่าสำหรับการตอบข้อที่ง่ายเกินไปสำหรับผู้ที่มีความสามารถสูง ซึ่งนำไปสู่การเกิดความคลาดเคลื่อนในการวัดที่สูงขึ้นได้ ความพยายามในการแก้ไขข้อจำกัดของแบบสอบประเพณีนิยมด้วยการจัดชุดแบบสอบให้เหมาะสมกับความสามารถของผู้สอบแต่ละบุคคลได้ เพื่อให้มีความคลาดเคลื่อนในการประมาณค่าลดลง (ศิริชัย กาญจนวาสี, 2545)

Lord and Novick (1968) ได้เสนอแนะว่า แบบสอบที่ใช้ในการทดสอบควรมีความเหมาะสมกับความสามารถของผู้สอบ นั่นคือ ผู้ที่มีความสามารถต่ำควรจะได้ทำข้อสอบข้อที่ง่าย และผู้ที่มีความสามารถสูงควรจะได้ทำข้อสอบข้อที่ยาก จากหลักการนี้ก่อให้เกิดการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive testing) ขึ้นมาเพื่อแก้ข้อจำกัดต่างๆที่พบข้างต้น โดยมีเป้าหมายที่จะดำเนินการสอบด้วยการคัดเลือกข้อสอบให้เหมาะสมกับความสามารถของผู้สอบแต่ละบุคคล โดยอาศัยทฤษฎีการตอบสนองข้อสอบ (Item response theory) มาใช้เป็นพื้นฐานในการดำเนินการสอบ ซึ่งจะมีการคัดเลือกข้อสอบที่สามารถให้สารสนเทศอันเป็นประโยชน์เกี่ยวกับความสามารถของผู้สอบแต่ละบุคคลและทำให้การทดสอบเป็นไปอย่างมีประสิทธิภาพ นอกจากนี้ยังใช้จำนวนข้อสอบที่สั้นลง ประหยัดเวลาของการทดสอบ มีความคล่องตัว ผลที่ได้มีความคลาดเคลื่อนต่ำ และสามารถนำผลมาเปรียบเทียบกันอย่างมีความหมาย (สายชล อบทม, 2539)

หลักการของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ เป็นการคัดเลือกข้อสอบให้เหมาะกับความสามารถของผู้สอบในแต่ละบุคคลโดยอยู่บนพื้นฐานของผลการตอบข้อสอบข้อที่ผ่านมาของผู้สอบนั้น เมื่อผู้สอบทำข้อสอบข้อเริ่มแรกหรือชุดแรกที่มาจากการคัดเลือกจากคลังข้อสอบแล้ว จะมีการวิเคราะห์ระดับความสามารถหรือประเมินความสามารถของผู้สอบเบื้องต้น เพื่อคัดเลือกข้อสอบข้อต่อไปที่มีค่าความยากและอำนาจจำแนกเหมาะสมที่จะใช้วัดระดับความสามารถของผู้สอบ และประมาณระดับความสามารถของผู้สอบใหม่ จากนั้นก็จะเลือกข้อสอบที่เหมาะสมข้อต่อไป โดยอาศัยหลักที่ว่า ถ้าการทำข้อสอบที่ผ่านมาถูก ข้อถัดไปจะยากขึ้น แต่ถ้าทำข้อสอบที่ผ่านมาผิด ข้อถัดไปจะง่ายลง กระบวนการนี้จะดำเนินการต่อไปเรื่อยจนสามารถประมาณ

ระดับความสามารถของผู้สอบได้อย่างน่าเชื่อถือและมีความคลาดเคลื่อนต่ำตามเกณฑ์ที่กำหนดไว้ การทดสอบจึงจะยุติ (ศิริชัย กาญจนวาสี, 2545)

ในปัจจุบันได้มีการนำคอมพิวเตอร์ และทฤษฎีการตอบสนองข้อสอบ (Item response theory) มาประยุกต์ใช้กับการทดสอบแบบปรับเหมาะ (Adaptive testing) ซึ่งเรียกว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ (Computerized adaptive testing) โดยนำคอมพิวเตอร์เข้ามาช่วยในการวิเคราะห์ข้อมูลและดำเนินการสอบ ซึ่งสามารถใช้ในการวัดความสามารถของผู้สอบในด้านต่าง ๆ รวมถึงการวัดบุคลิกภาพ และเจตคติ และใช้ได้ในการทดสอบแบบมีส่วนได้เสียสูง (High stakes tests) เช่น การสอบคัดเลือกขอหนังสือรับรองวุฒิ หรือการสอบใบอนุญาต เป็นต้น (Richard C., 2005)

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบนับว่ามีประโยชน์ต่อการวัดผลการศึกษาเป็นอย่างมาก เนื่องจากประหยัดค่าใช้จ่าย เพราะใช้ข้อสอบน้อยกว่าการทดสอบด้วยกระดาษและดินสอและไม่จำเป็นต้องสร้างข้อสอบใหม่ทุกครั้งที่มีการสอบ เนื่องจากมีข้อสอบเก็บไว้ในคลังข้อสอบอยู่แล้ว ผู้สอบไม่เกิดความเหน็ดเหนื่อยและความเบื่อหน่ายในการสอบ เพราะใช้เวลาในการสอบน้อยลงกว่าเดิม ข้อสอบที่ใช้มีความยากง่ายเหมาะสมกับความสามารถของผู้สอบ และสิ่งที่สำคัญที่สุด คือ ได้ผลการทดสอบที่มีความถูกต้องแม่นยำมากกว่าเดิม โดยเฉพาะกับกลุ่มผู้สอบที่มีความสามารถสูงหรือต่ำมาก ๆ (Hulin and others, 1983)

จากการศึกษาการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มีผู้ให้ความสนใจวิธีดำเนินการสอบ ทำการศึกษาค้นคว้า และพัฒนารูปแบบการทดสอบเพื่อให้เป็นการทดสอบที่มีประสิทธิภาพ ซึ่งผลการวิจัยมีด้วยกันหลายประเด็น เช่น

เรื่องโมเดลของการตอบสนองข้อสอบ (Item response model) พบว่า โมเดลการตอบสนองข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบโลจิสติก 3 พารามิเตอร์นั้นเหมาะสมกว่าโมเดลอื่นๆ (Urry, 1977)

เรื่องคลังข้อสอบ(Item pool) พบว่า ขนาดของคลังข้อสอบส่งผลต่อการประมาณค่าความสามารถของผู้สอบน้อยมาก (Ho,1989) และขนาดของคลังข้อสอบนั้นควรมีประมาณ 100-200 ข้อ และคลังข้อสอบ 116 -150 ข้อจะให้ผลดีที่สุด (Weiss,1988) ซึ่งสอดคล้องกับ Urry (1977) ที่พบว่า คลังข้อสอบที่วัดคุณลักษณะเดียวกันควรมีอย่างน้อย 100 ข้อ ส่วนข้อสอบในคลังข้อสอบนั้นควรมีค่าอำนาจจำแนกมากกว่า 0.8 ค่าความยากง่ายอยู่ระหว่าง -2.0 ถึง 2.0 และค่าสัมประสิทธิ์การเดาน้อยกว่า 0.3 นอกจากนั้น Ho (1989) พบว่า ชนิดของคลังข้อสอบ (ข้อสอบแบบสุ่ม ข้อสอบที่มีค่าอำนาจจำแนกสูง และข้อสอบที่ง่าย ๆ)นั้น มีปฏิสัมพันธ์กับวิธีการประมาณค่าความสามารถของผู้สอบ

เรื่องความยากง่ายของข้อสอบข้อแรก (Entry level) Lord (1980) พบว่าค่าความยากง่ายของข้อสอบข้อแรกที่เลือกให้แก่ผู้สอบนั้นไม่มีผลต่อความถูกต้องในการประมาณค่าความสามารถของผู้สอบ แต่ Ho (1989) พบว่า ถ้าข้อสอบข้อแรกมีความยากง่ายเท่ากับหรือต่ำกว่าความสามารถของผู้สอบแล้วจะทำให้การประมาณค่าความสามารถของผู้สอบมีความถูกต้องมากขึ้น

เรื่องเกณฑ์การคัดเลือกข้อสอบ (Item selection rule) Hulin and other (1983) พบว่า การใช้เกณฑ์การคัดเลือกข้อสอบโดยวิธีสารสนเทศสูงสุด (Maximum information) มีประสิทธิภาพกว่าการใช้เกณฑ์การคัดเลือกข้อสอบโดยวิธีอื่น

เรื่องเกณฑ์การคัดเลือกข้อสอบขั้นแรก (Item selection rules at the early stage) Veerkamp and Berger (1997) พบว่า วิธีสารสนเทศของฟิชเชอร์ที่มีการถ่วงน้ำหนัก (FII) มีประสิทธิภาพไม่ดีกว่าวิธีสารสนเทศของฟิชเชอร์ (FI) และ Fan and Hsu (1996 cited in Chen, Ankenmann and Chang, 2000) พบว่า วิธีสารสนเทศของคูลเบค-ไลเบลอร์ (KL) มีประสิทธิภาพไม่ดีกว่าวิธีสารสนเทศของฟิชเชอร์ (FI) ซึ่งไม่สอดคล้องกับ Chang and Ying (1996) พบว่า วิธีสารสนเทศของคูลเบค-ไลเบลอร์ (KL) มีประสิทธิภาพดีกว่าวิธีสารสนเทศของฟิชเชอร์ (FI) และChen, Ankenmann and Chang (2000) พบว่า วิธีสารสนเทศของฟิชเชอร์ที่มีการถ่วงน้ำหนัก (FII) วิธีสารสนเทศของฟิชเชอร์ที่มีการแจกแจงภายหลัง (FIP) วิธีสารสนเทศของคูลเบค-ไลเบลอร์ (KL) และวิธีสารสนเทศของคูลเบค-ไลเบลอร์ที่มีการแจกแจงภายหลัง (KLP) มีประสิทธิภาพและความแม่นยำในการประมาณค่าคุณลักษณะที่อยู่ปลายสุด (θ = -3,-2) มากกว่า วิธีสารสนเทศของฟิชเชอร์ (FI) และเมื่อแบบสอบที่มีความยาวมากกว่า 10 ข้อไม่ปรากฏข้อได้เปรียบเกี่ยวกับความแม่นยำในการประมาณคุณลักษณะสำหรับเกณฑ์การคัดเลือกข้อสอบ และ Chen and Ankenmann (2004) พบว่า เมื่อความยาวของแบบสอบน้อยกว่า 10 ข้อ วิธีสารสนเทศของฟิชเชอร์ที่มีการแจกแจงภายหลัง (FIP) และวิธีสารสนเทศของคูลเบค-ไลเบลอร์ที่มีการแจกแจงภายหลัง (KLP) มีแนวโน้มให้ผลดีกว่าวิธีสารสนเทศของฟิชเชอร์ (FI) ที่ระดับของคุณลักษณะที่อยู่ปลายสุด (extreme) อย่างไรก็ดีในการตั้งค่าให้สมจริง (realistic) จากการจำลองข้อมูลไม่สามารถสรุปได้ว่าวิธีสารสนเทศของฟิชเชอร์ที่มีการแจกแจงภายหลัง (FIP) และวิธีสารสนเทศของคูลเบค-ไลเบเลอร์ที่มีการแจกแจงภายหลัง (KLP) จะมีประสิทธิภาพดีกว่าวิธีสารสนเทศของฟิชเชอร์ (FI) โดยเฉพาะอย่างยิ่งเมื่อได้กำหนดการควบคุมอัตราการใช้ข้อสอบซ้ำเข้าไป เมื่อความยาวของแบบสอบที่มากกว่า 10 ข้อ วิธีการคัดเลือกข้อสอบแบบสารสนเทศของฟิชเชอร์ (FI) สารสนเทศของฟิชเชอร์ที่มีการแจกแจงภายหลัง (FIP) และสารสนเทศของคูลเบค-ไลเบลอร์ที่มีการแจกแจงภายหลัง (KLP) นี้จะมีลักษณะที่ไม่แตกต่างกัน ในขณะวิธีสารสนเทศของฟิชเชอร์ (FI) จะมีการใช้ข้อสอบสูงกว่าเล็กน้อย

เรื่องวิธีการประมาณค่าความสามารถของผู้สอบ (Ability estimation procedure) ต่าย เซี่ยงฉี (2534) พบว่า วิธีการประมาณค่าความสามารถตามวิธีของเบส์ที่ปรับปรุงใหม่ (Beyesian updating) และวิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข(Conditional maximum likelihood)ให้ผลไม่แตกต่างกัน และพบว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบรูปปิรามิดที่มีรูปแบบ จำนวนขั้น และวิธีการให้คะแนนต่างกันจะส่งผลต่อความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ Skaggs and Stevenson (1989 อ้างถึงในรังสรรค์ มณีเล็ก, 2540) ที่พบว่า หากผู้สอบมากกว่า 2000 คนแล้ว ผลที่ได้รับจากการประมาณค่าความสามารถทั้ง 2 วิธีไม่แตกต่างกัน แต่ถ้ามีผู้สอบน้อยกว่า 500 คนแล้ว วิธีการของเบส์จะมีความคงที่และถูกต้องมากกว่า และถ้าข้อสอบที่ใช้ในการสอบน้อยกว่า 15 ข้อแล้ว วิธีความเป็นไปได้สูงสุด จะมีความถูกต้องมากกว่า มีบางส่วนสอดคล้องกับ ข้อค้นพบ Gifford and Swaminathan (1990) พบว่า ถ้ากลุ่มผู้สอบมีขนาดเล็กแล้วจำนวนผู้ใช้มีจำนวนน้อยข้อแล้วการประมาณค่าความสามารถด้วยวิธีของเบส์แบบร่วมกัน (Joint bayesian) จะมีความถูกต้องมากกว่า

เรื่องประสิทธิภาพการทดสอบ Balckmore (1987) พบว่า คะแนนที่ได้จากการทดสอบด้วยกระดาษและดินสอ การทดสอบด้วยคอมพิวเตอร์ และการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ไม่แตกต่างกัน และนันทิยา พึ่งคำ (2531) พบว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จะให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่ำกว่าและใช้จำนวนข้อน้อยกว่าการทดสอบแบบประเพณีนิยม รวมถึง รังสรรค์ มณีเล็ก (2540) พบว่า วิธีการประมาณค่าความสามารถของผู้สอบ เกณฑ์ยุติการทดสอบ และความสามารถของผู้สอบ มีผลต่อความตรงตามสภาพ และความยากง่ายของข้อสอบข้อแรก เกณฑ์การคัดเลือกข้อสอบ วิธีประมาณค่าความสามารถของผู้สอบ เกณฑ์ยุติการทดสอบ และความสามารถของผู้สอบส่งผลต่อจำนวนข้อที่ใช้ในการทดสอบ นอกจากนี้กนกวรรณ รัตนธน (2544) พบว่า ความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำที่แตกต่างกันส่งผลต่อความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบ

เรื่องอัตราการใช้ข้อสอบซ้ำ Chang (1998) พบว่าวิธีการควบคุมอัตราการใช้ข้อสอบซ้ำด้วยวิธีของ Sympson and Hetter มีประสิทธิภาพมาก และวิธีของ Stocking and Lewis ทำให้มีความปลอดภัยของแบบสอบ เนื่องมาจากการใช้ข้อสอบซ้ำ แต่ลดความถูกต้องของการทดสอบโดยที่การสอบแต่ละครั้งควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 20 เปอร์เซ็นต์ของจำนวนข้อสอบที่ผู้สอบได้รับผ่านมา ส่วน Kalohn and Sparay (1998) พบว่า วิธีการคัดเลือกข้อสอบโดยมีอัตราการใช้ข้อสอบซ้ำ 10 และ 20 เปอร์เซ็นต์ มีความถูกต้องเหมือนกันและผู้สอบที่มีความสามารถระดับสูงควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 10 เปอร์เซ็นต์ของจำนวนข้อสอบที่ผู้สอบได้รับในการทดสอบครั้งที่ผ่านมา

เรื่องเกณฑ์ยุติการทดสอบ จากการศึกษาของ Dodd and Koch (1993) โดยใช้ Partial credit model พบว่า การยุติการทดสอบโดยดูจากค่าความคลาดเคลื่อนมาตรฐานจะให้ผลดีกว่าการยุติการทดสอบโดยการกำหนดจำนวนข้อสอบ ซึ่งสอดคล้องกับ การจำลองข้อมูลของ Dodd, Koch and De Alaya (1989 cited in Hot, Vorst and Mellenbergh, 2005) โดยใช้ Graded response model พบว่า เกณฑ์ยุติการทดสอบโดยใช้ค่าความคลาดเคลื่อนมาตรฐาน จะให้ผลดีกว่าเกณฑ์ยุติการทดสอบที่ใช้ค่าสารสนเทศของข้อสอบ และตรึงใจ พูลผลอำนวย (2534) ได้พัฒนาแบบสอบเฉพาะบุคคล (Tailored test) โดยใช้รูปแบบแยกทางคงที่รูปปิรามิดขนาดขั้นคงที่ 10 ข้น พบว่า ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของแบบสอบเฉพาะบุคคลท่ากับ 0.443 นอกจากนี้ผลการศึกษาของรังสรรค์ มณีเล็ก (2540) พบว่า การกำหนดค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 จะให้ค่าความตรงตามสภาพสูงสุด

จากผลการวิจัยดังกล่าว พบว่า บางประเด็นมีความชัดเจนว่าควรมีลักษณะเช่นไรจึงทำให้การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบมีประสิทธิภาพสูงสุด ส่วนประเด็นที่ยังไม่ค่อยชัดเจนคือ อัตราการใช้ข้อสอบซ้ำ และเกณฑ์การคัดเลือกข้อสอบขั้นแรกซึ่งเป็นประเด็นที่สำคัญในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (CAT) คือการที่เราสามารถที่จะเลือกข้อสอบที่ดีที่สุดจากคลังข้อสอบ (Item pool) ทั้งหมดให้เหมาะสมได้อย่างไร แต่เดิมขั้นตอนวิธีการเลือกข้อสอบต้องอาศัยสารสนเทศของข้อสอบอย่างมาก นั่นหมายความว่าข้อสอบที่ถูกคัดเลือกจะมีสารสนเทศของฟิชเชอร์ (FI) สูงในการประมาณค่าความสามารถของผู้สอบในระหว่างที่ตอบข้อสอบที่ได้ใช้ไปก่อนหน้านี้ นอกจากนี้ยังสังเกตได้ว่าเกณฑ์สารสนเทศนี้อาจเป็นสาเหตุที่ทำให้เกิดการใช้ข้อสอบซ้ำที่ไม่ได้สัดส่วน (Davey and Parshall, 1995; Sympson and Hetter, 1985; van der Linden, 1998 cited in Leung, Cnang and Hau, 2003) และข้อสอบแต่ละข้อจะถูกเลือกตามการประมาณค่าความสามารถอย่างต่อเนื่องและจับคู่ให้เหมาะสมกับระดับความสามารถ แต่ในการคัดเลือกข้อสอบขั้นแรกยังไม่ทราบการประมาณค่าความสามารถแน่ชัด ดังนั้นข้อสอบที่ถูกเลือกมาที่ขั้นแรกจึงมีแนวโน้มที่จะเหมาะสมน้อยกว่าข้อสอบที่ถูกเลือกในขั้นต่อมา จึงมีผู้ปรับปรุงเกณฑ์การคัดเลือกข้อสอบขั้นแรก และมีผู้เปรียบเทียบเกณฑ์การคัดเลือกข้อสอบขั้นแรกดังที่ได้กล่าวไว้ข้างต้น

จากการศึกษาพบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรก ได้แก่ วิธีสารสนเทศของฟิชเชอร์ (FI) วิธีสารสนเทศของฟิชเชอร์ที่มีการถ่วงน้ำหนัก (FII) วิธีสารสนเทศของฟิชเชอร์ที่มีการแจกแจงภายหลัง (FIP) วิธีสารสนเทศของคูลเบค-ไลเบลอร์ (KL) และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ที่มีการแจกแจงภายหลัง (KLP) ในวิธีการคัดเลือกข้อสอบขั้นแรกเหล่านี้ยังมีประเด็นที่ไม่ชัดเจนในเรื่องการเปรียบเทียบวิธีสารสนเทศของฟิชเชอร์ (FI) และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ (KL) ส่วนวิธีสารสนเทศของฟิชเชอร์ที่มีการถ่วงน้ำหนัก (FII) วิธีสารสนเทศของฟิชเชอร์ที่มีการ

แจกแจงภายหลัง (FIP) วิธีสารสนเทศของคูลเบค-ไลเบลอร์ (KL) และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ที่มีการแจกแจงภายหลัง (KLP) ให้ผลในลักษณะที่ไม่แตกต่างกัน ดังนั้นผู้วิจัยจึงสนใจที่จะศึกษาวิธีสารสนเทศของฟิชเชอร์ (FI) และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ (KL) และจากการศึกษาของ กนกวรรณ รัตนธน (2544) พบว่า อัตราการใช้ข้อสอบซ้ำที่ต่างกันส่งผลต่อความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบ นอกจากนี้ รังสรรค์ มณีเล็ก (2540) พบว่า เกณฑ์ยุติการทดสอบ มีผลต่อความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบ และเป็นที่น่าสังเกตว่าในงานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบนั้น มีการใช้เกณฑ์ยุติการทดสอบที่แตกต่างกัน โดยเฉพาะการกำหนดค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่างกัน เช่น การพัฒนาโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ สายชล อบทม (2539) ได้กำหนดค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่ำกว่า 0.3 แต่โสพส สุขานนท์สวัสดิ์ (2545) กำหนดค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่ำกว่า 0.45 เป็นเกณฑ์ยุติการทดสอบ และจากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับเกณฑ์ยุติการทดสอบ พบว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (CAT) ส่วนใหญ่จะมีค่าอยู่ในช่วงประมาณ 0.3 ถึง 0.5 ดังนั้นผู้วิจัยจึงสนใจที่จะศึกษาเกณฑ์ยุติการทดสอบโดยกำหนดค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ น้อยกว่าหรือเท่ากับ 0.3 และ 0.45 เป็นเกณฑ์ยุติการทดสอบ จากที่ได้กล่าวมาข้างต้น ผู้วิจัยจึงสนใจในประเด็นที่ว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกต่างกัน อัตราการใช้ข้อสอบซ้ำต่างกัน และเกณฑ์ยุติการทดสอบต่างกัน จะส่งผลต่อ ความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์หรือไม่ มากน้อยเพียงไร และจากการศึกษางานวิจัยที่เกี่ยวข้องยังไม่พบว่ามีผู้ใดศึกษาเรื่องนี้มาก่อน ดังนั้นผู้วิจัยจึงมีความสนใจที่จะเปรียบเทียบความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ที่ใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกต่างกัน อัตราการใช้ข้อสอบซ้ำต่างกัน และเกณฑ์ยุติการทดสอบต่างกัน เพื่อจะเป็นประโยชน์ต่อการศึกษาประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบและพิจารณาเลือกใช้หรือปรับปรุงเกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ต่อไป

คำถามวิจัย

ความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จะต่างกันหรือไม่อย่างไร เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่ต่างกัน

วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่ต่างกัน

สมมติฐานการวิจัย

จากการศึกษาเอกสาร ตำรา และงานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่า การเลือกข้อสอบข้อแรกให้กับผู้สอบนั้น ลักษณะของข้อสอบข้อแรกไม่มีผลต่อความถูกต้องในการประมาณค่าความสามารถของผู้สอบเมื่อสิ้นสุดการทดสอบ (Lord, 1980) และ ความถูกต้องในการประมาณค่าความสามารถของผู้สอบหรือค่าฟังก์ชันสารสนเทศของแบบสอบนั้นมีความสัมพันธ์แบบผกผันกับค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบ จึงทำให้ค่าความตรงของแบบทดสอบไม่ต่างกัน แต่เมื่อกำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่างกัน จึงทำให้ค่าความตรงของแบบทดสอบต่างกัน และ Kalohn and Sparay (1998) พบว่า วิธีการคัดเลือกข้อสอบโดยมีอัตราการใช้ข้อสอบซ้ำ 10 และ 20 เปอร์เซ็นต์ มีความถูกต้องเหมือนกันและผู้สอบที่มีความสามารถระดับสูงควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 10 เปอร์เซ็นต์ของจำนวนข้อสอบที่ผู้สอบได้รับในการทดสอบครั้งที่ผ่านมา ดังนั้นผู้วิจัยจึงตั้งสมมติฐานการวิจัยดังต่อไปนี้

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกต่างกัน อัตราการใช้ข้อสอบซ้ำต่างกัน และเกณฑ์ยุติการทดสอบต่างกัน น่าจะทำให้ค่าความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แตกต่างกัน

ขอบเขตของการวิจัย

1. แบบสอบวัดความสามารถทางคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 6 โดยวัดความเข้าใจเกี่ยวกับพื้นฐานทางจำนวน พีชคณิต การวัด เรขาคณิต เป็นแบบเลือกตอบ 4 ตัวเลือก จำนวน 40 ข้อ เพื่อเป็นตัวแปรเกณฑ์ในการศึกษาความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

2. คลังข้อสอบ เป็นกลุ่มข้อสอบที่ใช้วัดความเข้าใจ เรื่อง เศษส่วน ระดับชั้นประถมศึกษาปีที่ 6 มีทั้งหมด 244 ข้อพัฒนาโดย รังสรรค์ มณีเล็ก (2540) จำนวน 200 ข้อและเกียรติศักดิ์ ส่องแสง (2547) จำนวน 44 ข้อ ลักษณะของข้อสอบเป็นแบบเลือกตอบ 4 ตัวเลือก มีค่าอำนาจจำแนกอยู่ระหว่าง 0.41 ถึง 2.34 ค่าความยากง่ายอยู่ในช่วง -2.23 ถึง 2.72 และค่าการเดาอยู่ในช่วง 0.06 ถึง 0.30

3. โปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ คือ โปรแกรมที่มีการกำหนดรูปแบบของการทดสอบเป็นแบบหลายขั้นตอน (Multi-stage strategies) แบบทางแยกแปรผัน (Variable branching model) และประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์

4. ตัวแปรที่ศึกษาในครั้งนี้ประกอบด้วย

4.1 ตัวแปรอิสระ ได้แก่

4.1.1 เกณฑ์การคัดเลือกข้อสอบขั้นแรก

1) วิธีสารสนเทศของฟิชเชอร์

2) วิธีสารสนเทศของคูลเบค-ไลเบลอร์

4.1.2 อัตราการใช้ข้อสอบซ้ำ

1) อัตราการใช้ข้อสอบซ้ำ เท่ากับ 10 เปอร์เซ็นต์

2) อัตราการใช้ข้อสอบซ้ำ เท่ากับ 20 เปอร์เซ็นต์

3) อัตราการใช้ข้อสอบซ้ำ เท่ากับ 30 เปอร์เซ็นต์

4.1.3 เกณฑ์ยุติการทดสอบ

1) ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ เท่ากับหรือน้อยกว่า 0.3

2) ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ เท่ากับหรือน้อยกว่า 0.45

4.2 ตัวแปรตาม ได้แก่

ความตรงตามสภาพของการประมาณค่าความสามารถของผู้สอบ

ข้อตกลงเบื้องต้น

ผู้สอบมีความเข้าใจในวิธีการทดสอบด้วยเครื่องคอมพิวเตอร์และมีความสามารถในการใช้คอมพิวเตอร์

ข้อจำกัดของการวิจัย

1. ค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบมีการกระจายเป็นช่วง ๆ ด้วยจำนวนข้อสอบที่แตกต่างกัน จึงทำให้การคัดเลือกข้อสอบในบางครั้งจากคลังข้อสอบไม่ได้ข้อสอบตามที่ต้องการและจำเป็นต้องใช้ข้อสอบที่มีคุณลักษณะใกล้เคียงขึ้นมาแทน

2. โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ใช้ในการวิจัยนี้ ได้กำหนดค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ เท่ากับหรือน้อยกว่า 0.45 และ 0.3 และผู้วิจัยได้กำหนดจำนวนข้อสอบที่ใช้ในโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ไว้สูงสุดจำนวน 40 ข้อ สำหรับผู้สอบที่ได้ทำข้อสอบจำนวนมาก แล้วค่าความคลาดเคลื่อนมาตรฐานยังไม่น้อยกว่าหรือเท่ากับ 0.3 ซึ่งอาจทำให้ผู้สอบเกิดความเบื่อหน่ายและไม่ตั้งใจทำข้อสอบได้ จะถูกคัดชื่อออกจากการประมาณค่าความสามารถของผู้สอบ เนื่องจากข้อสอบชุดนี้ไม่เหมาะกับความสามารถของผู้สอบดังกล่าว

3. คลังข้อสอบที่ใช้ในการวิจัยครั้งนี้ เป็นกลุ่มข้อสอบที่ใช้วัดความเข้าใจ เรื่อง เศษส่วน ระดับชั้นประถมศึกษาปีที่ 6 จำนวนทั้งหมด 244 ข้อซึ่งพัฒนาโดยรังสรรค์ มณีเล็ก (2540) จำนวน 200 ข้อและเกียรติศักดิ์ ส่องแสง (2547) จำนวน 44 ข้อ แต่ข้อสอบจำนวน 44 ข้อซึ่งพัฒนาโดย เกียรติศักดิ์ ส่องแสง (2547) ไม่ได้ปรับเทียบคะแนนให้ข้อสอบอยู่บนสเกลเดียวกันเพื่อจัดเก็บเข้าคลังข้อสอบไปรวมกับข้อสอบจำนวน 200 ข้อซึ่งพัฒนาโดยรังสรรค์ มณีเล็ก (2540) แต่มีการตรวจความเป็นเอกมิติซึ่งตรงตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

คำจำกัดความที่ใช้ในการวิจัย

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ หมายถึง การทดสอบความรู้ทางคณิตศาสตร์ ด้วยข้อสอบที่มีความเหมาะสมกับความสามารถของผู้สอบและดำเนินการสอบโดยใช้โปรแกรมการทดสอบแบบปรับเหมาะตามระดับความสามารถของผู้สอบ ซึ่งพัฒนาโดย เกียรติศักดิ์ ส่องแสง (2547)

ความสามารถทางคณิตศาสตร์ หมายถึง คะแนนที่ได้จากการทดสอบด้วยแบบสอบวัดความสามารถทางคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 6 เป็นแบบตอบเลือกตอบ 4 ตัวเลือก จำนวน 40 ข้อ

ความตรงตามสภาพ หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าความสามารถที่ได้จากการทดสอบด้วยแบบสอบวัดความสามารถทางคณิตศาสตร์ที่ผู้วิจัยพัฒนาขึ้นเองกับค่าความสามารถเฉลี่ยที่ได้จากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

เกณฑ์การคัดเลือกข้อสอบขั้นแรก หมายถึง วิธีการคัดเลือกข้อสอบออกจากคลังข้อสอบเป็นข้อแรกและข้อต่อไปเพื่อให้ผู้สอบได้ตอบ 2 วิธี คือ วิธีสารสนเทศของฟิชเชอร์ (FI) และวิธีสารสนเทศของคูลเบค-ไลเบลอร์ (KL)

อัตราการใช้ข้อสอบซ้ำ หมายถึง สัดส่วนระหว่างจำนวนข้อสอบที่ใช้ซ้ำในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ กับจำนวนข้อสอบทั้งหมดที่ผู้สอบได้รับในการทดสอบครั้งที่ผ่านมา ซึ่งการวิจัยครั้งนี้กำหนดอัตราการใช้ข้อสอบซ้ำที่สูงสุด 3 ระดับ เท่ากับ 10, 20 และ 30 เปอร์เซ็นต์

ค่าความสามารถของผู้สอบ (θ) หมายถึง ระดับความสามารถของผู้สอบแต่ละคน ซึ่งวัดในหน่วย θ ที่มีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 สำหรับค่าความสามารถมีค่าตั้งแต่ $-\infty$ ถึง ∞

คลังข้อสอบ หมายถึง กลุ่มข้อสอบซึ่งสร้างขึ้นตามหลักเกณฑ์ที่ถูกต้องและถูกเก็บไว้อย่างเป็นระบบ คลังที่ใช้ในการวิจัยครั้งนี้ เป็นกลุ่มข้อสอบที่ใช้วัดความรู้ความเข้าใจวิชาคณิตศาสตร์ เรื่อง เศษส่วน ระดับชั้นประถมศึกษาปีที่ 6 ลักษณะของข้อสอบเป็นแบบเลือกตอบ 4 ตัวเลือก จำนวน 244 ข้อ มีค่าอำนาจจำแนก (a) อยู่ระหว่าง 0.41 ถึง 2.34 ค่าความยากง่าย (b) อยู่ในช่วง -2.23 ถึง 2.72 และค่าการเดา (c) อยู่ในช่วง 0.06 ถึง 0.30

โปรแกรมการทดสอบแบบปรับเหมาะตามระดับความสามารถของผู้สอบ หมายถึง ระบบการทดสอบด้วยคอมพิวเตอร์ที่มีการพัฒนากระบวนการสอบ ตั้งแต่การสร้างและจัดเก็บข้อสอบเป็นคลังข้อสอบ รวมถึงการทดสอบและประมวลผล โดยอาศัยคอมพิวเตอร์ พัฒนาโดย เกียรติศักดิ์ ส่องแสง (2547)

เกณฑ์ยุติการทดสอบ หมายถึง ค่าที่กำหนดไว้เพื่อเป็นตัวกำหนดการสิ้นสุดการทดสอบของผู้สอบแต่ละคน การวิจัยครั้งนี้ใช้ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ 2 เกณฑ์ คือ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถเท่ากับหรือน้อยกว่า 0.3 และ 0.45 เป็นเกณฑ์ยุติการทดสอบ

ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ หมายถึง ค่าส่วนเบี่ยงเบนมาตรฐานของผลต่างระหว่างค่าความสามารถจริงกับค่าความสามารถที่ได้จากการประมาณค่า ในการวิจัยครั้งนี้สามารถคำนวณจากการถอดรากที่สองของค่าความแปรปรวนในการทดสอบ โดยกำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 จึงยุติการทดสอบ

ประโยชน์ที่คาดว่าจะได้รับ

1. เป็นการขยายองค์ความรู้ในเชิงทฤษฎีเกี่ยวกับประสิทธิภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

2. เพื่อให้นักวัดผลการศึกษา ครู และอาจารย์ได้ข้อมูลสำหรับเป็นแนวทางในการเลือกใช้รูปแบบของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ที่มีประสิทธิภาพ และสารสนเทศที่ได้จากการวิจัยครั้งนี้จะเป็นประโยชน์ต่อการพิจารณาปรับปรุงหรือเปลี่ยนแปลงเกณฑ์การคัดเลือกข้อสอบขั้นแรกในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

3. ได้เผยแพร่ความรู้เกี่ยวกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ให้แพร่หลายออกไปมากกว่าเดิม เนื่องจากในปัจจุบันคอมพิวเตอร์เข้ามามีบทบาทต่อการจัดการเรียนการสอนในสถานศึกษาระดับต่าง ๆ มากขึ้น ดังนั้นผู้ที่มีหน้าที่เกี่ยวข้องกับการจัดการศึกษา โดยเฉพาะผู้ที่มีหน้าที่โดยตรงเกี่ยวกับการวัดผลการศึกษาควรที่จะมีความรู้ความเข้าใจในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

จากการศึกษาเอกสาร งานวิจัยและตำราต่าง ๆ ในเรื่องที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ผู้วิจัยได้นำเสนอเนื้อหาออกเป็น 3 ตอน คือ

ตอนที่ 1 ทฤษฎีการตอบสนองข้อสอบ

1.1. หลักการของทฤษฎีการตอบสนองข้อสอบ

1.2. ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

1.3. โมเดลของทฤษฎีการตอบสนองข้อสอบ

1.4. ปัจจัยที่มีผลต่อคุณภาพของแบบสอบ

ตอนที่ 2 การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

2.1. ความหมายของการทดสอบแบบปรับเหมาะ

2.2. หลักการของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

2.3. ประเภทของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

2.4. องค์ประกอบของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

2.5. วิธีการคัดเลือกข้อสอบขั้นแรก การควบคุมการใช้ข้อสอบซ้ำ และการสร้างความสมดุลของเนื้อหาในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

2.6. ขั้นตอนในการทดสอบ

2.7. การดำเนินการทดสอบ

2.8. ความแตกต่างระหว่างการทดสอบด้วยกระดาษและดินสอกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

ตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

ในแต่ละตอนมีรายละเอียดดังต่อไปนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 1 ทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ เป็นทฤษฎีที่อธิบายความสัมพันธ์ระหว่างลักษณะหรือความสามารถที่มีอยู่ภายในตัวบุคคลกับพฤติกรรมการตอบสนองข้อสอบของบุคคลนั้น ทฤษฎีนี้มีความเชื่อว่าพฤติกรรมการตอบสนองข้อสอบของผู้สอบ ซึ่งเป็นสิ่งที่สังเกตได้โดยตรงจะถูกกำหนดโดยคุณลักษณะ (Trait) หรือความสามารถ (Ability) ที่มีอยู่ในตัวบุคคลซึ่งไม่สามารถสังเกตได้ และยังเชื่อว่าค่าพารามิเตอร์ต่าง ๆ ของข้อสอบไม่ว่าจะเป็นค่าความยาก (b) ค่าอำนาจจำแนก (a) หรือค่าการเดา (c) ของข้อสอบแต่ละข้อเป็นคุณลักษณะที่มีอยู่ประจำ และคงที่พอสมควรในตัวข้อสอบนั้นจริง ฉะนั้นค่าพารามิเตอร์ของข้อสอบจึงไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่าง และค่าความสามารถของผู้สอบซึ่งเชื่อว่าเป็นคุณลักษณะที่มีอยู่ในตัวผู้สอบนั้นจริงจึงไม่ควรแปรเปลี่ยนไปตามค่าความยากของข้อสอบ เนื่องจากความสามารถของผู้สอบเป็นคุณลักษณะซึ่งไม่สามารถวัดหรือสังเกตได้โดยตรง นักวัดผลทางการศึกษาจึงได้พยายามหาความสัมพันธ์ระหว่างปริมาณการกระทำข้อสอบหรือคะแนน (Test performance or score) กับปริมาณความสามารถ (Ability) โดยแสดงความสัมพันธ์ดังกล่าวในลักษณะโมเดลทางคณิตศาสตร์ที่กำหนดด้วยฟังก์ชันทางคณิตศาสตร์เรียกว่า ฟังก์ชันลักษณะข้อสอบ (Item characteristic function) หรือเรียกว่าโค้งคุณลักษณะของข้อสอบ (Item characteristic curve) หรือฟังก์ชันการตอบสนองข้อสอบ (Item response function) (ศิริชัย กาญจนวาสี, 2538) ซึ่งมีการพัฒนาขึ้นหลายโมเดลด้วยกัน เช่น Nomal ogive model Logistic model และ Rasch model เป็นต้น แต่ละโมเดลนี้จะมีจำนวนพารามิเตอร์ของข้อสอบแตกต่างกันไป กล่าวคือ แบบพารามิเตอร์หนึ่งตัวจะแสดงเฉพาะพารามิเตอร์ค่าความยาก (Difficulty) เท่านั้น แบบพารามิเตอร์สองตัวจะเพิ่มพารามิเตอร์ค่าอำนาจจำแนก (Discrimination) เข้าไปอีกตัวหนึ่ง และแบบพารามิเตอร์สามตัวจะมีทั้งพารามิเตอร์ค่าความยาก ค่าอำนาจจำแนก และค่าการเดา สำหรับการเลือกโมเดลของทฤษฎีการตอบข้อสอบมาใช้นั้น จะต้องเลือกให้เหมาะสมกับสถานการณ์ต่าง ๆ ตามข้อตกลงเบื้องต้นของแต่ละโมเดล โดยทฤษฎีมีจุดเด่นหลายประการ ดังนี้ (Hambleton and Swaminathan, 1985)

ประการแรก ค่าพารามิเตอร์ของข้อสอบเป็นอิสระจากกลุ่มผู้สอบที่ใช้ในการประมาณค่า กล่าวคือ ค่าพารามิเตอร์ของข้อสอบแต่ละข้อ ได้แก่ ค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าการเดา (c) เป็นค่าคงที่ไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ ถึงแม้ว่าผู้สอบต่างกลุ่มกัน ค่าพารามิเตอร์ของข้อสอบจะมีค่าคงเดิมเสมอ และในทำนองเดียวกันค่าพารามิเตอร์ความสามารถของผู้สอบเป็นอิสระจากข้อสอบที่ใช้ในการประมาณค่า จะไม่แปรเปลี่ยนตามค่าพารามิเตอร์ของข้อสอบ ไม่ว่าผู้สอบจะทำข้อสอบข้อใด หรือได้ทำข้อสอบเมื่อใดก็ตาม

ประการที่สอง การเปรียบเทียบความสามารถของผู้สอบ จะไม่ขึ้นอยู่กับคำถามในแบบสอบ ถึงแม้ว่าผู้สอบจะใช้ข้อคำถามต่างกัน ก็สามารถนำค่าความสามารถมาเปรียบเทียบกันได้ ทั้งนี้

เนื่องจากค่าความสามารถที่ประมาณค่ามาได้นั้น เป็นคะแนนโลจิส (Logit) ซึ่งอยู่ในมาตรวัดเดียวกัน

ประการที่สาม การรายงานคุณภาพของข้อสอบในรูปค่าสารสนเทศ (Information) สามารถรายงานได้ทั้งเป็นรายข้อและทั้งฉบับ ค่าสารสนเทศเป็นตัวบ่งชี้ถึงความถูกต้องแม่นยำ (Accuracy) ในการประมาณค่าความสามารถ สามารถนำมาใช้แทนค่าความเที่ยง และค่าความคลาดเคลื่อนมาตรฐานในการวัดได้

จากคุณลักษณะดังกล่าว จึงทำให้นักวัดผลนำทฤษฎีการตอบสนองข้อสอบไปประยุกต์ใช้ในการวัดผลทางการศึกษาและจิตวิทยาในหลาย ๆ เรื่องด้วยกัน เช่น การสร้างคลังข้อสอบ (Item banking) การพัฒนาแบบทดสอบ (Test development) การหาความลำเอียงของข้อสอบ (Item bias) การทดสอบอิงเกณฑ์ (Criterion – reference testing) การเทียบคะแนนการสอบ (Test score equation) และการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive testing) (กนกวรรณ รัตนธน, 2544)

1.1. หลักการของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ เป็นทฤษฎีการวัดที่อธิบายความสัมพันธ์ระหว่างคุณลักษณะภายใน หรือความสามารถที่มีอยู่ในตัวบุคคลกับพฤติกรรมการตอบสนองข้อสอบของบุคคลนั้นว่ามีโอกาสตอบข้อสอบถูกมากน้อยเพียงไร ทฤษฎีนี้มีพื้นฐานความเชื่อว่า พฤติกรรมการตอบสนองต่อข้อสอบของผู้สอบ ซึ่งเป็นสิ่งที่สังเกตได้โดยตรงว่าถูกหรือผิด จะถูกกำหนดโดยคุณลักษณะภายในหรือความสามารถที่อยู่ภายในตัวบุคคลซึ่งเป็นสิ่งที่ไม่สามารถสังเกตได้โดยตรง (ศิริชัย กาจนวาสี, 2545) ดังนั้นทฤษฎีจึงอธิบายความสัมพันธ์ระหว่างปริมาณของการกระทำหรือคะแนนของผู้สอบกับระดับความสามารถของผู้สอบ ซึ่งแสดงได้ดังสมการต่อไปนี้ (สุพัฒน์ สุกมลสันต์, 2539)

$$P = f(\theta)$$

เมื่อ P แทน ผลการทดสอบ (Performance)

θ แทน ความสามารถ (Ability หรือ trait)

f แทน ความสัมพันธ์ (Function)

จากความสัมพันธ์ในสมการดังกล่าวเป็นการแสดงความสัมพันธ์ทั่ว ๆ ไปยังไม่เฉพาะเจาะจง จึงมีผู้สนใจศึกษาและพัฒนาเพื่อให้ได้สมการที่สามารถอธิบายลักษณะเฉพาะของข้อสอบแต่ละข้อที่สามารถอธิบายความสัมพันธ์ของสองตัวแปรนี้ได้เด่นชัดขึ้น โดยอาศัยข้อตกลงเบื้องต้นบางประการและวิธีการทางคณิตศาสตร์

1.2. ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ (IRT) มีข้อตกลงเบื้องต้น ดังต่อไปนี้ (Hambleton and Swaminathan, 1985)

1. ความเป็นมิติเดียว (Unidimension) คุณลักษณะภายในหรือความสามารถของผู้สอบมีอยู่หลายประการ ซึ่งคุณลักษณะภายในหรือความสามารถแต่ละอย่างสามารถกำหนดพฤติกรรมการตอบสนองข้อสอบ ถ้าผลการตอบสนองข้อสอบหรือคะแนนของผู้สอบสามารถอธิบายได้คุณลักษณะเดียวก็ถือว่ามีความเป็นมิติเดียว โมเดลของการตอบสนองข้อสอบมีข้อตกลงว่า ข้อสอบแต่ละข้อในแบบสอบนั้นวัดความสามารถหรือคุณลักษณะเดียวกัน แต่ข้อตกลงข้อนี้ไม่เข้มงวดนัก ถ้ามีลักษณะเด่นที่จะวัดองค์ประกอบใดองค์ประกอบหนึ่งก็ถือว่าเป็นมิติเดียวเช่นกัน

วิธีการตรวจสอบว่าแบบสอบนั้นวัดในมิติเดียวหรือไม่นั้น มีวิธีการทดสอบได้หลายอย่างแต่ที่นิยมมากมี 2 วิธี คือ (Warm, 1978 อ้างถึงใน สุพัฒน์ สุกมลสันต์, 2539)

1.1. โดยการหาค่าความสัมพันธ์ระหว่างค่าน้ำหนักองค์ประกอบรายข้อ (Factor loading) ขององค์ประกอบที่หนึ่งกับค่าสหพันธ์แบบไบซีเรียล (Biserial correlation) ของข้อสอบรายข้อกับคะแนนรวม หากค่าสหสัมพันธ์มีค่ามากกว่า .80 ก็แสดงว่าแบบสอบนั้นวัดมิติเดียว

1.2. โดยการวิเคราะห์องค์ประกอบ (Factor analysis) ของข้อสอบทั้งฉบับ ถ้าค่าไอเกน (Eigen value) ค่าสูงสุดแตกต่างจากค่าอื่นอย่างชัดเจนหรือไม่ ถ้าแตกต่างก็แสดงว่าแบบสอบนั้นวัดมิติเดียว

2. ความเป็นอิสระในการตอบข้อสอบ (Local independence) หมายถึง การตอบข้อสอบข้อต่าง ๆ ในแบบสอบของผู้สอบนั้นมีความเป็นอิสระในเชิงสถิติ กล่าวคือ การตอบข้อสอบข้อหนึ่ง ๆ จะไม่มีผลกระทบต่อการตอบข้อสอบข้ออื่น ๆ ในแบบสอบ เนื้อหาของข้อสอบข้อหนึ่ง ๆ จะต้องไม่มีเงื่อนไขในการตอบข้อสอบข้ออื่น ๆ (Hambleton and Swaminathan, 1985) ได้นิยามของความเป็นมิติเดียวของข้อสอบว่าควรอยู่บนพื้นฐานของหลักการความเป็นอิสระในการตอบข้อสอบ

3. โค้งลักษณะของข้อสอบ (Item characteristic curves : ICC) เป็นฟังก์ชันทางคณิตศาสตร์ที่แสดงถึงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบนั้นได้ถูกต้องกับระดับความสามารถที่วัดได้โดยชุดของข้อสอบหรือแบบสอบ จะเห็นได้ว่าโอกาสที่ผู้สอบจะตอบข้อสอบถูก (Probability) จะขึ้นอยู่กับโค้งลักษณะของข้อสอบ (ICC) ซึ่งเป็นอิสระจากการกระจายของความสามารถของผู้สอบ นั่นคือ โอกาสที่ผู้สอบจะตอบข้อสอบถูกไม่ขึ้นอยู่กับจำนวนของผู้สอบที่มีความสามารถเหมือนกัน รูปร่างของโค้งลักษณะของข้อสอบ (ICC) ในแต่ละข้อมีคุณสมบัติไม่

แปรเปลี่ยนไปตามกลุ่มผู้สอบ ดังนั้น จึงทำให้โอกาสในการตอบข้อสอบถูกในแต่ละข้อไม่แปรเปลี่ยน

นอกจากนี้แล้ว ถ้าข้อตกลงเบื้องต้นของการทดสอบตรงตามที่ทฤษฎีกำหนดไว้ ผลที่ได้จากการคำนวณจะมีลักษณะที่ดีดังนี้

1. ค่าพารามิเตอร์ของข้อสอบ คือ ค่าความยาก (a) ค่าอำนาจจำแนก (b) และค่าการเดา (c) จะเป็นค่าที่ไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ ไม่ว่าจะนำไปสอบกับผู้ใดก็ตามลักษณะเช่นนี้เรียกว่าความเป็นอิสระจากกลุ่มตัวอย่าง (Person free test calibration)

2. เมื่อทราบลักษณะการตอบข้อสอบแต่ละข้อของผู้สอบ เราสามารถจะคำนวณหาค่าความสามารถที่แท้จริงของผู้สอบคนนั้นได้ ค่าความสามารถที่แท้จริงนี้มีความสัมพันธ์กันโดยตรงกับคะแนนจริง การคำนวณความสามารถที่แท้จริงนั้นอาจใช้ข้อสอบข้อใดก็ได้ที่วัดสิ่งเดียวกัน ลักษณะเช่นนี้เรียกว่า ความเป็นอิสระของข้อสอบ (Item free person measurement)

โค้งลักษณะของข้อสอบมีหลายรูปแบบตามความเชื่อเกี่ยวกับความสัมพันธ์ระหว่างระดับความสามารถกับโอกาสในการตอบข้อสอบได้ถูกต้องจึงทำให้มีโมเดลของโค้งลักษณะของข้อสอบขึ้นหลายโมเดล ดังจะกล่าวต่อไปนี้

1.3. โมเดลของทฤษฎีการตอบสนองข้อสอบ

โค้งลักษณะของข้อสอบแสดงความสัมพันธ์ระหว่างระดับความสามารถกับโอกาสในการตอบข้อสอบได้ถูกต้อง ทำให้เกิดโมเดลของโค้งลักษณะของข้อสอบขึ้นหลายโมเดล แต่รูปแบบที่นิยมใช้กันแพร่หลาย คือ รูปแบบโลจิสติก (Logistic model) ซึ่งนิยมใช้กันทั่วไป 3 รูปแบบ ดังนี้ (Hambleton and Swaminathan, 1985)

1.3.1. รูปแบบโลจิสติกที่มี 1 พารามิเตอร์ (One - parameter logistic model)

ราส์ชได้พัฒนาทฤษฎีการตอบสนองข้อสอบ และเสนอการวิเคราะห์ข้อสอบแบบ 1 พารามิเตอร์ในปี ค.ศ.1960 โดยฟังก์ชันของแบบ 1 พารามิเตอร์นี้สามารถอธิบายได้ด้วยพารามิเตอร์ของข้อสอบเพียงตัวเดียวคือ ค่าความยาก (b) เขียนเป็นสมการได้ดังนี้

$$P_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}} \quad ; \ i = 1, 2, 3, \ldots, n$$

เมื่อ $P_i(\theta)$ แทน ความน่าจะเป็นของผู้สอบที่มีความสามารถ (θ) จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

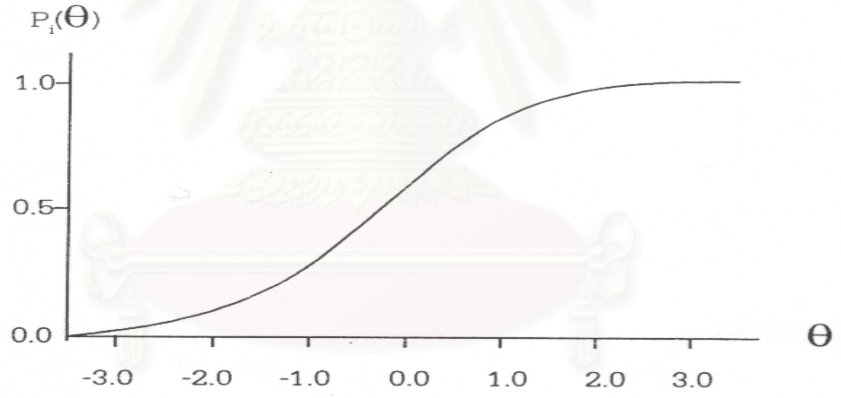
D แทน ค่าคงที่ของ Scale factor เพื่อปรับค่าของฟังก์ชั่นโลจิสติกให้ใกล้เคียงกับค่าที่ได้จากฟังก์ชันรูปตัวเอสปกติมีค่าประมาณ 1.7

b_i แทน ระดับความยากของข้อสอบข้อที่ i ที่แสดงถึงระดับความสามารถที่จุดโค้งลักษณะข้อสอบมีความชันมากที่สุด ในทางปฏิบัติค่าความยากจะมีค่าอยู่ระหว่าง -2 ถึง +2

θ แทน ระดับความสามารถที่แท้จริงที่คำนวณจากคะแนนรวม โดยการปรับให้เป็นคะแนนที่มีค่าเฉลี่ยเท่ากับ 0 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 โดยมีพิสัยการกระจายเท่ากับ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติระดับความสามารถจะมีค่าอยู่ระหว่าง -3 ถึง +3

e แทน ค่าคงที่ซึ่งมีค่าประมาณ 2.7182818

การวิเคราะห์ข้อสอบแบบ 1 พารามิเตอร์ (1 - Parameter model) นี้มีข้อตกลงว่าข้อสอบทุกข้อไม่มีโอกาสของการเดาถูก ($c_i = 0$) และค่าอำนาจจำแนกของข้อสอบ (a_i) ทุกข้อเท่ากันหมด โค้งลักษณะเฉพาะของข้อสอบแสดงในแผนภาพที่ 1



แผนภาพที่ 1 โค้งลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อสอบแบบ 1 พารามิเตอร์

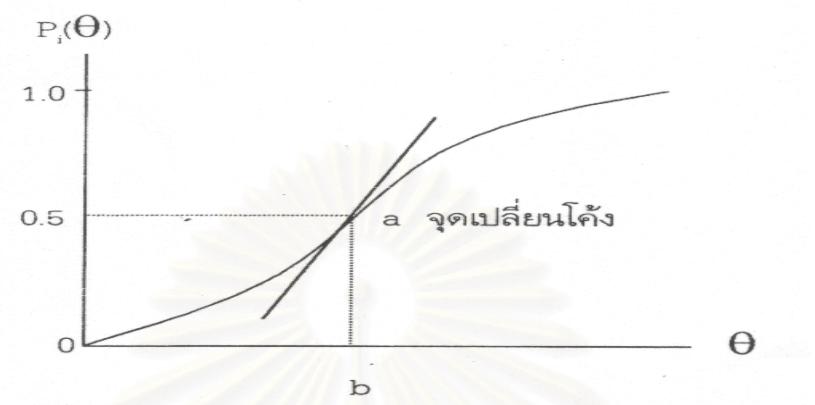
1.3.2. รูปแบบโลจิสติกที่มี 2 พารามิเตอร์ (Two – parameter logistic model)

Birnbaum (1957, 1958a, 1958b, 1968) ได้เสนอการวิเคราะห์ข้อสอบแบบโลจิสติกที่มีพารามิเตอร์ 2 ตัว คือ ค่าความยาก (b) และค่าอำนาจจำแนก (a) และสามารถเขียนเป็นสมการได้ดังนี้

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \qquad i = 1, 2, 3 \ldots, n$$

เมื่อ a_i แทน ค่าอำนาจจำแนกของข้อสอบข้อที่

การวิเคราะห์ข้อสอบแบบ 2 พารามิเตอร์ (2 - Parameter model) นี้มีข้อตกลงว่าข้อสอบทุกข้อไม่มีโอกาสของการเดาถูก คือ ค่า $c_i = 0$ โค้งลักษณะเฉพาะของข้อสอบแบบนี้มีลักษณะดังแผนภาพที่



แผนภาพที่ 2 โค้งลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อสอบแบบ 2 พารามิเตอร์

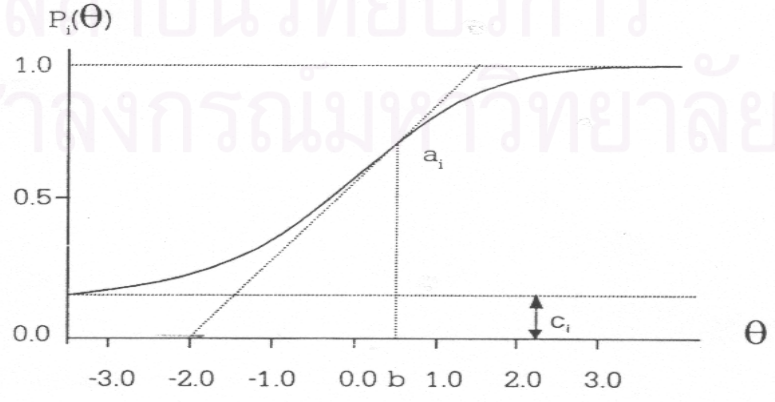
1.3.3. รูปแบบโลจิสติกที่มี 3 พารามิเตอร์ (Three – parameter logistic model)

การวิเคราะห์ข้อสอบแบบ 3 พารามิเตอร์ (3 – parameter model) นี้ดัดแปลงมาจากรูปแบบโลจิสติกที่มี 2 พารามิเตอร์ พารามิเตอร์ที่เพิ่มขึ้น คือ ค่าการเดา (c_i) และมีสมการ ดังนี้

$$P_i(\theta) = c_i + (1 - c_i)\frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad i = 1, 2, 3, \ldots, n$$

เมื่อ c_i แทน ค่าการเดา

โค้งลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อสอบแบบ 3 พารามิเตอร์ มีลักษณะดังแผนภาพที่ 3



แผนภาพที่ 3 โค้งลักษณะเฉพาะของข้อสอบของการวิเคราะห์ข้อสอบแบบ 3 พารามิเตอร์

1.4. ปัจจัยที่มีผลต่อคุณภาพของแบบสอบ

ตามแนวคิดของทฤษฎีการตอบสนองข้อสอบเชื่อว่าแบบสอบจะมีคุณภาพดีเพียงใดขึ้นอยู่กับคุณภาพของข้อสอบแต่ละข้อว่าสามารถทดสอบสิ่งที่ต้องการทดสอบได้มากน้อยเพียงใด หรือกล่าวอีกนัยหนึ่งก็คือ ข้อสอบสามารถทดสอบได้ถูกต้องแม่นยำมากน้อยเพียงใด ซึ่งขึ้นอยู่กับปัจจัยหลายประการ (ต่าย เซี่ยงฉี, 2534; สุพัฒน์ สุกมลสันต์, 2539; รังสรรค์ มณีเล็ก, 2540)

1.4.1. ความถูกต้องแม่นยำ (Accuracy)

ในการประมาณค่าความสามารถของผู้สอบ โดยปกติไม่ว่าจะเป็นการวัดหรือการประมาณค่าใด ๆ ก็ตามจะต้องมีความคลาดเคลื่อนอยู่เสมอ โดยเฉพาะอย่างยิ่งในการวัดหรือการประมาณค่าความสามารถในทางการศึกษาและจิตวิทยาจะมีความคลาดเคลื่อนค่อนข้างสูง ถ้าความคลาดเคลื่อนในการประมาณค่ามีค่าสูง ความถูกต้องแม่นยำ (Accuracy) ในการประมาณค่าก็จะต่ำ ซึ่งสามารถเขียนเป็นสมการแสดงความสัมพันธ์ระหว่างความสามารถจริง (θ) ความสามารถที่ได้จากการประมาณค่า $(\hat{\theta})$ และ ค่าความคลาดเคลื่อน (E) ได้ดังนี้ (Lord and Novick, 1968)

$$E = \hat{\theta} - \theta$$

จากสมการนี้จะเห็นว่าถ้าความคลาดเคลื่อนมีค่าน้อยหรือเข้าใกล้ศูนย์ ความสามารถที่ได้จากการประมาณค่าก็จะเข้าใกล้ความสามารถจริง นั่นคือ มีความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบสูง

1.4.2. ความคลาดเคลื่อนมาตรฐานในการประมาณค่า (Standard error of estimate)

ความคลาดเคลื่อนมาตรฐานในการประมาณค่า คือ ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในการประมาณค่า ซึ่งแตกต่างจากความคลาดเคลื่อนมาตรฐานในการวัด (Standard error of measurement: SEM) กล่าวคือ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าเป็นความคลาดเคลื่อนของค่าความสามารถของผู้สอบที่คลาดเคลื่อนไปจากความสามารถจริง ($E = \hat{\theta} - \theta$) ส่วนความคลาดเคลื่อนมาตรฐานในการวัด คือความคลาดเคลื่อน (E) ที่เกิดจากคะแนนดิบหรือคะแนนที่ได้จากการวัด (X) ที่คลาดเคลื่อนไปจากคะแนนจริง (T) นั่นคือ E = X – T ความคลาดเคลื่อนมาตรฐานในการประมาณค่ามีความสัมพันธ์กับค่าฟังก์ชันสารสนเทศของแบบสอบ (Test information function) คือ $SEE = 1/\sqrt{I(\theta)}$ ค่าฟังก์ชันสารสนเทศของแบบสอบจะเป็นตัวบ่งชี้ถึงคุณภาพของการประมาณค่าและยังสามารถใช้แทนค่าความเที่ยงหรือค่าความคลาดเคลื่อนมาตรฐานในการวัดทฤษฎีการทดสอบแบบประเพณีนิยมได้ด้วย (Hambleton and Cook, 1977 อ้างถึงใน สุพัฒน์ สุกมลสันต์, 2539)

Warm (1978 อ้างถึงใน ต่าย เซี่ยงฉี, 2534 ; รังสรรค์ มณีเล็ก, 2540) ได้แสดงความสัมพันธ์ระหว่างค่าความคลาดเลื่อนมาตรฐานในการประมาณค่าความสามารถกับค่าความเชื่อมั่นของแบบทดสอบไว้ ดังนี้

$$r_{tt} = 1 - \overline{SEE}^2$$

เมื่อ $\overline{SEE}$ แทน ค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ

Samejima Green and others. (1977, 1984 อ้างถึงใน ต่าย เซี่ยงฉี, 2534 ; รังสรรค์ มณีเล็ก, 2540) มีความเห็นว่า ไม่ควรนำค่าความเที่ยงมาใช้กับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ แต่ Divgi (1989) เห็นว่าควรใช้ค่าความเที่ยงของแบบทดสอบกับการทดสอบลักษณะนี้ ซึ่งมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าหาได้จากสูตร (อ้างถึงใน ต่าย เซี่ยงฉี, 2534 ; รังสรรค์ มณีเล็ก, 2540)

$$SEE = \sqrt{\frac{N\sum E^2 - \left(\sum E\right)^2}{N(N-1)}}$$

เมื่อ SEE แทน ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ

E แทน ค่าความคลาดเคลื่อนในการประมาณค่า

N แทน จำนวนผู้สอบ

1.4.3. ฟังก์ชันสารสนเทศของข้อสอบ (Item information function)

ฟังก์ชันสารสนเทศของข้อสอบ คือ ค่าที่แสดงถึงความถูกต้องแม่นยำในการประมาณค่าความสามารถจริง (θ_i) ของผู้สอบในการตอบข้อสอบแต่ละข้อ ซึ่งสามารถคำนวณได้จากสูตรต่อไปนี้

$$I(\theta, U_i) = \frac{P_i'(\theta)^2}{P_i(\theta)Q_i(\theta)}$$

เมื่อ $I(\theta, U_i)$ แทน ฟังก์ชันสารสนเทศของข้อสอบ

$P_i'(\theta)$ แทน ความชันของโค้งลักษณะเฉพาะของข้อสอบที่ระดับความสามารถ θ

$P_i(\theta)$ แทน ความน่าจะเป็นที่ผู้สอบที่มีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ถูก

$Q_i(\theta)$ แทน $1 - P_i(\theta)$

จากสมการข้างต้น จะเห็นได้ว่า ฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อขึ้นอยู่กับความชันของโค้งลักษณะเฉพาะของข้อสอบ ถ้าโค้งลักษณะเฉพาะของข้อสอบชันมากขึ้นในขณะที่ความแปรปรวนของการตอบข้อสอบถูกน้อยลง โค้งสารสนเทศของข้อสอบที่ระดับความสามารถนั้น ๆ จะยิ่งสูงขึ้น ความสูงของโค้งสารสนเทศของข้อสอบอยู่ที่ระดับความสามารถใดแสดงว่าสามารถจำแนกระดับความสามารถของผู้สอบได้ดี ณ ระดับความสามารถนั้น

1.4.4. ฟังก์ชันสารสนเทศของแบบสอบ (Test information function)

ฟังก์ชันสารสนเทศของแบบสอบเป็นสัดส่วนกลับกันกับกำลังสองของความยาวของช่วงความเที่ยง ซึ่งเป็นผลมาจากการประมาณค่าความสามารถของผู้สอบจากการตอบข้อสอบทั้งฉบับ ค่านี้แสดงถึงความถูกต้องแม่นยำในการประมาณค่าความสามารถจริง (θ) ของผู้สอบแบบสอบทั้งฉบับว่า มีมากน้อยเพียงใด ดังนั้นโค้งสารสนเทศของแบบสอบจึงเป็นตัวบ่งชี้ถึงความถูกต้องแม่นยำของค่าความสามารถที่ประมาณได้ ซึ่งแสดงในรูปสมการได้ดังนี้

$$I(\theta)=\sum_{i=1}^{n} I(\theta,\mu_i)=\sum \frac{P_i^{'}(\theta)^2}{P_i(\theta)Q(\theta)}$$

เมื่อ $I(\theta)$ แทน ฟังก์ชันสารสนเทศของแบบสอบ

$I(\theta,U_i)$ แทน ฟังก์ชันสารสนเทศของข้อสอบ

ถ้าเรามีกลุ่มข้อสอบที่ทราบค่าสารสนเทศของข้อสอบ เราสามารถสร้างแบบสอบให้มีโค้งสารสนเทศของแบบสอบ ณ ความสามารถระดับใดระดับหนึ่งที่เราต้องการ เช่น การสร้างแบสอบเพื่อคัดเลือกนักเรียนเข้าศึกษาต่อ ก็ต้องสร้างแบบสอบที่มีประสิทธิภาพสูงสุดที่ระดับความสามารถสูง ๆ นั่นคือ ให้มีโค้งสารสนเทศของแบบสอบสูง ณ ระดับความสามารถสูง ๆ

ตอนที่ 2 การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

2.1. ความหมายของการทดสอบแบบปรับเหมาะ

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive testing) หมายถึง การทดสอบที่ใช้แบบสอบต่างชุดกันสำหรับผู้สอบต่างกัน โดยมีการคัดเลือกข้อสอบที่มีความยากเหมาะสมกับระดับความสามารถของผู้สอบ

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ มีความแตกต่างที่สำคัญจากการทดสอบแบบประเพณีนิยม ตรงที่การทดสอบแบบประเพณีนิยมเป็นการสร้างแบบสอบขึ้นมาชุด

เดียวให้ครอบคลุมองค์ประกอบของคุณลักษณะหรือเนื้อหาที่ต้องการวัด มีค่าความยากของข้อสอบที่หลากหลายโดยไม่มีการกำหนดสัดส่วนของค่าความยากที่แน่นอน โดยทั่วไปนิยมให้มีค่าความยาก (b) อยู่ระหว่าง 0.2 - 0.8 สำหรับนำไปใช้สอบกับผู้สอบกลุ่มใดก็ได้ ไม่ว่าผู้สอบจะมีความสามารถสูง ปานกลางหรือต่ำ แต่การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบเป็นการใช้ข้อสอบจากคลังข้อสอบมาสร้างเป็นแบบสอบ แบบสอบแต่ละชุดมีการออกแบบให้เหมาะสมกับความสามารถของผู้สอบแต่ละบุคคล (ศิริชัย กาญจนวาสี, 2538) ซึ่งการทดสอบดังกล่าวนี้มีชื่อและวิธีการแตกต่างกันออกไป เช่น Adaptive testing, Branched testing, Individualized testing, Response contingent testing, Tailor testing computerized testing, Programmed testing, Computerized automate testing, Sequential item testing ชื่อแรกที่ใช้เรียกการทดสอบลักษณะนี้ คือ การทดสอบเทเลอร์ (Tailor testing) ซึ่ง William W. turnbull เป็นผู้คิดขึ้นในปี 1951 (Lord, 1980)

2.2. หลักการของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

หลักการของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ เป็นการคัดเลือกข้อสอบสำหรับแต่ละบุคคลโดยอยู่บนพื้นฐานของผลการตอบข้อสอบข้อที่ผ่านมาของผู้สอบนั้น เมื่อผู้สอบทำข้อสอบข้อแรกซึ่งได้จากคลังข้อสอบแล้ว จะมีการวิเคราะห์ระดับความสามารถหรือประเมินความสามารถของผู้สอบเบื้องต้น เพื่อคัดเลือกข้อสอบข้อต่อไปที่มีค่าความยาก และค่าอำนาจจำแนกเหมาะสมที่จะใช้วัดระดับความสามารถของผู้สอบและทำการประมาณค่าระดับความสามารถของผู้สอบใหม่ จากนั้นก็จะเลือกข้อที่เหมาะสมข้อต่อไป โดยอาศัยหลักการที่ว่าถ้าการทำข้อที่ผ่านมาถูก ข้อถัดไปจะยากขึ้น แต่ถ้าการทำข้อที่ผ่านมาผิดข้อถัดไปจะง่ายลง กระบวนการนี้จะดำเนินต่อไปเรื่อยจนสามารถประมาณค่าระดับความสามารถของผู้สอบได้อย่างเชื่อถือได้ตามเกณฑ์ที่กำหนดไว้ การทดสอบจึงจะยุติลง (ศิริชัย กาญจนวาสี, 2545)

2.3. ประเภทของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

Weiss (1974 อ้างถึงใน ต่าย เชี่ยงฉี, 2534) ได้แบ่งวิธีการคัดเลือกข้อสอบที่มีระดับความยากเหมาะสมกับความสามารถของผู้สอบเป็น 2 วิธี คือ

1. ยุทธวิธีสองขั้นตอน (Two–stage strategies)
2. ยุทธวิธีหลายขั้นตอน (Multi–stage strategies)
    - 2.1 รูปแบบแยกทางคงที่ (Fixed branching model)
        - 2.1.1 รูปแบบปิรามิด (Pyramidal model) ได้แก่
            1) รูปแบบปิรามิดขนาดขั้นคงที่ (Constant step size pyramidal model)

2) รูปแบบปิรามิดขนาดขั้นแปรผัน (Variable step size pyramidal model)
3) รูปแบบปิรามิดข้างตัด (Truncated pyramidal model)
4) รูปแบบปิรามิดแบบมีหลายข้อในแต่ละขั้น (Multi - item pyramidal model)
5) รูปแบบปิรามิดแบบให้น้ำหนักตัวเลือกเพื่อแยกทาง (Differential response option branching pyramidal model)

2.1.2 รูปแบบเฟล็กซิเลเวล (Flexilevel model)
2.1.3 รูปแบบปรับระดับชั้น (Stradaptive model)

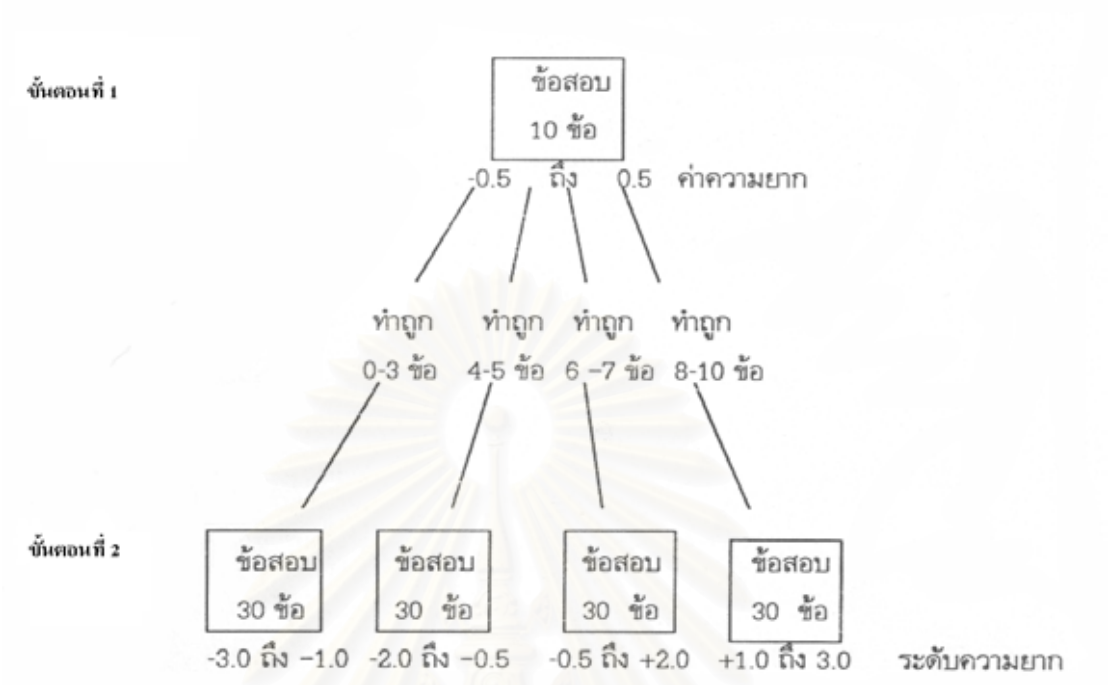
2.2 รูปแบบแยกทางแปรผัน (Variable branching model)
2.2.1 ยุทธวิธีแบบเบส์ (Bayesian strategies)
2.2.2 ยุทธวิธีความเป็นไปได้สูงสุด (Maximum likelihood estimation strategies)

รายละเอียดของการทดสอบแบบต่าง ๆ พอสรุปได้ดังนี้

ยุทธวิธีสองขั้นตอน

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้ยุทธวิธีสองขั้นตอน โดยมีการดำเนินการสอบเป็น 2 ขั้นตอน ขั้นตอนแรกทำการทดสอบด้วยแบบสอบกำหนดทิศทาง (Routing system) ซึ่งเป็นแบบสอบที่มีจำนวนข้อสอบประมาณ 10 ข้อ ประกอบด้วยข้อสอบที่มีค่าเฉลี่ยความยากอยู่ระดับปานกลาง ผู้สอบทุกคนต้องตอบข้อสอบเหมือนกัน เพื่อกำหนดทิศทางของผู้สอบแต่ละคนว่าจะต้องทำการทดสอบในขั้นตอนที่สองด้วยแบบสอบที่มีความยากระดับใดหลังจากทราบทิศทางหรือระดับความสามารถเบื้องต้นของผู้สอบแล้วก็ทำการทดสอบในขั้นตอนที่สอง ด้วยแบบสอบวัดผล (Measurement test) ซึ่งจะแบ่งเป็นแบบสอบชุดย่อย ๆ หลายชุดตามระดับความยากโดยปกติจะมีชุดละประมาณ 20 – 30 ข้อ ผู้สอบแต่ละคนจะได้รับแบบสอบชุดที่สองไม่เหมือนกัน ผู้สอบที่ได้รับการประเมินจากการทดสอบในขั้นตอนแรกว่ามีความสามารถสูงก็จะได้รับแบบสอบชุดที่สองที่ยาก ผู้สอบที่ความสามารถปานกลางจะได้รับแบบสอบชุดที่สองที่มีระดับความยากปานกลางและผู้สอบที่มีความสามารถต่ำก็จะได้รับแบบสอบชุดที่สองที่ง่าย

โครงสร้างของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้ยุทธวิธีสองขั้นตอน แสดงดังแผนภาพที่ 4



แผนภาพที่ 4 โครงสร้างของการทดสอบโดยใช้ยุทธวิธี 2 ขั้นตอน

ยุทธวิธีหลายขั้นตอน

การทดสอบแบบหลายขั้นตอน เป็นการทดสอบมากกว่า 2 ขั้นตอน โดยมีการจัดโครงสร้างของขั้นตอนการทดสอบ มีการจัดเรียงข้อสอบตามความยาก ในแต่ละขั้นตอนของการทดสอบมีวิธีดำเนินการแตกต่างกัน หลักการโดยทั่วไปเริ่มต้นจากข้อสอบที่มีความยากปานกลาง ถ้าตอบถูกข้อต่อไปจะยากขึ้น แต่ถ้าตอบผิดข้อต่อไปจะง่ายลง การทดสอบจะดำเนินเช่นนี้ไปเรื่อยและจะยุติตามเกณฑ์ที่กำหนดไว้ การทดสอบแบบหลายขั้นตอนนี้พอจะจำแนกเป็นแบบต่าง ๆ กันได้ดังนี้ (ศิริชัย กาญจนวาสี, 2538)

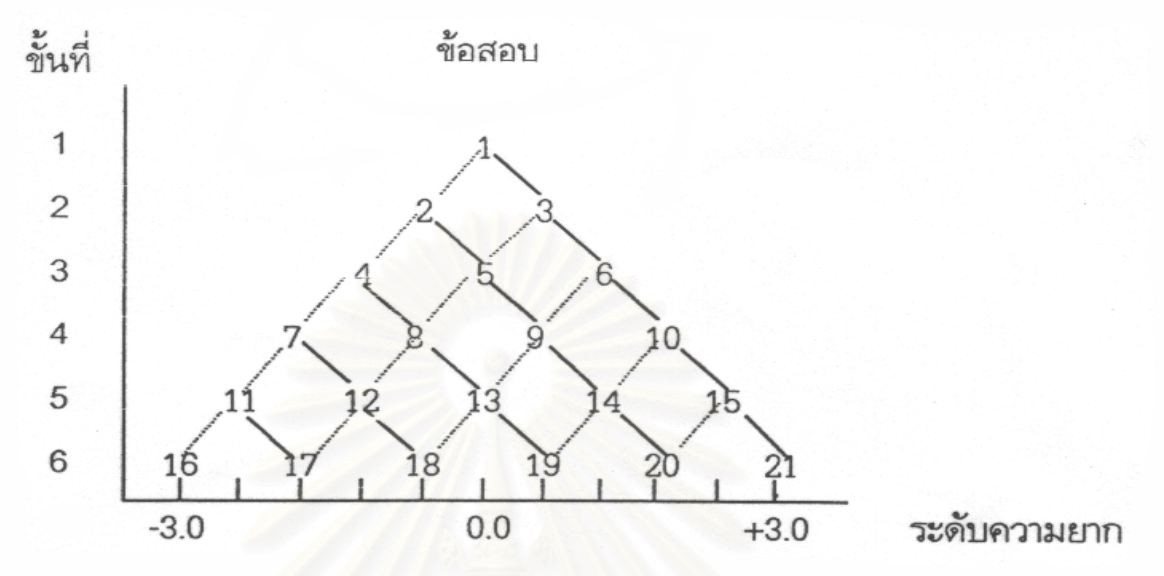
1. แบบทางแยกคงที่

เป็นการทดสอบที่กำหนดทางแยกของการเลือกข้อสอบไว้คงที่ตายตัว มีรูปแบบของการดำเนินการทดสอบแบบทางแยกคงที่นี้หลายรูปแบบดังนี้

1.1 รูปแบบปิรามิด จำแนกเป็นหลายแบบ ได้แก่

1.1.1 รูปแบบปิรามิดขนาดขั้นคงที่ เป็นการจัดโครงสร้างข้อสอบตามลำดับความยากเป็นรูปสามเหลี่ยมหรือเรียกว่าทางแยกขึ้นลง (up - and – down

branching) โดยมีจำนวนข้อทดสอบในแต่ละขั้นเท่ากับลำดับที่ของขั้น กล่าวคือ ขั้นที่ 1 จะมีข้อสอบ 1 ข้อ ขั้นที่ 2 จะมีข้อสอบ 2 ข้อ ในขั้นที่ 6 จะมีข้อสอบ 6 ข้อ ดังแผนภาพที่ 5

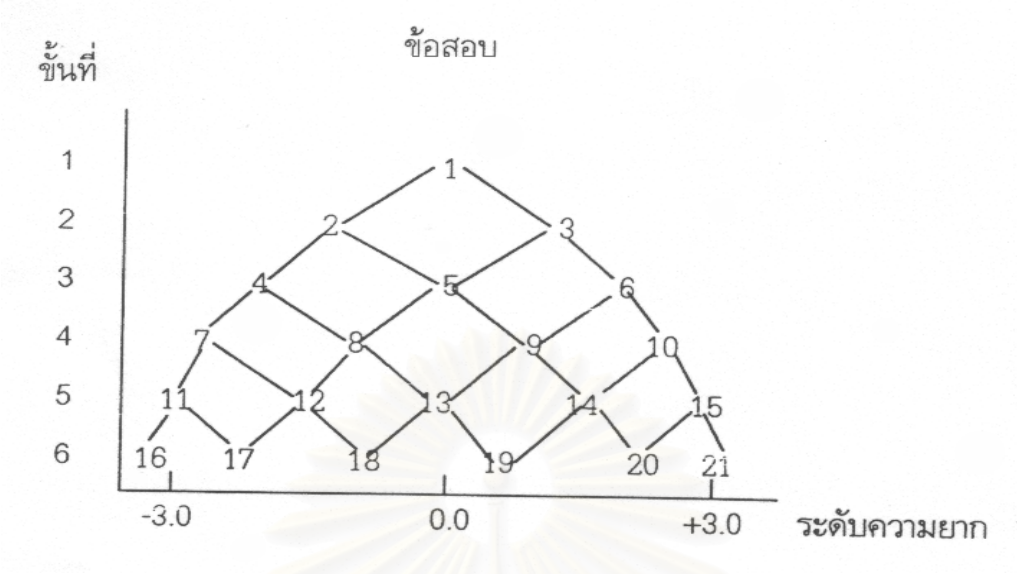


แผนภาพที่ 5 โครงสร้างของการทดสอบรูปแบบปิรามิดขนาดขั้นคงที่

ในการตอบแบบสอบ ผู้สอบจะต้องตอบข้อสอบขั้นละหนึ่งข้อ โดยเริ่มตอบข้อที่อยู่บนยอดสามเหลี่ยมซึ่งมีความยากปานกลางก่อน ถ้าตอบถูกขั้นต่อไปจะแยกไปตอบข้อที่ยากขึ้น แต่ถ้าตอบผิดขั้นต่อไปจะไปตอบข้อที่ง่ายกว่า กระบวนการทดสอบจะเป็นเช่นนี้จนถึงขั้นสุดท้าย

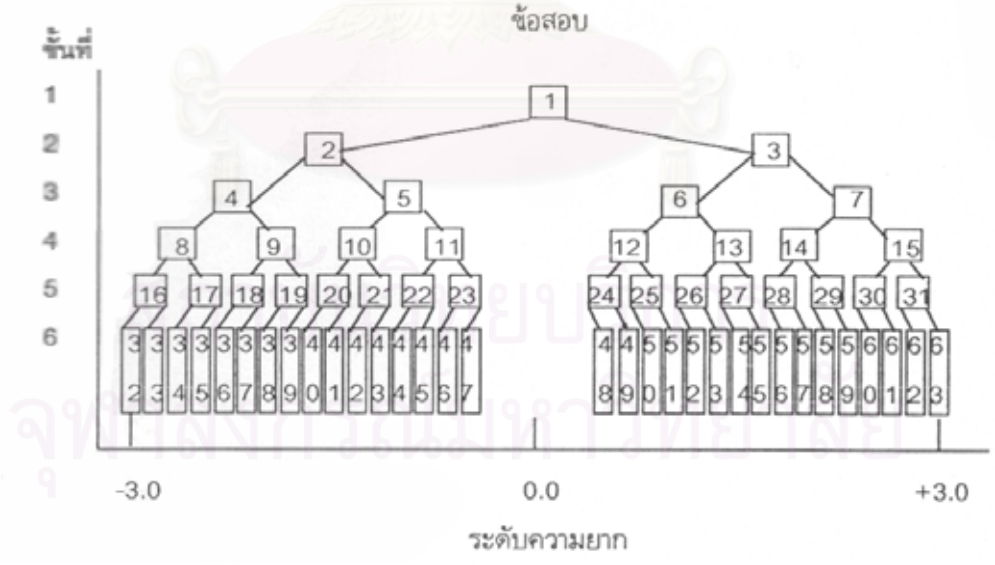
1.1.2 รูปแบบปิรามิดขนาดขั้นแปรผัน เป็นการจัดโครงสร้างข้อสอบตามลำดับความยาก โดยข้อสอบแต่ละข้อมีค่าความยากเพิ่มขึ้นหรือลดลงไม่คงที่ หรือเรียกว่า วิธี H – L โดยมีรูปแบบการปรับค่าความยากของข้อสอบเป็น 2 แบบ ดังนี้

1) วิธี H – L กำหนดช่วงห่างระดับความยากของข้อที่อยู่ติดกันภายในขั้นให้มีค่าไม่เท่ากัน โดยให้ข้อที่อยู่กลาง ๆ มีช่วงห่างของระดับความยากมาก แล้วค่อย ๆ ลดลงในข้อที่ค่อนไปทางง่ายหรือยาก ดังแผนภาพที่ 6



แผนภาพที่ 6 โครงสร้างของการทดสอบรูปแบบปิรามิดขนาดขั้นแปรผัน

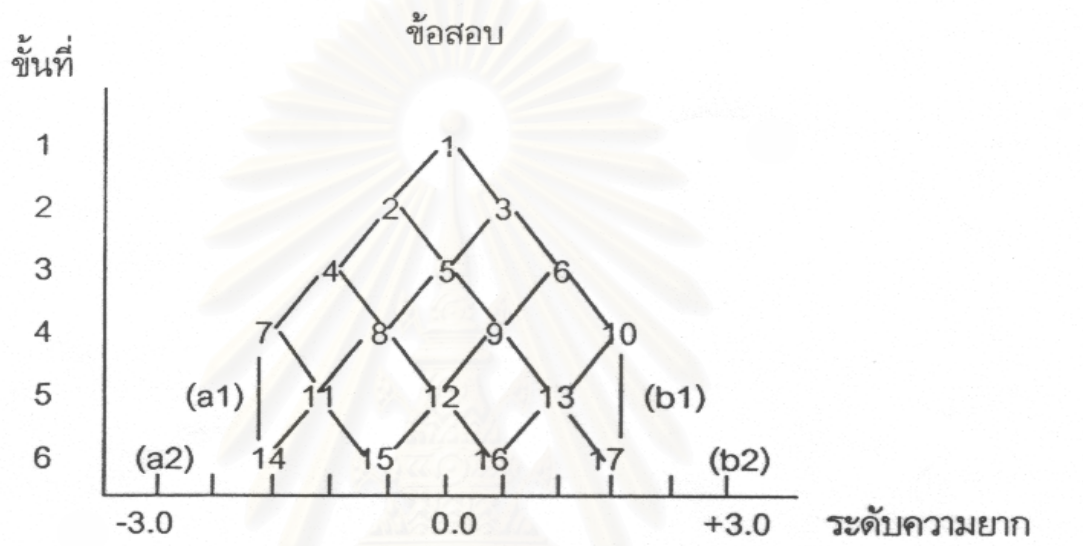
2) วิธีของรอบินและมอนโร (Robbins - Monro procedure) นี้มีการปรับช่วงการเพิ่มขึ้นหรือลดลงของค่าความยากให้แคบลงอย่างต่อเนื่องตลอดการทดสอบ ซึ่งเป็นการทำให้การเปลี่ยนแปลงค่าความยากของข้อสอบหดตัวลงระหว่างการทดสอบ ดังแผนภาพที่ 7



แผนภาพที่ 7 โครงสร้างของการทดสอบรูปปิรามิดชนิดรอบิน - มอนโร

การจัดชุดข้อสอบสำหรับการทดสอบลักษณะนี้ จึงต้องเตรียมข้อสอบที่มีช่วงห่างของค่าความยากไม่เท่ากันในแต่ละขั้นของการทดสอบ โดยในขั้นตอนแรก ๆ ต้องใช้ข้อสอบที่มีความยากปานกลาง ข้อสอบในขั้นตอนนี้จะมีช่วงห่างของค่าความยากค่อนข้างกว้าง ขั้นตอนต่อไปข้อสอบที่ง่ายลงหรือยากขึ้น จะต้องมีช่วงห่างของค่าความยากที่แคบลง ตามลำดับ

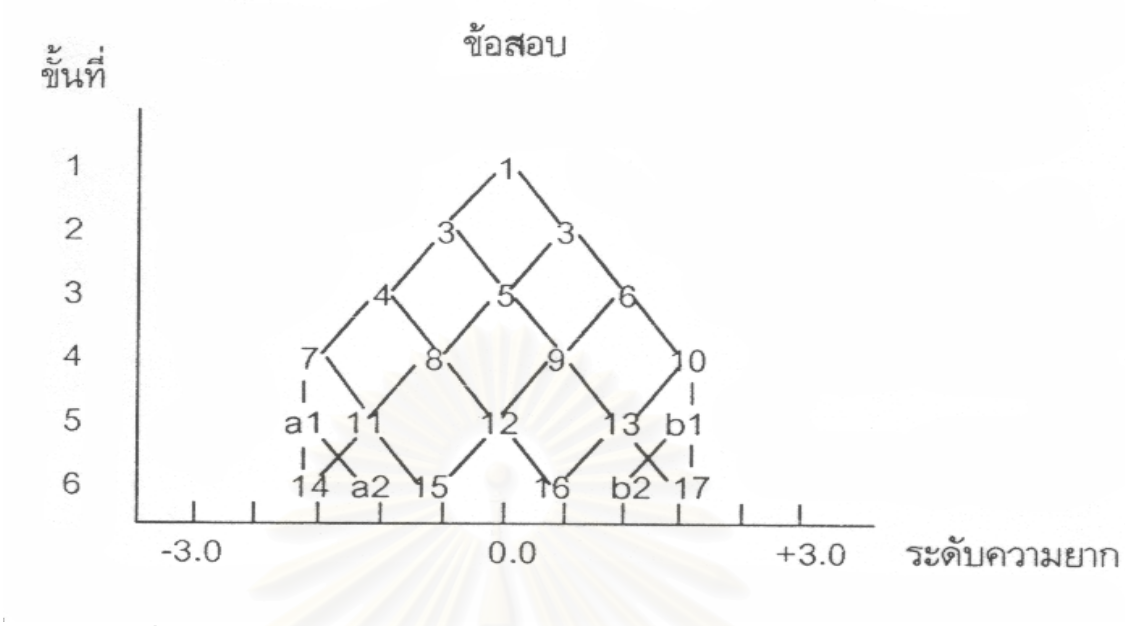
1.1.3 รูปแบบปิรามิดข้างตัด เป็นแบบที่ลดจำนวนข้อสอบลงจากแบบปิรามิดขนาดขั้นคงที่ด้วยวิธีการสกัดการสะท้อนกลับและวิธีรักษาการสะท้อนกลับ ดังแผนภาพที่ 8 และ 9



แผนภาพที่ 8 โครงสร้างของการทดสอบรูปแบบปิรามิดข้างตัดตามวิธีสกัดการสะท้อนกลับ

จากแผนภาพที่ 8 การตอบข้อสอบในขั้นที่ 1 ถึงขั้นที่ 3 มีวิธีตอบเช่นเดียวกับแบบปิรามิดขนาดขั้นคงที่ จนถึงจุดสกัดในขั้นที่ 4 ถ้าตอบข้อสอบข้อที่ 7 ถูก ก็จะแยกไปตอบข้อที่ 11 แต่ถ้าตอบข้อที่ 7 ผิด ก็จะไปตอบข้อที่ 14 และในทำนองเดียวกัน ถ้าตอบข้อที่ 10 ผิดก็จะแยกไปตอบข้อที่ 13 แต่ถ้าตอบข้อที่ 10 ถูกก็จะไปตอบข้อที่ 17

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



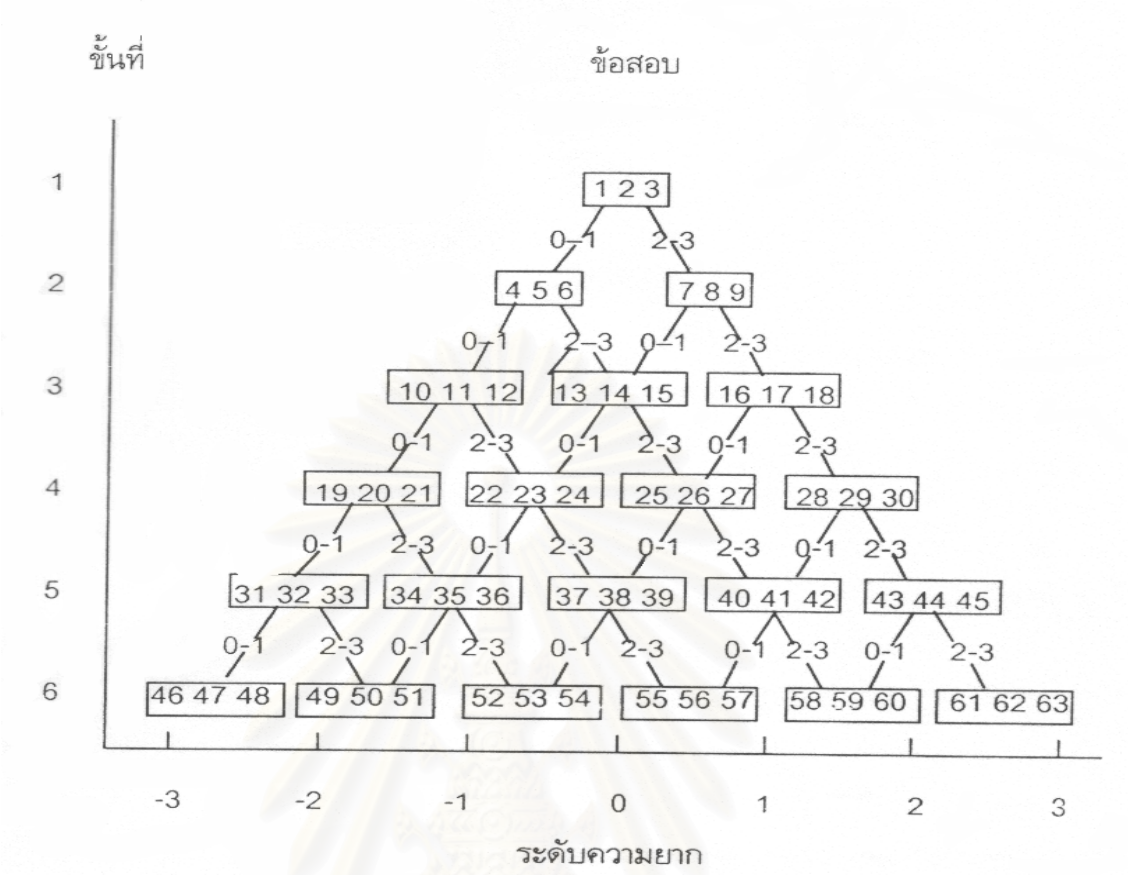
แผนภาพที่ 9 โครงสร้างของการทดสอบรูปแบบปิรามิดข้างตัดตามวิธีรักษาการสะท้อนกลับ

จากแผนภาพที่ 9 การตอบข้อสอบในขั้นที่ 1 ถึงขั้นที่ 3 มีวิธีตอบเช่นเดียวกับแบบปิรามิดชนิดขนาดขั้นคงที่ จนถึงจุดสกัดในขั้นที่ 4 ถ้าตอบข้อสอบข้อที่ 7 ถูกก็จะแยกไปตอบข้อที่ 11 แต่ถ้าตอบข้อที่ 7 ผิด ก็จะไปตอบข้อที่ a1 ซึ่งมีระดับความยากเท่ากับข้อที่ 7 และถ้าตอบข้อที่ a1 ผิดอีกก็จะไปตอบข้อที่ 14 แต่ถ้าตอบข้อที่ a1 ถูก ก็จะแยกไปตอบข้อที่ a2 ซึ่งมีระดับความยากเท่ากับข้อที่ 11

1.1.4 รูปแบบปิรามิดแบบมีหลายข้อในแต่ละขั้น เป็นแบบที่ให้มีจำนวนข้อในแต่ละขั้นมากกว่าหนึ่งข้อ ทั้งนี้เพื่อให้การตอบเพื่อแยกทางได้ถูกต้องมากยิ่งขึ้น ดังแผนภาพที่ 10

จากแผนภาพที่ 10 ในขั้นที่ 1 ถ้าผู้สอบตอบข้อสอบถูก 0 ถึง 1 ข้อ จะแยกไปตอบข้อสอบข้อที่ 4, 5 และ 6 ถ้าตอบข้อสอบถูก 2 ถึง 3 ข้อ ก็จะแยกไปตอบข้อสอบข้อที่ 7, 8 และ 9 สำหรับในขั้นอื่น ๆ ก็ในทำนองเดียวกัน

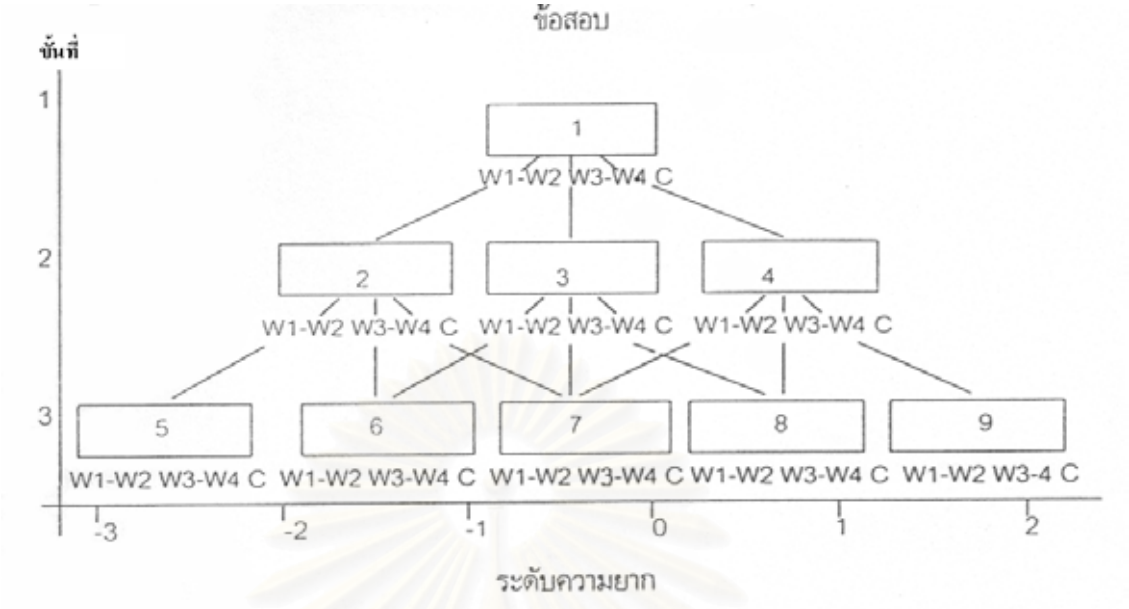
ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



แผนภาพที่ 10 โครงสร้างของการทดสอบรูปแบบปิรามิดมีหลายข้อในแต่ละขั้น

1.1.5 รูปแบบปิรามิดชนิดให้น้ำหนักตัวเลือกของข้อสอบเพื่อแยกทาง เป็นแบบที่คล้ายกับแบบที่ใช้ข้อสอบหลาย ๆ ข้อในแต่ละขั้น คือ แทนที่จะใช้ข้อสอบหลาย ๆ ข้อในแต่ละขั้นก็ใช้ตัวเลือกของข้อสอบแทนรูปแบบนี้เหมาะสำหรับข้อสอบเลือกตอบที่สามารถกำหนดน้ำหนักคะแนนของตัวเลือกได้ ดังแผนภาพที่ 11

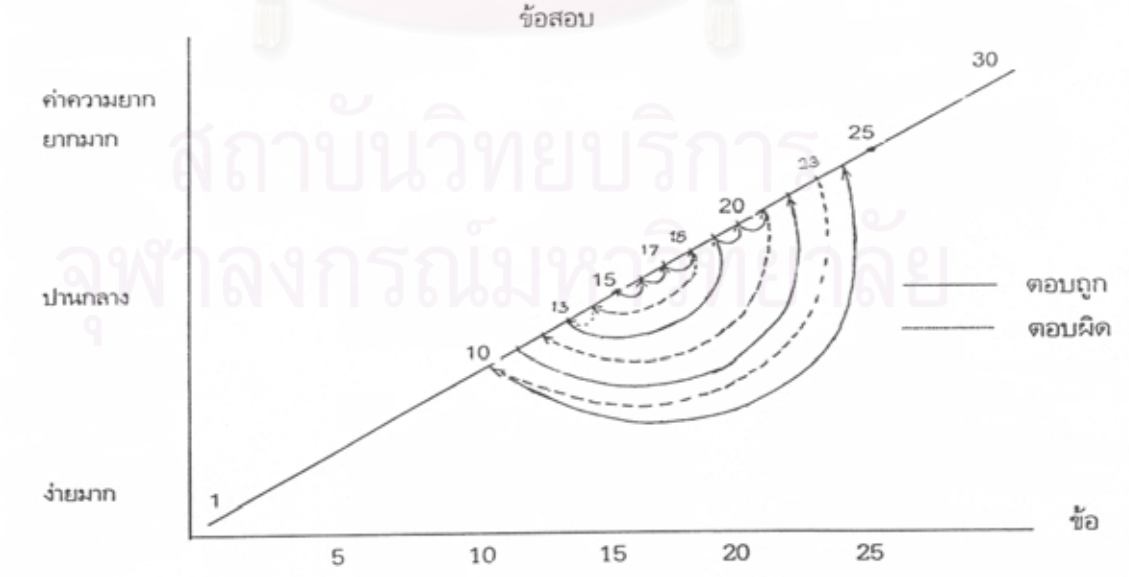
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



แผนภาพที่ 11 โครงสร้างการทดสอบรูปแบบปิรามิดชนิดให้น้ำหนักตัวเลือกข้อสอบเพื่อแยกทาง

1.2 รูปแบบเฟล็กซิเลเวล

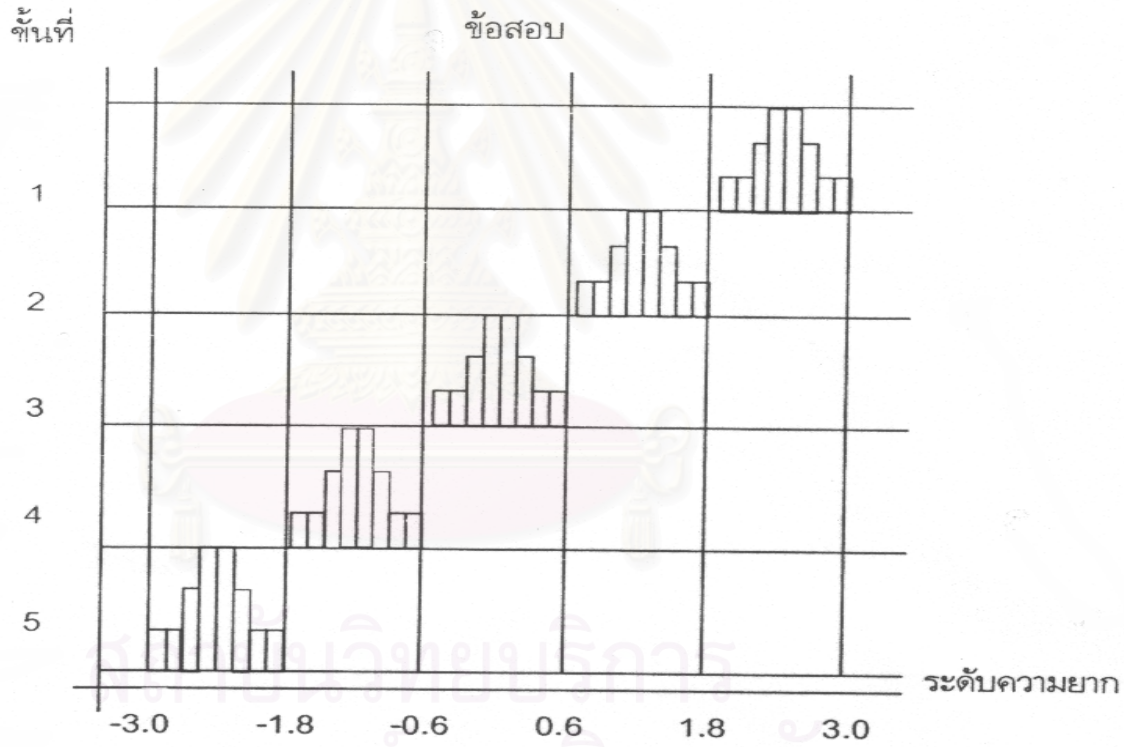
การทดสอบแบบนี้เป็นแบบที่ลอร์ดคิดขึ้นมาในครั้งแรกเพื่อใช้กับการทดสอบที่ใช้กระดาษกับดินสอ (Paper and pencil test) ซึ่งประกอบด้วยชุดของข้อสอบจำนวนหนึ่ง ข้อสอบแต่ละข้อมีช่วงห่างของค่าความยากเท่ากัน โดยทำการเรียงจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุด การทดสอบเริ่มต้นด้วยการให้ผู้ตอบทำข้อสอบข้อที่มีความยากปานกลาง ถ้าตอบถูกก็จะไปตอบข้อที่ยากขึ้น แต่ถ้าตอบผิดข้อต่อไปก็จะง่ายลง ดังแผนภาพที่ 12



แผนภาพที่ 12 โครงสร้างของการทดสอบรูปแบบเฟล็กซิเลเวล

1.3 รูปแบบปรับระดับชั้น แบบสอบปรับระดับแบบแบ่งชั้นพัฒนามาจากแบบสอบปรับระดับแบบเฟล็กซิเลเวล และมีลักษณะคล้ายกับลักษณะการใช้แบบสอบวัดเชาว์ปัญญาของ Binet คือ มีการดำเนินการทดสอบโดยใช้ข้อสอบจากคลังข้อสอบที่มีการจัดแบ่งความยากของข้อสอบเป็นระดับชั้น (Stratified) แต่ละระดับชั้นประกอบด้วยชุดของข้อสอบที่มีค่าความยากอยู่ในช่วงที่กำหนดเดียวกัน เช่น ค่าความยากอยู่ระหว่าง .20 ถึง .30, .31 ถึง .40 เป็นต้น

การใช้แบบสอบปรับระดับชั้นนี้เริ่มต้นที่ข้อสอบระดับความยากใดก็ได้ ถ้าทำข้อนั้นถูก ข้อต่อไปจะเป็นข้อที่มีอำนาจจำแนกสูงสุดที่มีความยากถัดขึ้น แต่ถ้าทำข้อนั้นผิด ข้อต่อไปจะเป็นข้อที่มีอำนาจจำแนกสูงสุดที่มีความง่ายถัดลงไป การทดสอบจะดำเนินเช่นนี้ไปเรื่อยจนถึงความยากระดับเพดาน (Ceiling stratum) ซึ่งผู้สอบไม่สามารถทำข้อสอบชุดนั้นได้เลยสักข้อ หรือทำคะแนนได้ไม่เกินคะแนนที่ได้จากการเดา (Chance score) ดังแผนภาพที่ 13



แผนภาพที่ 13 โครงสร้างของการทดสอบรูปแบบปรับระดับชั้น

2. รูปแบบแยกทางแปรผัน

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้กลวิธีหลายขั้นตอนตามแบบทางแยกแปรผัน หมายถึง รูปแบบการตอบข้อสอบหลายขั้นตอนที่ไม่ได้กำหนดข้อสอบและเส้นทางในการตอบข้อสอบไว้ล่วงหน้าว่าถ้าผู้สอบตอบถูกจะต้องไปทำข้อสอบข้อใดหรือถ้าตอบผิดจะต้องไปทำข้อใดต่อไป ในแบบทางแยกแปรผันนี้จะไม่มีขนาดของขั้น แต่จะ

ดำเนินการทดสอบจากกลุ่มข้อสอบที่คำนวณระดับความยากและค่าอำนาจจำแนกของข้อสอบไว้แล้ว กฎการคัดเลือกข้อสอบถัดไปเพื่อให้เหมาะสมกับความสามารถของผู้สอบนั้นจะใช้แบบทางคณิตศาสตร์เข้ามาช่วยได้แก่ การใช้กลวิธีของเบส์และกลวิธีความเป็นไปได้สูงสุด ซึ่งมีรายละเอียดดังนี้

2.1 ยุทธวิธีของเบส์

ยุทธวิธีของเบส์นี้ประยุกต์ทฤษฎีของเบส์ในกระบวนการตอบข้อสอบและอาศัยเครื่องคอมพิวเตอร์ในการทดสอบ จึงมักเรียกว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้เครื่องคอมพิวเตอร์ช่วยมีขั้นตอนดังนี้

1) ในแต่ละขั้นของการทดสอบจะมีการประมาณค่าความสามารถของผู้สอบและความคลาดเคลื่อนมาตรฐานในการประมาณค่าไว้จากข้อมูลต่าง ๆ ที่มีอยู่เกี่ยวกับผู้สอบและข้อสอบ

2) คัดเลือกข้อสอบจากกลุ่มข้อสอบที่ได้คำนวณค่าพารามิเตอร์ของข้อสอบไว้แล้ว ข้อสอบในกลุ่มข้อสอบทุกข้อที่ยังไม่ได้นำมาให้ผู้สอบคนนั้น ๆ สอบจะถือว่าเป็นข้อสอบที่มีโอกาสนำมาใช้สอบได้ กระบวนการนี้แสดงให้เห็นว่าข้อสอบข้อใด ๆ ในกลุ่มข้อสอบที่นำมาใช้ในการทดสอบกับผู้สอบคนใด ๆ ก็ตามจะเป็นข้อสอบที่มีระดับความยากใกล้เคียงกับระดับความสามารถของผู้สอบมากที่สุด หลังจากดำเนินการสอบโดยใช้ข้อสอบที่คัดเลือกไว้ก็จะประมาณความสามารถของผู้สอบไว้ก่อน แล้วการตอบถูกหรือการตอบผิดจะนำมารวมกันเพื่อคำนวณโดยใช้ทฤษฎีบทของเบส์หาค่าความสามารถในภายหลัง การประมาณค่าครั้งหลังนี้เป็นการประมาณค่าที่ปรับจากค่าที่ได้คำนวณไว้แต่เดิม กระบวนการนี้จะสิ้นสุดลงเมื่อความคลาดเคลื่อนในการประมาณค่าความสามารถมีค่าน้อยกว่าที่กำหนดไว้ล่วงหน้า

2.2 ยุทธวิธีการประมาณค่าความเป็นไปได้สูงสุด

วิธีนี้ได้พัฒนาขึ้นโดยใช้วิธีการประมาณค่าความเป็นไปได้สูงสุดตามทฤษฎีการตอบสนองข้อสอบ ซึ่งมีวิธีการประมาณค่าความสามารถของผู้สอบที่คล้ายคลึงกับกระบวนการตามทฤษฎีบทของเบส์ แม้ว่าเหตุผลทางคณิตศาสตร์ต่างกัน หลังจากที่ผู้สอบตอบข้อสอบเพียง 1 ข้อก็จะนำผลการทดสอบไปแก้สมการความเป็นไปได้สูงสุด และจะได้ค่าประมาณค่าความสามารถและความคลาดเคลื่อนมาตรฐาน ข้อสอบข้อถัดไปที่เลือกมาใช้ทดสอบจะเป็นข้อสอบจากกลุ่มข้อสอบและเป็นข้อสอบที่มีระดับความยากเหมาะสมกับความสามารถของผู้สอบมากที่สุด เมื่อผู้สอบตอบข้อสอบข้อนั้นแล้วก็จะมีการประเมินค่าความสามารถของผู้สอบทันทีจากข้อมูลการตอบทั้งหมดของผู้สอบ ซึ่งรวมถึงข้อสอบข้อสุดท้ายที่ผู้สอบตอบ จากนั้นก็ประมาณค่าความสามารถและความคลาดเคลื่อนมาตรฐานใหม่โดยใช้สมการความเป็นไปได้สูงสุด จนกว่าการทดสอบจะสิ้นสุดตามเกณฑ์ที่กำหนดไว้

2.4. องค์ประกอบของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Weiss and Kingsbury, 1984 อ้างถึงในรังสรรค์ มณีเล็ก, 2540)

1. โมเดลการตอบสนองข้อสอบ (Item response model) ภายใต้บริบทของทฤษฎีการตอบสนองข้อสอบแล้ว การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ หรือการทดสอบแบบซี เอ ที (CAT) นั้นจะใช้โมเดลโลจิสติก (Logistic model) 1, 2 หรือ 3 พารามิเตอร์ หรือใช้โมเดลโอไจฟ์ (Ogive model) ก็ได้ การเลือกโมเดลมาใช้ในการวิเคราะห์ข้อมูลนั้นขึ้นอยู่กับลักษณะข้อมูลว่ามีลักษณะอย่างไร มีความเหมาะสมกับโมเดลใด

2. คลังข้อสอบ (Item pool) ประกอบด้วยชุดของข้อสอบที่มุ่งวัดลักษณะหรือความสามารถที่ต้องการ ข้อสอบจะต้องครอบคลุมช่วงของระดับความสามารถต่าง ๆ กัน มีจำนวนข้อสอบที่มากพอ โดยข้อสอบแต่ละข้อควรมีอำนาจจำแนกสูง มีค่าความยากที่ครอบคลุมช่วงระดับต่าง ๆ อย่างเหมาะสม และควรมีจำนวนข้อสอบเท่า ๆ กันในแต่ละระดับค่าความยาก

3. ความยากง่ายของข้อสอบข้อแรก (Entry level) ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ข้อสอบข้อแรกจะมีระดับความยากง่ายแตกต่างกันสำหรับผู้สอบแต่ละคน การเริ่มทำข้อสอบที่คลาดเคลื่อนไปจากความสามารถจริงไม่มีผลกระทบต่อการประมาณค่าความสามารถ แต่ถ้าหากมีการเลือกข้อสอบข้อแรกให้กับผู้สอบอย่างเหมาะสมแล้วก็จะช่วยประหยัดเวลาในการทดสอบได้มาก เพราะลดจำนวนข้อสอบลง

4. เกณฑ์การคัดเลือกข้อสอบ (Item selection rule) มีวิธีการเลือกข้อสอบจากคลังข้อสอบที่มีประสิทธิภาพ 2 วิธีการ คือ วิธีการเลือกข้อสอบที่มีค่าสารสนเทศสูงสุด (Maximum information item selection) และวิธีการเลือกข้อสอบแบบเบส์ (Bayesian item selection) ซึ่งวิธีแรกจะให้ค่าสารสนเทศสูงสุด ส่วนวิธีหลังจะให้ค่าความแปรปรวนที่คาดหวังของการประมาณค่าความสามารถต่ำสุด ทั้ง 2 วิธีจะคล้ายคลึงกันเนื่องจากว่าค่าสารสนเทศสูงสุดและความแปรปรวนที่คาดหวังของการประมาณค่ามีความสัมพันธ์กัน

5. วิธีการให้คะแนน (Scoring method) ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ มีวิธีการให้คะแนนหลายวิธีด้วยกัน คือ (Weiss, 1974 อ้างถึงใน ต่าย เฉียงซี่, 2534)

5.1 การให้คะแนนตามระดับความยากของข้อที่ยากที่สุดที่ตอบถูก เป็นการให้คะแนนโดยดูจากข้อสอบที่มีระดับความยากสูงที่สุดที่ผู้สอบสามารถตอบได้ถูกต้อง

5.2 การให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ทำถูก เป็นการให้คะแนนโดยการนำระดับความยากของข้อสอบทุกข้อที่ผู้สอบตอบถูกมาหาค่าเฉลี่ย

5.3 การให้คะแนนตามค่าเฉลี่ยความยากของทุกข้อที่ทำ เป็นการนำระดับความยากของข้อสอบทุกข้อที่ผู้สอบได้ตอบโดยไม่คำนึงว่าจะตอบถูกหรือไม่มาหาค่าเฉลี่ย

5.4 การให้คะแนนตามระดับความยากของข้อสุดท้ายที่ตอบ ไม่ว่าขั้นสุดท้ายจะตอบถูกหรือไม่

5.5 การให้คะแนนตามระดับความยากของข้อสอบที่ต่อจากข้อสุดท้ายที่ตอบ ($(N+1)^{th}\ item$) กล่าวคือ ถ้าขั้นสุดท้ายที่ผู้สอบสามารถตอบถูก คะแนนในขั้นต่อไปก็จะเพิ่มขึ้น แต่ถ้าขั้นสุดท้ายตอบผิดขั้นต่อไปคะแนนก็จะลดลง

5.6 การให้คะแนนตามวิธีของเบส์ มีหลายวิธี แต่วิธีที่นิยมใช้กันมาก คือ วิธีของเบส์ที่ปรับปรุงให้ทันสมัย (Bayesian updating) ซึ่งสามารถประมาณค่าความสามารถของผู้สอบได้ค่อนข้างคงที่ ซึ่งในการวิจัยนี้ผู้วิจัยเลือกวิธีการประมาณค่าความสามารถของผู้สอบ โดยใช้วิธีของเบส์ (Bayesian) โดยเลือกใช้วิธีของ โอเวน (Owen) ซึ่งมีวิธีการดังนี้ (Owen, 1975) โดยมีสูตรดังนี้ (อ้างถึงในสุพัฒน์ สุกมลสันต์, 2539; รังสรรค์ มณีเล็ก, 2540)

2.1 ในกรณีที่ตอบถูก

$$\theta_{m+1} = \theta_m + (1-c)\left(\frac{\sigma_m^2}{\sqrt{\frac{1}{a^2}+\sigma_m^2}}\right)\left(\frac{\Theta(D)}{c+(1-c)A(-D)}\right)$$

$$\sigma_{m+1}^2 = \sigma_m^2\left(1-\left(\frac{1-c}{1+\frac{1}{a^2\sigma_m^2}}\right)\left(\frac{\Theta(D)}{B}\right)\left(\frac{(1-c)\Theta(D)}{B-D}\right)\right)$$

เมื่อ $$B = c(1-c)\times A(-D)$$

2.2 ในกรณีที่ตอบผิด

$$\theta_{m+1} = \theta_m - \left(\frac{\sigma_m^2}{\sqrt{\frac{1}{a^2}+\sigma_m^2}}\right)\left(\frac{\Theta(D)}{A(D)}\right)$$

$$\sigma_{m+1}^2 = \sigma_m^2\left(1-\left(\frac{\Theta(D)}{1+\frac{1}{a^2\sigma_m^2}}\right)\left(\frac{\Theta(D)}{A(D)}+D\right)\div A(D)\right)$$

เมื่อ θ_m แทน ความสามารถของผู้สอบที่ประมาณค่าได้ก่อนตอบข้อสอบข้อที่ m+1 ซึ่งตามปกติแล้ว ถ้าไม่ทราบค่าความสามารถเบื้องต้นของผู้สอบ ก็มักกำหนดให้เท่ากับ 0

σ_m^2 แทน ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบก่อนตอบข้อที่ m+1 ถ้าไม่ทราบค่าความแปรปรวนดังกล่าวมาก่อน ก็มักกำหนดให้เท่ากับ 1

θ_{m+1} แทน ค่าความสามารถของผู้สอบโดยประมาณหลังจากที่ตอบข้อที่ m+1

σ_{m+1}^2 แทน ค่าความแปรปรวนในการประมาณค่าความสามารถของผู้สอบเมื่อตอบข้อสอบข้อที่ m+1

a_i แทน พารามิเตอร์ค่าอำนาจจำแนกข้อสอบข้อที่ m+1

b_i แทน พารามิเตอร์ค่าความยากของข้อสอบข้อที่ m+1

c_i แทน พารามิเตอร์ระดับโอกาสการเดาข้อสอบได้ถูกข้อที่ m+1

D แทน จุดบนแกน X

$\Theta(D)$ แทน ค่าออร์ดิเนท (ordinate) ของโค้งปกติที่จุด D

$A(D)$ แทน พื้นที่ใต้โค้งปกติจากค่า D จนถึงจุด D

5.7 การให้คะแนนโดยวิธีการประมาณค่าความเป็นไปได้สูงสุด (Maximum likelihood) มีหลายวิธี แต่วิธีที่นิยมใช้กันมากคือวิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข (Conditional maximum likelihood) การประมาณค่าโดยวิธีนี้มีข้อจำกัด คือ ถ้าผู้สอบตอบข้อสอบถูกหมดหรือผิดหมด จะไม่สามารถประมาณค่าได้ ขั้นตอนในการให้คะแนนมีดังนี้ (Hambleton and Swanminathan, 1985 อ้างถึงในสุพัฒน์ สุกมลสันต์, 2539; รังสรรค์ มณีเล็ก, 2540)

ขั้นที่ 1 ประมาณค่าความสามารถเริ่มต้น $(\theta_m; m = 0)$ โดยใช้สูตรดังนี้

$$\theta_0 = \ln\left(\frac{r_a}{k - r_a}\right)$$

เมื่อ $r_a = \sum a_i U_l$

U_l = 1 เมื่อตอบข้อทดสอบถูก

U_l = 0 เมื่อตอบข้อทดสอบผิด

a_i คือ อำนาจจำแนกของข้อทดสอบข้อที่ i

k คือ จำนวนข้อทดสอบทั้งหมดที่ผู้สอบตอบ

ขั้นที่ 2 หาค่า $P_i(\theta_m)$ และ $Q_i(\theta_m)$ โดยใช้สูตรดังนี้

$$P_i(\theta_m) = c_i + (1-c_i)\frac{e^{Da_i(\theta_m - b_i)}}{Da_i(\theta_m - b_i)}$$

$$Q_i(\theta_m) = 1 - P_i(\theta_m) = \frac{1-c_i}{1+e^{Da_i(\theta_m - b_i)}}$$

เมื่อ θ_m คือ ความสามารถของผู้สอบที่ประมาณค่าได้ในครั้งที่ m

a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i

b_i คือ ระดับความยากของข้อสอบข้อที่ i

c_i คือ ค่าการเดาของข้อสอบข้อที่ i

D คือ ค่าคงที่ซึ่งประมาณ 1.7

e คือ ค่าคงที่ซึ่งมีค่าประมาณ 2.7182818

ขั้นที่ 3 หาค่าปรับแก้ (h_m) โดยใช้สูตร

$$h_m = \frac{D\left[r_a - \sum P_i(\theta_m)\right]}{-D^2 \sum P_i(\theta_m) Q_i(\theta_m)}$$

ขั้นที่ 4 ประมาณค่าความสามารถของผู้สอบใหม่ (θ_{m+1}) โดยใช้สูตร

$$\theta_{m+1} = \theta_m - h_m$$

ขั้นที่ 5 ทำการคำนวณซ้ำในขั้นที่ 2,3 และ 4 จนกระทั่ง h_m เข้าใกล้ศูนย์ $(h_m < 0.001)$

6. เกณฑ์ยุติการทดสอบ (Termination criterion) ลักษณะสำคัญของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ คือ การทดสอบจะดำเนินไปเรื่อยจนกระทั่งถึงเกณฑ์ที่ระบุไว้ให้ยุติการทดสอบ การดำเนินการสอบก็จะยุติลง เกณฑ์ที่ใช้กันอยู่ทั่วไปมีอยู่ 2 เกณฑ์ คือ

6.1 การกำหนดจำนวนข้อสอบคงที่ เป็นการกำหนดจำนวนข้อสอบที่ใช้ในการทดสอบให้คงที่สำหรับทุกคน เช่น กำหนดให้การทดสอบแบบปรับเหมาะใช้ข้อสอบจำนวน 25 ข้อ (n = 25) เมื่อผู้สอบทำข้อสอบได้ครบ 25 ข้อ การทดสอบก็จะยุติลง เกณฑ์นี้ค่อนข้างเป็นประโยชน์ในการศึกษาภายใต้สถานการณ์จำลองแบบมอนติคาร์โล (Monte Carlo simulation) เนื่องจากจำนวนข้อสอบเท่ากันทำให้สามารถเปรียบเทียบสารสนเทศของแบบสอบได้โดยตรง ในทางปฏิบัติการกำหนดให้ทุกคนทำข้อสอบจำนวนเท่ากัน อาจมีคุณภาพของการวัดผลได้แตกต่างกัน (ศิริชัย กาญจนวาสี, 2538)

6.2 การกำหนดค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ในทางปฏิบัติการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ ควรให้ดำเนินการไปเรื่อยจนกว่าการประมาณค่า θ มีความคลาดเคลื่อนมาตรฐาน (Standard Error: SE (θ)) ลดต่ำลงจนถึงระดับที่ยอมรับได้ การทดสอบจึงควรยุติลง

2.5. วิธีการคัดเลือกข้อสอบขั้นแรก การควบคุมการใช้ข้อสอบซ้ำ และการสร้างความสมดุลของเนื้อหาในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (CAT)

วิธีการคัดเลือกข้อสอบขั้นแรก

การคัดเลือกข้อสอบมีความสัมพันธ์กับคุณภาพของการประมาณค่าคุณลักษณะ ($\hat{\theta}$) เป็นอย่างมาก การประมาณค่าคุณลักษณะยิ่งมีความถูกต้องแม่นยำมากยิ่งขึ้น ถ้าข้อสอบที่ถูกคัดเลือกมีความเหมาะสมกับความสามารถของผู้สอบ การประมาณค่าคุณลักษณะของผู้สอบด้วยข้อสอบที่เป็นขั้นแรกของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มีความคลาดเคลื่อนมากกว่าข้อสอบที่เป็นขั้นต่อมาจนถึงข้อสอบที่เป็นขั้นสุดท้ายของการทดสอบ ดังนั้นข้อสอบที่ถูกคัดเลือกที่ขั้นแรกมีแนวโน้มที่จะเหมาะสมน้อยกว่าข้อสอบที่ถูกคัดเลือกในขั้นต่อมา เพื่อปรับแก้ให้การคัดเลือกข้อสอบขั้นแรกของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มีประสิทธิภาพ ดังนั้นในไม่กี่ปีที่ผ่านมาได้มีการเสนอเกณฑ์การคัดเลือกข้อสอบที่ได้รับการปรับแก้แล้วหลายเกณฑ์ ได้แก่ สารสนเทศฟิชเชอร์แบบถ่วงน้ำหนัก (FII; Veerkamp and Berger, 1997) สารสนเทศของฟิชเชอร์ที่มีการแจกแจงภายหลัง (FIP; van der Linden, 1995 cited in Chen, Ankenmamm and Chang, 2000) สารสนเทศของคูลเบค-ไลเบลอร์ (KL; Chan and Ying, 1996) และสารสนเทศของคูลเบค-ไลเบอร์ที่มีการแจกแจงภายหลัง (KLP; Chang and Ying, 1996) ซึ่งมีรายละเอียดดังต่อไปนี้

1. วิธีสารสนเทศของฟิชเชอร์

ให้ $P_j(\theta)$ แทน ฟังก์ชันการตอบสนองข้อสอบข้อที่ j

$Q_j(\theta)$ แทน 1 - $P_j(\theta)$

ซึ่งฟังก์ชันสารสนเทศมีดังต่อไปนี้

$$I_j(\theta) = \frac{\left[\frac{\partial P_j(\theta)}{\partial \theta}\right]}{P_j(\theta)Q_j(\theta)} = \frac{\left[P_j'(\theta)\right]^2}{P_j(\theta)Q_j(\theta)} \qquad (1)$$

สมการที่ 1 เป็นการอธิบายเกณฑ์การคัดเลือกสารสนเทศของฟิชเชอร์ โดยกำหนดว่าข้อสอบที่ถูกเลือกจะให้ค่าสารสนเทศสูงสุด $I_j(\hat{\theta}_n)$ สำหรับการประมาณคุณลักษณะชั่วคราว ($\hat{\theta}_n$) ตามการใช้ข้อสอบ n ข้อ สำหรับโมเดลการตอบสนองข้อสอบโลจิสติก 3 พารามิเตอร์แบบทวิภาค (Dichotomous) สมการที่ 1 ทำให้เข้าใจง่าย ดังนี้ (Hambleton, Swaminathan, and Rogers, 1991)

$$I_j(\theta) = \frac{2.89a_j^2(1-c_j)}{\left[c_j + \exp\left[1.7a_j(\theta - b_j)\right]\right]\left[1+\exp\left[-1.7a_j(\theta-b_j)\right]\right]^2} \qquad (2)$$

ซึ่ง a_j เป็น พารามิเตอร์ค่าอำนาจจำแนกข้อสอบ

b_j เป็น พารามิเตอร์ค่าความยากของข้อสอบ

c_j เป็น พารามิเตอร์ระดับโอกาสการเดาข้อสอบได้ถูก

สมการที่ 2 แสดงให้เห็นว่าสารสนเทศที่เพิ่มขึ้น คือ b_j เข้าใกล้ (θ) , a_j เพิ่มขึ้น, และ c_j เข้าใกล้0 (Hambleton et al.,1991) เกณฑ์การคัดเลือกข้อสอบสารสนเทศของฟิชเชอร์ที่ใช้ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีข้อตกลงว่าเมื่อการประมาณค่าคุณลักษณะความสามารถของผู้สอบ $(\hat{\theta})$ ตามข้อสอบที่ถูกใช้ n ข้อซึ่งข้อสอบที่ถูกคัดเลือกเป็นข้อสอบที่ให้ค่าสารสนเทศสูงสุด $I_j(\hat{\theta})$

2. วิธีสารสนเทศของฟิชเชอร์ที่มีการถ่วงน้ำหนัก (FII)

Veerkamp and Berger (1997) ได้เสนอ เกณฑ์การคัดเลือกข้อสอบโดยทั่วไป ซึ่งเรียกว่า เกณฑ์สารสนเทศถ่วงน้ำหนักโดยทั่วไป (General weighted information criterion; GWIC)

$$GWIC_j(\theta) = \int_{-\infty}^{\infty} W(\theta) I_j \, d\theta \qquad (3)$$

ซึ่ง $W(\theta)$ เป็น ฟังชันถ่วงน้ำหนัก

$I_j(\theta)$ เป็น สารสนเทศของฟิชเชอร์ (นิยามไว้ในสมการที่ 1)

บนฐานของ GWI ค่าฟังก์ชันสารสนเทศของฟิชเชอร์ที่มากเกินช่วงระดับคุณลักษณะจะถูกรวมให้เป็นค่าเดียว คิดเป็นค่าเฉลี่ยถ่วงน้ำหนัก (Weighted average) เมื่อใช้เป็นเกณฑ์การคัดเลือกข้อสอบแล้ว ข้อสอบที่มีค่าเฉลี่ยถ่วงน้ำหนักสูงที่สุดจะเป็นทางเลือกหนึ่งสำหรับผู้สอบ

การสร้างเกณฑ์การคัดเลือกข้อสอบที่แตกต่างกันนั้น สามารถพัฒนาเป็นเกณฑ์คัดเลือกที่เหมาะสม โดยการปรับฟังก์ชันถ่วงน้ำหนัก เช่น เมื่อการประมาณคุณลักษณะเข้าใกล้ระดับ

คุณลักษณะที่แท้จริง ฟังชันสารสนเทศของฟิชเชอร์ที่สูงที่สุดเป็นการคัดเลือกข้อสอบที่เหมาะสม และฟังก์ชันถ่วงน้ำหนักในสมการที่ 3 สามารถแสดงได้ดังนี้

$$W(\theta) = \begin{cases} 1, & \theta = \hat{\theta}_n \\ 0, \text{ otherwise} \end{cases} \qquad (4)$$

(เช่น การใช้การประมาณคุณลักษณะแบบจุดเท่านั้น และไม่สนใจระดับคุณลักษณะที่เป็นไปได้อย่างอื่น) เป็นที่น่าสังเกตว่า เมื่อสมการที่ 4 ถูกใช้เป็นฟังชันถ่วงน้ำหนักใน GWIC สมการที่ (3) ซึ่งได้เป็นเกณฑ์การคัดเลือกสารสนเทศของฟิชเชอร์ (F-rule)

3. สารสนเทศของฟิชเชอร์ที่มีการแจกแจงภายหลัง (FIP)

จากวิธีสารสนเทศของฟิชเชอร์ที่มีการถ่วงน้ำหนัก (FII) เมื่อมีการอธิบายถึงความไม่แน่นอนของการประมาณคุณลักษณะในขั้นแรกของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ วิธีแก้ปัญหา คือ การกำหนดการแจกแจงภายหลังเป็นฟังก์ชันถ่วงน้ำหนักอาจจะเป็นฟังก์ชันที่เหมาะสมสำหรับการพิจารณาระดับคุณลักษณะทั้งหมด ในกรณีเกณฑ์การคัดเลือกข้อสอบ (แทนด้วย FIP) สามารถแสดงได้ดังนี้

$$FIP_j(X_n) = GWIC_j(X_n) = \int_{-\infty}^{\infty} P(\theta|X_n) I_j(\theta) d\theta \qquad (5)$$

ให้ $I_j(\theta)$ เป็น ฟังก์ชันสารสนเทศของฟิชเชอร์ สำหรับข้อที่ j

$P(\theta/X_n)$ เป็น การแจกแจงคุณลักษณะภายหลังหลังจากที่ข้อสอบ n ข้อถูกใช้สอบ

4. สารสนเทศคูลเบค-ไลเบลอร์ (KL)

Chang and Ying (1996) ได้แสดงให้เห็นฟังก์ชันสารสนเทศคูลเบค-ไลเบลอร์ดังต่อไปนี้

$$KL_j(\theta\|\theta_0) = P_j(\theta_0)\log\left[\frac{P_j(\theta_0)}{P(\theta)}\right] + [1-P_j(\theta_0)]\log\left[\frac{1-P_j(\theta_0)}{1-P_j(\theta)}\right] \qquad (6)$$

ให้ $P_j(\theta)$ แทน ฟังก์ชันการตอบสนองข้อสอบสำหรับข้อที่ j ซึ่ง θ และ θ_n แทน 2 ระดับคุณลักษณะ Chang and Ying สังเกตว่ามีคุณลักษณะที่สำคัญหลายประการของฟังก์ชันสารสนเทศคูลเบค-ไลเบลอร์ (KL) ได้แก่

1. $KL_j(\theta\|\theta_0) \neq KL(\theta_0\|\theta)$
2. $KL_j(\theta\|\theta_0) \geq 0$ และ $KL_j(\theta_0\|\theta_0) = 0$

3. คล้ายกับผลรวมของค่าฟังก์ชันสารสนเทศของฟิชเชอร์ กล่าวคือ สารสนเทศระดับแบบสอบเป็นผลรวมของสารสนเทศระดับข้อสอบ ดังนี้

$$KL^{n}(\theta\|\theta_0) = \sum_{j=1}^{n} KL_j(\theta\|\theta_0) \tag{7}$$

ในขณะที่ GWIC คือ ฟังก์ชันถ่วงน้ำหนักที่สามารถประยุกต์ใช้กับ $KL_j(\theta\|\theta_0)$ เพื่อที่จะหาเกณฑ์การคัดเลือกข้อสอบ ซึ่งฟังก์ชันถ่วงน้ำหนักทั่วไป คือ

$$W(\theta,\theta_0) = \begin{cases} 1, & \theta \in (\hat{\theta}-\delta_1, \hat{\theta}+\delta_2) \text{ และ } \theta_0 \in (\hat{\theta}-\delta_3, \hat{\theta}+\delta_4) \\ 0, & \text{อื่นๆ} \end{cases} \tag{8}$$

ซึ่งการประมาณค่าความสามารถ ($\hat{\theta}$) เป็น ค่าความสามารถที่แท้จริง (θ) ตามข้อสอบ n ข้อ และพื้นที่ที่ใช้เป็นเกณฑ์ในการคัดเลือกข้อสอบ กล่าวคือ ข้อสอบที่มีพื้นที่สูงสุด ซึ่งเท่ากับสารสนเทศ $KL_j(\theta\|\theta_0)$ ที่เฉลี่ยสูงสุดจะถูกคัดเลือก โดยมีสูตรดังนี้

$$KL_j(\hat{\theta}) = \int_{\hat{\theta}_l}^{\hat{\theta}_u} KL_j(\theta\|\hat{\theta})d\theta \tag{9}$$

เมื่อ $(\hat{\theta}_l, \hat{\theta}_u) = \left(\hat{\theta} - \frac{z}{\sqrt{n}}, \hat{\theta} + \frac{z}{\sqrt{n}}\right)$

$\frac{z}{\sqrt{n}}$ เป็น ลิมิตความเชื่อมั่น (confidence limit)

5. สารสนเทศคูลเบค-ไลเบลอร์ที่มีการแจกแจงภายหลัง (KLP)

Chang and Ying (1996) ได้แทนฟังก์ชันสารสนเทศฟิชเชอร์ ($I_j(\theta)$) ในสมการที่ 5 ด้วยสารสนเทศคูลเบค-ไลเบลอร์ ($KL(\theta\|\hat{\theta}_n)$) ซึ่งได้เกณฑ์การคัดเลือกข้อสอบดังนี้

$$KLP_j(\hat{\theta}_n, U_n) = \int_{-\infty}^{\infty} P(\theta|U_n) KL_j(\theta\|\hat{\theta}_n)d\theta \tag{10}$$

จากที่กล่าวมาสามารถสรุปความแตกต่างระหว่างเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์และคูลเบค-ไลเบลอร์ ได้ดังนี้

วิธีสารสนเทศของฟิชเชอร์	วิธีสารสนเทศของคูลเบค-ไลเบลอร์
Local information	Global information
ค่าฟังก์ชันสารสนเทศของข้อสอบ $Ij(\theta) = \frac{2.89a_j^2(1-c_j)}{\left[C_j + \exp(1.7a_j(\theta-b_j))(1+\exp(-1.7a_j(\theta-b_j)))\right]^2}$	ค่าฟังก์ชันสารสนเทศของข้อสอบ $KL_j(\theta\|\hat{\theta}) = P_j(\hat{\theta})\log\left[\frac{P_j(\hat{\theta})}{P(\theta)}\right] + [1-P_j(\hat{\theta})]\log\left[\frac{1-P_j(\hat{\theta})}{1-P_j(\theta)}\right]$
ค่าฟังก์ชันสารสนเทศของแบบสอบ $I^{(n)}(\theta) = \sum I_j(\theta)$	ค่าฟังก์ชันสารสนเทศของแบบสอบ $KL^{(n)}\left(\theta\|\hat{\theta}\right) = \sum KL_j\left(\theta\|\hat{\theta}\right)$
การคัดเลือกข้อสอบที่มีค่าสารสนเทศที่สูงที่สุดเพียงข้อเดียว (The item information at point) $Ij(\theta) = \frac{2.89a_j^2(1-c_j)}{\left[C_j + \exp(1.7a_j(\theta-b_j))(1+\exp(-1.7a_j(\theta-b_j)))\right]^2}$	การคัดเลือกข้อสอบที่มีเกณฑ์สารสนเทศแบบช่วง (Interval information criterion) $KL_j\left(\hat{\theta}\right) = \int_{\theta_l}^{\theta_u} KL_j(\theta\|\hat{\theta})d\theta$ $\left(\hat{\theta}_l, \hat{\theta}_u\right) = \left(\hat{\theta} - \frac{3}{\sqrt{n}+1}, \hat{\theta} + \frac{3}{\sqrt{n}+1}\right)$

การควบคุมการใช้ข้อสอบซ้ำ

ได้มีการเสนอวิธีการควบคุมการใช้ข้อสอบซ้ำต่างๆกันหลายวิธีเพื่อที่จะควบคุมการใช้ข้อสอบซ้ำในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (CATs) (Chang and Ying, 1999; Davey and Parshall, 1995; McBride and Martin, 1983; Revuelte and Ponsoda, 1998; Stocking and Lewis, 1998, 2000; Sympson and Hetter, 1985) ถึงแม้ว่าวิธีการส่วนใหญ่นี้สามารถที่จะควบคุมการใช้ข้อสอบซ้ำได้เป็นอย่างดีในลักษณะที่ว่าข้อสอบทั้งหลายที่ถูกใช้แล้วมีอัตราการใช้ข้อสอบซ้ำน้อยกว่าค่าที่ได้กำหนดไว้ล่วงหน้า จึงได้มีการบรรยายถึงวิธีการใช้ข้อสอบซ้ำหลายวิธีอย่างกระชับใจความที่ใช้กันแพร่หลายในทางปฏิบัติและมีส่วนเกี่ยวข้องกันมากกับการศึกษานี้

วิธีการของ Sympson and Hetter (1985) (SH)

การใช้ข้อสอบซ้ำสามารถที่จะควบคุมได้เป็นอย่างดีภายใต้วิธีการของ Sympson and Hetter (1985) ซึ่งจะเป็นพื้นฐานสำหรับวิธีการในการควบคุมการใช้ข้อสอบซ้ำที่มีประสิทธิภาพมากยิ่งขึ้น วิธีSHได้ถูกนำเสนอเพื่อที่จะควบคุมการใช้ข้อสอบซ้ำโดยตรงในลักษณะของความเป็นไปได้ เช่นข้อสอบที่ถูกใช้มากที่สุดด้วยความถี่ที่น้อยกว่าอัตราการใช้ข้อสอบซ้ำที่สูงสุดที่ได้มีการกำหนดไว้ล่วงหน้า ($r_{\max}$) ซึ่งมีความน่าจะเป็น 3 อย่างดังต่อไปนี้

$P(S)$ คือ ความน่าจะเป็นที่ข้อสอบที่ถูกเลือกเป็นข้อสอบที่ดีที่สุดตามวิธีการคำนวณ (Algorithm)

$P(A)$ คือ ความน่าจะเป็นที่ได้ถูกใช้จริงกับผู้เข้าสอบ

$P(A/S)$คือ ความน่าจะเป็นแบบมีเงื่อนไขที่ข้อสอบ 1 ข้อถูกใช้ถ้าข้อสอบนั้นได้ถูกเลือกเป็นข้อสอบที่ดีที่สุดซึ่งได้ถูกเรียกว่าเป็น “พารามิเตอร์การใช้ข้อสอบซ้ำ” ด้วยการที่จะเข้าถึงความต้องการ (met) ที่ไม่มีข้อสอบที่ถูกใช้มากกว่า $r_{\max}$ ความสัมพันธ์ระหว่างความน่าจะเป็นทั้ง 3 อย่างนี้สามารถอธิบายได้ดังนี้ คือ $P(A) = P(A/S) \times P(S) \leq r_{\max}$

พารามิเตอร์การใช้ข้อสอบซ้ำ คือ $P(A/S)$ ซึ่งมีบทบาทสำคัญในการควบคุมอัตราการใช้ข้อสอบซ้ำในวิธีการของ SH จุดประสงค์ของ $P(A/S)$ เพื่อปรับ $P(S)$ ในลักษณะที่สามารถทำให้ $P(A)$ น้อยกว่าหรือเท่ากับ $r_{\max}$ ถึงแม้จะเป็นการง่ายที่จะทำให้ทราบค่าของ $P(A/S)$ แต่ในการกำหนด $P(S)$ ก็ไม่ใช่เป็นเรื่องง่าย ยิ่งไปกว่านั้นเมื่อ $P(A/S)$ ได้ถูกกำหนดขึ้นสำหรับข้อสอบข้อใดข้อหนึ่งแล้วค่าของ $P(S)$ สำหรับข้อส่วนที่เหลือในคลังข้อสอบจะถูกเปลี่ยนแปลงด้วยค่าของ $P(A/S)$ ที่มีความเกี่ยวข้องกับข้อสอบเหล่านั้น ดังนั้นการจำลองข้อมูลซ้ำที่ต่อเนื่องกันเป็นชุดมีความจำเป็นที่ต้องทำเพื่อที่จะทำให้ค่าของ $P(S)$ และ $P(A/S)$ คงที่ซึ่งอยู่ภายใต้เงื่อนไขที่ค่าทั้งหมดของ $P(A)$ น้อยกว่าหรือเท่ากับ $r_{\max}$ ถ้าต้องการคำอธิบายมากกว่านี้สำหรับการจำลองทวนซ้ำให้ดูที่ Sympson and Hetter (1985) หรือ Sympson and Hetter (1997)

วิธีการแบบมีเงื่อนไขของ Stocking and Lewis (1998) (SLC)

ถึงแม้ว่าการใช้ข้อสอบซ้ำสามารถที่จะควบคุมได้ดีสำหรับข้อมูลทั้งหมดที่น่าสนใจตามวิธีของSH ซึ่งไม่สามารถรับประกันได้ว่าอัตราการใช้ข้อสอบซ้ำที่ระดับคุณลักษณะที่กำหนดให้ได้ถูกควบคุมเป็นอย่างดีด้วย ข้อสอบ 1 ข้ออาจจะถูกใช้ซ้ำกับผู้สอบเข้าสอบทั้งหมดที่ระดับคุณลักษณะเฉพาะ ถึงแม้ว่าอัตราการใช้ข้อสอบซ้ำโดยทั่ว ๆ ไปจะมีอัตราต่ำสำหรับผู้เข้าสอบในระดับคุณลักษณะทั้งหมด ดังนั้นการควบคุมอัตราการใช้ข้อสอบซ้ำที่ระดับคุณลักษณะแต่ละระดับดูเหมือนจำเป็นในทางปฏิบัติ

Stocking and Lewis (1998) ได้เสนอ วิธีการMultinomialแบบมีเงื่อนไขเพื่อที่จะควบคุมอัตราการใช้ข้อสอบซ้ำที่ระดับคุณลักษณะตามวิธีการของ Stocking และ Lewis (1998) การจำลองข้อมูลทวนซ้ำสำหรับการหาค่า $P(A/S)$ คงที่ได้ถูกดำเนินการภายใต้ระดับคุณลักษณะเฉพาะมากกว่าข้อมูลทั้งหมดที่น่าสนใจ โดยอาศัยเมตริกซ์ของค่า $P(A/S)$ ที่คงที่อัตราการใช้ข้อสอบซ้ำสามารถควบคุมได้อย่างมีประสิทธิภาพที่ระดับคุณลักษณะเฉพาะ อย่างไรก็ตามการจำลองข้อมูลทวนซ้ำสำหรับวิธี SLC เป็นการใช้เวลามาก

วิธีการของ Davey and Parshall (1995) (DP)

Stocking and Lewis (1998) ได้เสนอแนะให้ใช้วิธีการควบคุมการใช้ข้อสอบซ้ำที่ถูกกำหนดให้ที่ระดับคุณลักษณะ (Trait levels) แต่ว่า Davey and Parshall (1995) ได้เสนอวิธีการควบคุมการใช้ข้อสอบซ้ำที่ถูกกำหนดที่ข้อสอบ (Items) ที่ได้ใช้ในการทดสอบแล้วที่ Davey and Parshall (1995) ใช้วิธีนี้เพื่อที่จะหลีกเลี่ยงข้อสอบที่จะปรากฏเป็นกลุ่มย่อยมากเกินไปสำหรับผู้เข้าสอบ ตามวิธีของ Davey and Parshall (1995) การควบคุมการใช้ข้อสอบซ้ำสามารถที่จะปรับปรุงได้ที่ระดับแบบสอบและระดับข้อสอบแต่ละข้อ อย่างไรก็ตามวิธีนี้ไม่สามารถรับประกันได้ว่าอัตราคาบเกี่ยวกันระหว่างแบบสอบที่ได้กำหนดไว้ล่วงหน้าสามารถที่จะตอบสนองได้อย่างแน่ชัดเพื่อที่จะให้มีความปลอดภัยในแบบสอบ (Test security) การจำลองข้อมูลซ้ำสำหรับหาค่า P(A/S) ที่คงที่เป็นการใช้เวลามากด้วยเหมือนกัน โดยเฉพาะอย่างยิ่งเมื่อมีคลังข้อสอบจำนวนมาก ดังนั้นวิธี DP จึงเป็นการยากที่จะนำมาใช้ในการปฏิบัติ (Stocking and Lewis ,1998)

การสร้างความสมดุลของเนื้อหา (Content balancing)

การทดสอบมักจะกำหนดแบบแผนของเนื้อหา(Content specifications) เพื่อที่จะให้ได้แน่ใจได้ว่าผู้ที่เข้าสอบได้รับการทดสอบในเนื้อหาเท่ากันและเหมาะสม ในระหว่างการใช้ CAT วิธีการคัดเลือกข้อสอบควรจะเลือกข้อสอบที่ให้สารสนเทศมากที่สุดในการประมาณค่าความสามารถของผู้ที่เข้าสอบในช่วงเวลานั้น แต่ไม่บรรลุถึงการกำหนดแบบแผนของเนื้อหา (Content specifications) ของข้อสอบทั้งหมด ตัวอย่างเช่น การทดสอบทางด้านการคำนวณเพื่อที่จะให้ทราบถึงความสามารถ ของผู้เข้าสอบในเรื่อง บวก, ลบ, คูณ และหาร ทั้งหมด แต่ใน CAT อาจจะคัดเลือกให้ทำข้อสอบแต่เรื่อง ลบ และคูณเท่านั้น ซึ่งผู้เข้าสอบที่ไม่มีความรู้เรื่อง บวก และหาร พอดีก็จะถูกประเมินค่าความสามารถสูงเกินไป สำหรับผู้เข้าสอบที่มีความรู้เกี่ยวกับเรื่อง บวก และหาร ดีก็จะได้รับการประเมินค่าความสามารถที่แท้จริงของผู้เข้าสอบที่ต่ำเกินไป

การสร้างความสมดุลในเนื้อหา (Content balancing) มิได้เป็นปัญหาของการวัดทางจิต ซึ่งไม่เกี่ยวกับโมเดลการวัด IRT / TRT ซึ่งได้คัดเลือกสำหรับการทดสอบ เพื่อที่จะให้มีการวัดที่ถูกต้องและเป็นธรรมสำหรับเนื้อหาทั้งหมด การสร้างความสมดุลของเนื้อหาโดยระบบของ CAT ย่อมเป็นสิ่งที่จำเป็น (Aimee Michelle Boyd, 2003) ในตอนต่อไปนี้จะมีการกล่าวถึงวิธีการการสร้างความสมดุลของเนื้อหา(Content balancing)ในระบบCAT ซึ่งได้แก่ วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีเงื่อนไขบังคับ (Constrained CAT:CCAT), วิธีโมเดลมัลติโนเมลที่ได้แก้ไขปรับปรุงแล้ว (Modified multinomial model), วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีเงื่อนไขบังคับที่ได้มีการแก้ไขปรับปรุงแล้ว (Modified constrained CAT: MCCAT)

วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีเงื่อนไขบังคับ (Constrained CAT: CCAT)

ปรัชญาพื้นฐานของ CCAT คือ ข้อสอบควรจะถูกคัดเลือกจากขอบเขตของเนื้อหา (Content area) ซึ่งขอบข่ายนั้นมีคนทำถูกหรือข้อสอบถูกใช้ได้น้อยที่สุด ซึ่งในทางปฏิบัติข้อสอบที่คิดว่าเหมาะสมที่สุดจะได้รับการคัดเลือกจากขอบเขตเนื้อหา (Content area) ที่มีอัตราการใช้ข้อสอบซ้ำในปัจจุบันน้อยที่สุด คิดจากอัตราการใช้ข้อสอบซ้ำตามเป้าหมายที่ได้กำหนดไว้ล่วงหน้า

วิธีโมเดลมัลติโนเมลที่ได้แก้ไขปรับปรุงแล้ว (Modified multinomial model: MMM)

ทำการแจกแจงสะสม (Cumulative distribution) ขึ้นก่อนโดยเทียบอัตราการใช้ข้อสอบซ้ำตามเป้าหมายที่ได้กำหนดไว้ล่วงหน้าของขอบเขตเนื้อหา (Content areas) ทั้งหมดซึ่งรวมเป็นหน่วยเดียวกันแล้วตามด้วยการแจกแจงแบบ Multinomial (โดยเทียบเปอร์เซ็นต์เป้าหมายของขอบเขตของเนื้อหา (Content areas) ซึ่งมีผลรวมเป็น 1.0 เสร็จแล้วให้ใช้ตัวเลขสุ่มจากการแจกแจงแบบ Uniform U(0,1) เพื่อใช้กำหนดขอบเขตของเนื้อหา (Content areas) ให้ตรงกันในการแจกแจงแบบสะสม (Cumulative distribution) ซึ่งการแจกแจงสะสมดังกล่าวให้คัดเลือกข้อสอบที่เหมาะสมที่สุด อย่างไรก็ดีเนื่องจากความคลาดเคลื่อนในการสุ่ม เปอร์เซ็นต์ตามเป้าหมายของขอบเขตของเนื้อหา (Content areas) อาจจะไม่เป็นไปตามที่คาดไว้ในการทดสอบจริง ดังนั้นเมื่อใดก็ตามที่ขอบเขตของเนื้อหา (Content areas) เข้าถึงเปอร์เซ็นต์ที่ตั้งเป้าไว้ก็ต้องทำการปรับการแจกแจงแบบ Multinomial ใหม่โดยการใช้เปอร์เซ็นต์ของการที่บรรลุเป้าหมายของขอบเขตของเนื้อหานั้นยังเหลืออยู่ ในขณะที่กลไกการสุ่มได้รวบรวมอยู่ในวิธีนี้แล้วลำดับของขอบเขตเนื้อหา(Content areas) ต่างกัน

วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีเงื่อนไขบังคับที่ได้มีการปรับแก้แล้ว (Modified Constrained CAT: MCCAT)

วิธีนี้เป็นวิธีที่ได้ปรับปรุงมาจาก CCAT แทนที่จะพยายามคัดเลือกข้อสอบจากขอบเขต (Content areas) ซึ่งอยู่ต่ำสุดจากอัตราการใช้ข้อสอบซ้ำที่ได้กำหนดไว้ แต่ก็จะคัดเลือกเอาข้อสอบที่เหมาะสมจากขอบข่ายของเนื้อหาที่ยังไม่ได้ใช้ ด้วยวิธีนี้ผลที่ไม่พึงประสงค์จากการจัดลำดับโดยวิธี CCAT ก็จะสามารถหลีกเลี่ยงได้

2.6. ขั้นตอนในการทดสอบ (Testing Algorithm)

กระบวนการในการทดสอบเป็นหัวใจสำคัญของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ซึ่งจะต้องทราบว่าจะเริ่มต้น ดำเนินการทดสอบและสิ้นสุดการทดสอบอย่างไร

2.6.1. การเริ่มต้นการทดสอบ (Starting)

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จะให้ความสำคัญกับค่าความยากของข้อสอบข้อต่อไปที่จะนำมาเสนอแก่ผู้สอบ การพิจารณาความยากของข้อสอบจะพิจารณาจากการประมาณค่าความสามารถของผู้สอบในขณะนั้น เมื่อผู้เข้าสอบตอบข้อสอบในแต่ละครั้งความสามารถของผู้สอบจะถูกประมาณค่าใหม่ทุกครั้งเสมอ เมื่อเริ่มต้นการทดสอบมักจะไม่มีข้อมูลเกี่ยวกับความสามารถของผู้เข้าสอบ ดังนั้นการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ส่วนมากจะคัดเลือกข้อสอบที่มีความยากปานกลางเสนอแก่ผู้เข้าสอบเป็นข้อแรกในการเริ่มต้นการทดสอบ หรืออาจคัดเลือกข้อสอบที่ง่ายเป็นข้อสอบในการเริ่มต้นสอบเพื่อลดความวิตกกังวลและสร้างแรงจูงใจในการสอบของผู้เข้าสอบ

2.6.2. การดำเนินการสอบ (Continuing)

การดำเนินการสอบด้วยการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จะต้องคำนึงถึงประสิทธิภาพของการทดสอบ กล่าวคือ การทดสอบจะต้องใช้ข้อสอบน้อยข้อ และลดเวลาที่ใช้ในการทดสอบลง เพื่อลดความเครียดและความเหนื่อยล้าจากการทดสอบ แต่ผลการประมาณค่าความสามารถของผู้สอบได้ตรงกับระดับความสามารถที่แท้จริง ดังนั้นในขั้นตอนการดำเนินการทดสอบจะต้องกระทำอย่างรัดกุม กระบวนการคัดเลือกข้อสอบจะต้องสามารถคัดเลือกข้อสอบที่มีสารสนเทศ โดยเฉพาะค่าความยากจะต้องสอดคล้องกับค่าความสามารถของผู้สอบที่ถูกประมาณค่าครั้งสุดท้ายมากที่สุด การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์สามารถคัดเลือกข้อสอบเสนอแก่ผู้สอบได้ถูกต้องรวดเร็ว

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จึงเป็นกระบวนการทดสอบที่มีประสิทธิภาพสูง

2.6.3. การสิ้นสุดการทดสอบ (Stopping)

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จะสิ้นสุดลงเมื่อผลการสอบที่ผ่านมาเป็นไปตามเกณฑ์ยุติการทดสอบ ซึ่งเกณฑ์ยุติการทดสอบนั้นอาจจะกำหนดโดยใช้เกณฑ์จำนวนข้อสอบ เกณฑ์เวลาที่ใช้ในการทดสอบ เกณฑ์คะแนนจุดตัด (Cut off score) เพื่อแยกระดับความสามารถของผู้สอบออกเป็นกลุ่มรอบรู้และไม่รอบรู้ และเกณฑ์ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบเป็นไปตามเกณฑ์ยุติการทดสอบที่ตั้งไว้ การทดสอบแต่ละคนก็จะสิ้นสุดลง

2.7. **การดำเนินการทดสอบ**

การดำเนินการทดสอบเริ่มต้นโดยการให้ผู้สอบนั่งหน้าจอคอมพิวเตอร์ที่มีเครื่องขับแผ่นบันทึก (Disk drive) ในแผ่นบันทึกข้อมูลจะมีโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ที่มีองค์ประกอบต่าง ๆ ตามที่กำหนด ผู้ดำเนินการทดสอบจัดเตรียมเครื่องให้อยู่ในสภาพที่พร้อมจะทดสอบ และชี้แจงเกี่ยวกับวิธีการตอบข้อสอบ จากนั้นเครื่องคอมพิวเตอร์จะแสดงข้อสอบข้อแรกบนจอภาพเพื่อให้ผู้สอบได้ตอบ ข้อสอบข้อแรกนี้จะเป็นข้อสอบที่มีระดับความยากง่ายปานกลาง ถ้าผู้สอบตอบข้อสอบข้อแรกถูกข้อสอบข้อต่อไปจะเป็นข้อสอบที่มีระดับความยากเพิ่มขึ้น แต่ถ้าผู้สอบข้อสอบข้อแรกผิด ข้อสอบข้อต่อไปจะเป็นข้อสอบที่มีระดับความยากลดลง จากนั้นก็จะประมาณค่าความสามารถของผู้สอบ การทดสอบจะดำเนินต่อไปในตามแนวทางต่อไปนี้ (สุพัฒน์ สุกมลสันต์, 2539)

1. หลังจากประมาณค่าความสามารถของผู้สอบแล้ว เครื่องคอมพิวเตอร์ก็จะคัดเลือกข้อสอบข้อใหม่ที่มีความยากง่ายใกล้เคียงกับความสามารถของผู้สอบ
2. ผู้สอบตอบข้อสอบข้อใหม่ที่ปรากฏในหน้าจอคอมพิวเตอร์ คอมพิวเตอร์รับคำตอบและตรวจคำตอบ
3. ประมาณค่าความสามารถของผู้สอบอีกครั้งหนึ่งตามผลการตอบข้อสอบข้อที่ผ่านมา
4. พิจารณาว่าผู้สอบควรทำข้อสอบต่อไปหรือไม่ตามเกณฑ์ยุติการทดสอบที่กำหนดไว้

5. ถ้าผู้สอบต้องทำสอบข้อต่อไปก็ย้อนกลับไปขั้นที่ 1 อีกครั้งเมื่อสิ้นสุดการทดสอบ สารสนเทศเกี่ยวกับการตอบข้อสอบของผู้สอบก็จะถูกบันทึกไว้ เครื่องคอมพิวเตอร์ และแสดงผลการทดสอบให้ผู้สอบทราบ

2.8. ความแตกต่างระหว่างการทดสอบด้วยกระดาษและดินสอกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computer Adaptive Testing)

Straetmans and Eggen (1998 อ้างถึงใน เกียรติศักดิ์ ส่องแสง, 2547) กล่าวว่า การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แตกต่างจากแบบทดสอบที่ใช้กระดาษและดินสอ ดังนี้

2.8.1. ความยาก (Difficulty)

แบบทดสอบที่ใช้กระดาษและดินสอ โดยเฉพาะข้อสอบแบบเลือกตอบ มีการใช้กันอย่างแพร่หลายในปัจจุบัน นักเรียนทุกคนจะต้องทำการทดสอบในเวลาเดียวกันด้วยข้อสอบชุดเดียวกัน ระดับความยากของแบบทดสอบจะถูกปรับตามค่าเฉลี่ยความสามารถของกลุ่มผู้เข้าสอบ ซึ่งมีผลทำให้ผู้เข้าสอบที่มีความสามารถสูงหรือต่ำกว่าค่าเฉลี่ยเกิดความสะเพร่าหรือมีความเครียดเป็นอย่างมาก ผู้เข้าสอบจะเผชิญกับข้อสอบที่ยากหรือง่ายเกินไป ตามหลักจิตวิทยาแล้วความถูกต้องของการวัดผลการทดสอบจะลดลง ส่วนการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ผู้เข้าสอบแต่ละคนจะได้รับแบบทดสอบที่แตกต่างกัน ซึ่งเป็นแบบทดสอบที่ไม่ยากหรือง่ายเกินไป ดังนั้นแบบทดสอบจึงท้าทายความรู้สึกของผู้เข้าสอบ ผู้เข้าสอบไม่วิตกกังวล ความถูกต้องของการวัดของผู้สอบแต่ละคนจะมีค่าสูง

2.8.2. ความยาวของแบบทดสอบ (Test length)

การทดสอบที่ใช้กระดาษและดินสอจะกำหนดความยาวของแบบทดสอบตายตัว ความยาวของแบบทดสอบนั้นจะถูกกำหนดจากกลุ่มผู้เข้าสอบทั้งหมด ผู้เข้าสอบทุกคนจะต้องใช้ข้อสอบชุดเดียวกัน ถ้าความสามารถของกลุ่มผู้เข้าสอบกระจายเป็นโค้งปกติแล้ว แบบทดสอบที่ดีจะต้องมีความยากกระจายไปให้ครอบคลุมความสามารถของกลุ่มผู้เข้าสอบด้วย ดังนั้นจึงมีข้อสอบที่ไม่เหมาะกับความสามารถของผู้เข้าสอบแต่ละคนรวมอยู่ด้วย ส่วนการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ผู้เข้าสอบแต่ละคนจะได้รับแบบทดสอบที่แตกต่างกัน ซึ่งจะถูกปรับตามความสามารถของผู้เข้าสอบ หลังจากทำข้อสอบในแต่ละข้อแล้วความสามารถของผู้เข้าสอบแต่ละคนจะถูกประมาณค่าใหม่ทุกครั้ง จนกระทั่งประมาณค่าความสามารถของผู้สอบครั้งหลังสุดเป็นไปตามเกณฑ์ยุติการทดสอบ การทดสอบก็จะสิ้นสุดลง

2.8.3. ความปลอดภัย (Security)

การทดสอบที่ใช้กระดาษและดินสอ ผู้เข้าสอบจะต้องทำแบบทดสอบชุดเดียวกันในเวลาเดียวกันด้วยจึงจะทำให้ข้อสอบไม่รั่ว ถ้าทำการทดสอบในเวลาที่ต่างกัน ความเชื่อมั่นในการทดสอบจะลดลง แต่การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์สามารถแก้ปัญหานี้ได้ โดยการให้ผู้เข้าสอบแต่ละคนทดสอบโดยใช้ข้อสอบที่ต่างกัน ในเวลาที่ต่างกันได้ด้วย แต่ผลการทดสอบยังคงเปรียบเทียบกันได้

2.8.4. ความสมดุลของเนื้อหา (Balancing of Content)

การสร้างแบบทดสอบที่ใช้กระดาษและดินสอ จะสร้างตามแผนการออกข้อสอบทั้งนี้เพื่อให้เกิดความมั่นใจว่า ข้อสอบที่เขียนออกมานั้นเป็นตัวแทนของความรู้และทักษะที่กำหนดให้ในจุดประสงค์ของหลักสูตร แต่ข้อสอบในแบบทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์จะถูกคัดเลือกโดยยุทธศาสตร์การคัดเลือกจากสารสนเทศสูงสุด เนื้อในแบบทดสอบจึงถูกกำหนดโดยพิจารณาตามหลักการทางจิตวิทยา ดังนั้นผู้เข้าสอบจะได้รับข้อสอบที่มีความสมดุล ถึงแม้ว่าการทดสอบจะสิ้นสุดหลังจากที่ทำแบบทดสอบที่มีจำนวนข้อสอบเพียงไม่กี่ข้อก็ตาม

2.8.5. วิธีการควบคุม (Locus of Control)

การทดสอบที่ใช้กระดาษและดินสอผู้ทำข้อสอบจะต้องตัดสินว่า คำตอบที่ได้จากแบบทดสอบจะนำไปทำอะไร มีความเป็นไปได้ที่จะมีคำตอบที่แตกต่างกันในคำถามเดียวกัน จึงอาจมีการทดสอบโดยให้เปิดหนังสือตอบและมีการทวนคำตอบต่าง ๆ ซึ่งปกติแล้ว ผู้บริหารการสอบจะเป็นผู้อนุญาตให้ทำในลักษณะใดลักษณะหนึ่ง ส่วนการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ คอมพิวเตอร์จะเป็นผู้ตัดสินว่า คำถามข้อต่อไปควรจะเป็นอะไร การควบคุมกระบวนการทดสอบจะรัดกุมยิ่งขึ้น

2.8.6. การบริหารจัดการ (Logistics)

การทดสอบที่ใช้กระดาษเขียนตอบจะต้องมีครูผู้ควบคุมการสอบหรือผู้ที่เกี่ยวข้องกับการสอบจำนวนมาก แบบทดสอบจะนำมาจากธนาคารข้อสอบหรือถ่ายสำเนา ต้องมีการใช้ห้องสอบและจ้างผู้ควบคุมการสอบ หลังจากทำการทดสอบจะมีการเก็บรวบรวมกระดาษคำตอบและข้อสอบพร้อมทั้งทำเครื่องหมายต่าง ๆ ไว้ ส่วนการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ คอมพิวเตอร์ช่วยลดภาระงานต่าง ๆ เหล่านี้มาจากครูและลดอุปสรรคในการทดสอบลง

จากที่กล่าวข้างต้น Straetmans and Eggen (1998 อ้างถึงใน เกียรติศักดิ์ ส่องแสง, 2547) สามารถสรุปได้ดังตารางที่ 1

ตารางที่ 1 ความแตกต่างระหว่างแบบทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์กับแบบทดสอบที่ใช้กระดาษและดินสอ

ประเด็น	แบบทดสอบที่ใช้กระดาษและดินสอ	แบบทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์
ความยาก	ปรับความสามารถของผู้เข้าสอบให้อยู่ในระดับปานกลาง	ปรับความสามารถของผู้เข้าสอบเป็นรายบุคคล
ความยาวของแบบทดสอบ	ระบุเจาะจงและข้อสอบมีจำนวนมากข้อ	ไม่ระบุเจาะและข้อสอบจำนวนน้อยข้อ
ความปลอดภัย	มีปัญหาเรื่องความปลอดภัยเมื่อผู้เข้าสอบไม่ได้สอบในเวลาเดียวกัน	ป้องกันความปลอดภัยได้ดีกว่าเพราะผู้เข้าสอบแต่ละคนจะใช้แบบทดสอบที่ต่างกัน
ความสมดุลของเนื้อหา	ไม่มีปัญหา	ไม่ได้อยู่ภายใต้การควบคุมโดยปราศจากการวัดผลเป็นพิเศษ
วิธีการควบคุม	ผู้เข้าสอบอยู่ในการควบคุม	คอมพิวเตอร์อยู่ในการควบคุม
การบริหารจัดการ	ทำงานหนัก	กระจายการทำงาน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

จากการศึกษางานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ผู้วิจัยสรุปเป็นประเด็นต่าง ๆ ได้ ดังนี้

โมเดลการตอบสนองข้อสอบ (Item response model)

จากการสร้างสถานการณ์จำลองข้อมูลของ Urry (1977) พบว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบนั้นจะไม่ค่อยมีประสิทธิภาพหากเลือกใช้โมเดลการวิเคราะห์ข้อมูลไม่ค่อยเหมาะสม เช่น ค่าความตรงเชิงโครงสร้างของแบบสอบจะลดลงเมื่อเทียบกับการใช้แบบสอบที่มีความยาวเท่ากันแต่เลือกใช้โมเดลการวิเคราะห์ข้อสอบที่ดีกว่า ซึ่งโมเดลโลจิสติก 3 พารามิเตอร์มีความเหมาะสมมากกว่าโมเดลอื่น ๆ

คลังข้อสอบ (Item pool)

Ho (1989) พบว่า ขนาดของคลังข้อสอบส่งผลต่อการประมาณค่าความสามารถของผู้สอบน้อยมาก เว้นแต่กลุ่มผู้สอบที่มีความสามารถต่ำ และเมื่อใช้ข้อสอบเริ่มต้นที่มีค่าความยากง่ายสูงกว่าความสามารถที่แท้จริงของผู้สอบ และยังพบว่า ชนิดของคลังข้อสอบ (ข้อสอบแบบสุ่ม ข้อสอบที่มีค่าอำนาจจำแนกสูง และข้อสอบที่ง่าย ๆ) นั้น มีปฏิสัมพันธ์กับวิธีการประมาณค่าความสามารถของผู้สอบ (วิธีของเบส์ (Bayesian) วิธีโมดัลเบส์ (Modal bayesian) และวิธีความเป็นไปได้สูงสุด Maximum likelihood)) และชนิดของคลังข้อสอบมีปฏิสัมพันธ์กับระดับความสามารถของผู้สอบ (สูง ปานกลาง ต่ำ) Weiss (1988) ได้เสนอแนะว่า ขนาดของคลังข้อสอบที่มีขนาดพอเหมาะนั้นควรมีประมาณ 100-200 ข้อ ซึ่ง 116 -150 ข้อจะให้ผลดีที่สุด ซึ่งสอดคล้องกับ Urry (1977) ที่พบว่าคลังข้อสอบที่วัดคุณลักษณะเดียวกันควรมีอย่างน้อย 100 ข้อ ส่วนข้อสอบในคลังข้อสอบนั้น ควรมีค่าอำนาจจำแนกมากกว่า 0.8 ค่าความยากง่ายอยู่ระหว่าง -2.0 ถึง 2.0 และค่าสัมประสิทธิ์การเดาน้อยกว่า 0.3

ความยากง่ายของข้อสอบข้อแรก (Entry level)

ในการเลือกข้อสอบข้อแรกให้กับผู้สอบนั้น พบว่า ลักษณะของข้อสอบข้อแรกไม่มีผลต่อความถูกต้องในการประมาณค่าความสามารถของผู้สอบเมื่อสิ้นสุดการทดสอบ (Lord, 1980) แต่ Ho (1989) พบว่า ถ้าข้อสอบข้อแรกมีความยากง่ายเท่ากับหรือต่ำกว่าความสามารถของผู้สอบแล้วจะทำให้การประมาณค่าความสามารถของผู้สอบมีความถูกต้องมากขึ้น และรังสรรค์ มณีเล็ก (2540) พบว่า ความยากง่ายของข้อสอบข้อแรกส่งผลต่อจำนวนข้อที่ใช้ในการทดสอบ

เกณฑ์การคัดเลือกข้อสอบจากคลังข้อสอบ (Selection rule)

Hulin and other (1983) พบว่า การใช้เกณฑ์การคัดเลือกข้อสอบโดยการจับคู่ระหว่างค่าความสามารถของผู้สอบกับค่า θ_{max} เพื่อแก้ปัญหาการใช้ข้อสอบมากเกินไปนั้น พบว่า มีการสูญเสียสารสนเทศมากกว่าการคัดเลือกข้อสอบโดยวิธีสารสนเทศสูงสุด (I_{max}) หรือเรียกอีกอย่างว่าวิธีสารสนเทศของฟิชเชอร์ (FI) ส่วน รังสรรค์ มณีเล็ก (2540) พบว่า เกณฑ์การคัดเลือกข้อสอบจากคลังข้อสอบ มีผลต่อจำนวนข้อสอบที่ใช้ในการทดสอบ

เกณฑ์การคัดเลือกข้อสอบขั้นแรก (Item selection rules at the early stage)

Veerkamp and Berger (1997) พบว่า วิธีสารสนเทศของฟิชเชอร์ที่มีการถ่วงน้ำหนัก (FII) มีประสิทธิภาพไม่ดีกว่าวิธีสารสนเทศของฟิชเชอร์ (FI) และ Fan and Hsu (1996 cited in Chen, Ankenmann and Chang, 2000) พบว่า วิธีสารสนเทศของคูลเบค-ไลเบลอร์ (KL) มีประสิทธิภาพไม่ดีกว่าวิธีสารสนเทศของฟิชเชอร์ (FI) ซึ่งไม่สอดคล้องกับ Chang and Ying (1996) พบว่า วิธีสารสนเทศของคูลเบค-ไลเบลอร์ (KL) มีประสิทธิภาพดีกว่าวิธีสารสนเทศของฟิชเชอร์ (FI) และ Chen, Ankenmann and Chang (2000) พบว่า วิธีสารสนเทศของฟิชเชอร์ที่มีการถ่วงน้ำหนัก (FII) วิธีสารสนเทศของฟิชเชอร์ที่มีการแจกแจงภายหลัง (FIP) วิธีสารสนเทศของคูลเบค-ไลเบลอร์ (KL) และวิธีสารสนเทศของคูลเบค-ไลเบลอร์ที่มีการแจกแจงภายหลัง (KLP) มีประสิทธิภาพและความแม่นยำในการประมาณค่าคุณลักษณะที่อยู่ปลายสุด (θ = -3,-2) มากกว่า วิธีสารสนเทศของฟิชเชอร์ (FI) และเมื่อแบบสอบที่มีความยาวมากกว่า 10 ข้อไม่ปรากฏข้อได้เปรียบเกี่ยวกับความแม่นยำในการประมาณคุณลักษณะสำหรับเกณฑ์การคัดเลือกข้อสอบ และ Chen and Ankenmann (2004) พบว่า เมื่อความยาวของแบบสอบน้อยกว่า 10 ข้อ วิธีสารสนเทศของฟิชเชอร์ที่มีการแจกแจงภายหลัง (FIP) และวิธีสารสนเทศของคูลเบค-ไลเบลอร์ที่มีการแจกแจงภายหลัง (KLP) มีแนวโน้มให้ผลดีกว่าวิธีสารสนเทศของฟิชเชอร์ (FI) ที่ระดับของคุณลักษณะที่อยู่ปลายสุด (extreme) อย่างไรก็ดีในการตั้งค่าให้สมจริง (realistic) จากการจำลองข้อมูลไม่สามารถสรุปได้ว่าวิธีสารสนเทศของฟิชเชอร์ที่มีการแจกแจงภายหลัง (FIP) และวิธีสารสนเทศของคูลเบค-ไลเบลอร์ที่มีการแจกแจงภายหลัง (KLP) จะมีประสิทธิภาพดีกว่าวิธีสารสนเทศของฟิชเชอร์ (FI) โดยเฉพาะอย่างยิ่งเมื่อได้กำหนดการควบคุมอัตราการใช้ข้อสอบซ้ำเข้าไป เมื่อความยาวของแบบสอบที่มากกว่า 10 ข้อ วิธีการคัดเลือกข้อสอบแบบสารสนเทศของฟิชเชอร์ (FI) สารสนเทศของฟิชเชอร์ที่มีการแจกแจงภายหลัง (FIP) และสารสนเทศของคูลเบค-ไลเบลอร์ที่มีการแจกแจงภายหลัง (KLP) นี้จะมีลักษณะที่ไม่แตกต่างกัน ในขณะวิธีสารสนเทศของฟิชเชอร์ (FI) จะมีการใช้ข้อสอบสูงกว่าเล็กน้อย

วิธีการประมาณค่าความสามารถของผู้สอบ (Ability estimation procedure)

ต่าย เซี่ยงฉี (2534) พบว่า วิธีการประมาณค่าความสามารถตามวิธีของเบส์ที่ปรับปรุงใหม่ (Beyesian updating) และวิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข(Conditional maximum likelihood)ให้ผลไม่แตกต่างกัน และพบว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบรูปปิรามิดที่มีรูปแบบ จำนวนขั้นและวิธีการให้คะแนนต่างกันจะส่งผลต่อความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ นอกจากนี้ยังพบว่าความตรงเชิงโครงสร้างของแบบสอบปิรามิดแบบขั้นคงที่ที่มี 16 ขั้น ประมาณค่าความสามารถด้วยวิธีของเบส์ มีประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงสุด ซึ่งข้อค้นพบดังกล่าวมีบางส่วนที่สอดคล้องกับ Skaggs and Stevenson (1989 อ้างถึงในรังสรรค์ มณีเล็ก, 2540) ที่พบว่า หากผู้สอบมากกว่า 2000 คนแล้ว ผลที่ได้รับจากการประมาณค่าความสามารถทั้ง 2 วิธีไม่แตกต่างกัน แต่ถ้ามีผู้สอบน้อยกว่า 500 คนแล้ว วิธีของเบส์จะมีความคงที่และถูกต้องมากกว่า และถ้าข้อสอบที่ใช้ในการสอบน้อยกว่า 15 ข้อแล้ว วิธีความเป็นไปได้สูงสุด จะมีความถูกต้องมากกว่า มีบางส่วนสอดคล้องกับ ข้อค้นพบ Gifford and Swaminathan (1990) พบว่า ถ้ากลุ่มผู้สอบมีขนาดเล็กแล้วจำนวนผู้ใช้มีจำนวนน้อยข้อแล้วการประมาณค่าความสามารถด้วยวิธีของเบส์แบบร่วมกัน (Joint bayesian) จะมีความถูกต้องมากกว่า นอกจากนี้ รังสรรค์ มณีเล็ก (2540) พบว่า วิธีการประมาณค่าความสามารถของผู้สอบมีผลต่อค่าความตรงเชิงสภาพของการทดสอบจำนวนข้อสอบที่ใช้ในการทดสอบ

ประสิทธิภาพของการทดสอบ

Balckmore (1987) พบว่า คะแนนที่ได้จากการทดสอบด้วยกระดาษและดินสอ การทดสอบด้วยคอมพิวเตอร์ และการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ไม่แตกต่างกัน และนันทิยา พึ่งคำ (2531) พบว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จะให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่ำกว่าและใช้จำนวนข้อสอบน้อยกว่าการทดสอบแบบประเพณีนิยม แต่ให้ความตรงในการประมาณค่าความสามารถเหมือนเดิม และจากการศึกษาของ ต่าย เซี่ยงฉี (2534) พบว่า ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้แบบทดสอบรูปปิรามิด เมื่อจำนวนขั้นเพิ่มขึ้น (ขนาดขั้นคงที่และขนาดขั้นแปรผัน) จะทำให้ประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงขึ้น รวมถึง รังสรรค์ มณีเล็ก (2540) พบว่า วิธีการประมาณค่าความสามารถของผู้สอบ เกณฑ์ยุติการทดสอบ และความสามารถของผู้สอบ มีผลต่อความตรงตามสภาพ และความยากง่ายของข้อสอบข้อแรก เกณฑ์การคัดเลือกข้อสอบ วิธีประมาณค่าความสามารถของผู้สอบ เกณฑ์ยุติการสอบ และความสามารถของผู้สอบส่งผลต่อจำนวนข้อที่ใช้ในการทดสอบ จากการศึกษาของ กนกวรรณ รัตนธน (2544) พบว่า ระดับความสามารถของผู้สอบ

และอัตราการใช้ข้อสอบซ้ำที่แตกต่างกันส่งผลต่อความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบ นอกจากนี้ เกียรติศักดิ์ ส่องแสง (2547) พบว่า การจัดชุดข้อสอบให้ตอบขั้นละ 3 ข้อ ให้ค่าฟังก์ชันสารสนเทศของแบบสอบสูงที่สุด รองลงมาได้แก่ การจัดชุดข้อสอบให้ตอบขั้นละ 2 ข้อ ส่วนการกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อนั้น พบว่า การกำหนดเวลาให้ตอบข้อสอบเฉลี่ยเท่ากันทุกข้อ หรือ การกำหนดเวลาให้ตอบข้อสอบในแต่ละข้อตามค่าความยากของข้อสอบ จะทำให้ผู้สอบใช้เวลาในการทดสอบน้อยลง และการจัดชุดของข้อสอบที่มีสื่อประกอบการทดสอบ จะทำให้ผู้สอบใช้เวลาในการทดสอบน้อยลง และผู้สอบมีความพึงพอใจในการทดสอบมากยิ่งขึ้น

อัตราการใช้ข้อสอบซ้ำ

Chang (1998) พบว่าวิธีการควบคุมอัตราการใช้ข้อสอบซ้ำด้วยวิธีของ Sympson and Hetter มีประสิทธิภาพมาก และวิธีของ Stocking and Lewis ทำให้มีความปลอดภัยของแบบสอบเนื่องมาจากการใช้ข้อสอบซ้ำ แต่ลดความถูกต้องของการทดสอบโดยที่การสอบแต่ละครั้งควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 20 เปอร์เซ็นต์ของจำนวนข้อสอบที่ผู้สอบได้รับผ่านมา ส่วน Kalohn and Sparay (1998) พบว่า วิธีการคัดเลือกข้อสอบโดยมีอัตราการใช้ข้อสอบซ้ำ 10 และ 20 เปอร์เซ็นต์ มีความถูกต้องเหมือนกันและผู้สอบที่มีความสามารถระดับสูงควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 10 เปอร์เซ็นต์ของจำนวนข้อสอบที่ผู้สอบได้รับในการทดสอบครั้งที่ผ่านมา และ Revuelta and Ponsoda (1998) พบว่า วิธี Restricted ใช้ในการลดอัตราการใช้ข้อสอบซ้ำได้สูงสุด และ วิธี Progressive ลดจำนวนข้อสอบภายในคลังข้อสอบที่ยังไม่ได้ใช้ ซึ่งการควบคุมอัตราการใช้ข้อสอบซ้ำมีจุดมุ่งหมายที่สำคัญ 2 ประการ คือ เพื่อป้องกันอัตราการใช้ข้อสอบซ้ำที่มากเกินไป และเพิ่มอัตราการใช้ข้อสอบที่ไม่ได้ใช้ภายในคลังข้อสอบ รวมถึง Chang and Twu (1998) ได้ศึกษาการจำลองข้อมูลโดยใช้วิธีมอนติคาร์โล (Monte Carlo simulation) ทำการเปรียบเทียบวิธีการควบคุมอัตราการใช้ข้อสอบซ้ำ 5 วิธี โดยการพิจารณาภายใต้ขนาดของคลังข้อสอบที่แตกต่างกัน อัตราการใช้ข้อสอบซ้ำสูงสุดที่แตกต่างกัน และความคลาดเคลื่อนมาตรฐานของการวัด ผลการศึกษาพบว่าวิธีการแบบมีเงื่อนไขที่พัฒนาโดย Stocking and C.Lewis (1995) เป็นวิธีการที่ดีที่สุด มีความถูกต้องแม่นยำของการวัด นอกจากนี้ Chen, Ankenmann and Spray (2003) ได้เสนอการวิเคราะห์การคำนวณของดัชนีการคาบเกี่ยวกันของแบบสอบโดยเฉลี่ยเป็นฟังก์ชันของดัชนีการใช้ข้อสอบซ้ำในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีจำนวนข้อสอบที่แน่นอน ผลการวิจัยชี้ให้เห็นว่าการควบคุมโดยการใช้ข้อสอบซ้ำโดยการระบุ r_{max} (เช่น วิธีของ Sympson and Hetter ,1985) ให้การควบคุมโดยตรงมากที่สุดทั้งระดับข้อสอบและระดับแบบสอบ รวมทั้ง Chen and Lei (2005) ได้เสนอวิธีการของ Sympson and Hetter ซึ่งสามารถใช้ควบคุมการใช้ข้อสอบซ้ำทั้งระดับข้อสอบแต่

ละข้อ และระดับแบบสอบได้ในเวลาเดียวกัน ผลการวิจัยแสดงให้เห็นว่าการควบคุมอัตราการใช้ข้อสอบซ้ำ (Item exposure rate) จะได้รับการปรับปรุงให้ดีขึ้นและความแม่นยำ (Precision) ของการประมาณค่าคุณลักษณะจะลดลงเมื่ออัตราคาบเกี่ยวกันระหว่างแบบสอบที่สูงสุดที่ได้กำหนดไว้ล่วงหน้ามีความเข้ม (Stringent) มากขึ้น

การสร้างความสมดุลของเนื้อหา (Content balancing)

Leung, Chang and Hua (2003) ได้ศึกษาโดยการสร้างสถานการณ์จำลองข้อมูล ทำการเปรียบเทียบวิธีการสร้างความสมดุลของเนื้อหา(Content balancing) 3 วิธีประกอบด้วย วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีเงื่อนไขบังคับ (Constrained CAT:CCAT), วิธีโมเดลมัลติโนเมลที่ได้แก้ไขปรับปรุงแล้ว (Modified multinomial model), วิธีการทดสอบแบบปรับเมาะด้วยคอมพิวเตอร์ที่มีเงื่อนไขบังคับที่ได้มีการแก้ไขปรับปรุงแล้ว (Modified constrained CAT: MCCAT) ภายใต้เงื่อนไขจำนวนข้อสอบในแบบสอบ(Test length) และอัตราการใช้ข้อสอบซ้ำสูงสุดที่ได้กำหนดไว้ (Target maximum exposure rate) ที่แตกต่างกัน ผลของการศึกษาการจำลองข้อมูลหลายชุดชี้ให้เห็นว่า ไม่มีผลที่เป็นระบบ (Systematic) ของวิธีการสร้างความสมดุลของเนื้อหา ในประสิทธิภาพการวัดและการใช้คลังข้อสอบ อย่างไรก็ดี ระหว่าง 3 วิธีนี้ วิธี MMM มีการใช้ข้อสอบน้อยกว่า วิธี CCAT และ MCCAT

เกณฑ์ยุติการทดสอบ

Dodd, Koch and De Alaya (1993) ได้ศึกษาผลกระทบของลักษณะของกลุ่มข้อสอบและความแตกต่างของเกณฑ์ยุติการทดสอบ จากการจำลองข้อมูลโดยใช้ Partial credit model พบว่า การยุติโดยการดูค่าความคลาดเคลื่อนมาตรฐาน จะให้ผลดีกว่าการยุติโดยดูค่าสารสนเทศของข้อสอบ และยังพบว่ากลุ่มข้อสอบ (Item pool) ที่มีข้อสอบน้อยที่สุดที่สามารถใช้ได้ คือ 30 ข้อ และควรมีค่าความยากง่ายปานกลาง ซึ่งสอดคล้องกับ การจำลองข้อมูลของ Dodd, Koch and De Alaya (1989 cited in Hot, Vorst and Mellenbergh, 2005) โดยใช้ Graded response model พบว่า เกณฑ์ยุติการทดสอบโดยใช้ค่าความคลาดเคลื่อนมาตรฐาน จะให้ผลดีกว่าเกณฑ์ยุติการทดสอบที่ใช้ค่าสารสนเทศของข้อสอบ ส่วน De Alaya, Dodd and Koch (1990) ได้เปรียบเทียบการทดสอบโดยใช้แบบสอบเฟล็กซิเลเวลกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ที่ประมาณค่าความสามารถด้วยวิธีของเบส์ โดยใช้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถเป็นเกณฑ์ยุติการทดสอบ ซึ่งได้เปรียบเทียบที่จุดต่าง ๆ ของค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ผลจากการศึกษาพบว่า ค่าความสามารถของผู้สอบที่ได้จากการทดสอบสองวิธีมีความถูกต้องแม่นยำไม่ต่างกัน และตรึงใจ พูลผลอำนวย (2534) ได้พัฒนา

แบบสอบเฉพาะบุคคล (Tailored test) โดยใช้รูปแบบแยกทางคงที่รูปปิรามิดขนาดขั้นคงที่ 10 ขั้น ในวิชาวิทยาศาสตร์โดยใช้คอมพิวเตอร์ในการทดสอบ พบว่า ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของแบบสอบเฉพาะบุคคลท่ากับ 0.443 รวมถึงผลการศึกษาของรังสรรค์ มณีเล็ก (2540) พบว่า การกำหนดค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 จะให้ค่าความตรงตามสภาพสูงสุด นอกจากนี้ Gushta (2003) ได้แสดงตัวอย่างโดยการจำลองข้อมูลเปรียบเทียบเกณฑ์ยุติการทดสอบ โดยใช้การกำหนดจำนวนข้อสอบให้เท่ากับ 25 ข้อเป็นเกณฑ์ยุติการทดสอบ และใช้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบเป็นเกณฑ์ยุติการทดสอบ ซึ่งกำหนดให้น้อยกว่าหรือเท่ากับ 0.30 พบว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบจากการกำหนดจำนวนข้อสอบเป็นเกณฑ์ยุติการทดสอบมีค่าอยู่ในช่วง 0.365 - 0.434 แต่ถ้ากำหนดค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบเป็นเกณฑ์ยุติการทดสอบ จะได้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับเกณฑ์ยุติที่กำหนด ซึ่งมีค่าอยู่ในช่วง 0.298 – 0.3 เนื่องจากกำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.3 เป็นเกณฑ์ยุติการทดสอบ

การพัฒนาโปรแกรมคอมพิวเตอร์

สายชล อบทม (2539) ได้พัฒนาโปรแกรมสำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ที่ใช้รูปแบบทางแยกแปรผัน การประมาณค่าความสามารถของผู้สอบโดยวิธีของเบส์ ยุติการทดสอบเมื่อความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่ำกว่า 0.3 ประกอบด้วยโปรแกรมย่อย 2 โปรแกรม คือ โปรแกรมย่อยการพิมพ์ข้อสอบและโปรแกรมย่อยการทดสอบ ที่ปฏิบัติการบนระบบปฏิบัติการ Windows 95 แต่โสพส สุขานนท์สวัสดิ์ (2545) ยุติการทดสอบเมื่อความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่ำกว่า 0.45 หรือกดปุ่มยุติการทดสอบเมื่อต้องการยุติการทดสอบ ซึ่งพัฒนาให้สามารถใช้ได้บนระบบปฏิบัติการ Windows 98 ขึ้นไป นอกจากนี้เกียรติศักดิ์ ส่องแสง (2547) ได้พัฒนาโปรแกรมสำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ที่สามารถทำงานบนระบบปฏิบัติการ Windows 98 ขึ้นไป โปรแกรมสามารถจัดชุดข้อสอบให้ตอบในแต่ละขั้นได้ 3 แบบ (ชุดข้อสอบให้ตอบขั้นละ 1 ข้อ, 2 ข้อ, 3 ข้อ) กำหนดเวลาในการตอบแต่ละข้อได้ 3 แบบ (ไม่กำหนดเวลา, กำหนดเวลาเฉลี่ยเท่ากันทุกข้อ, กำหนดเวลาตามน้ำหนักค่าความยากของข้อสอบ) และจัดรูปแบบของข้อสอบได้ 2 รูปแบบ (ไม่มีสื่อประสม, มีสื่อประสม) โปรแกรมเริ่มต้นการทดสอบด้วยข้อสอบที่มีความยากง่ายปานกลาง ประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์ คัดเลือกข้อสอบข้อ

ต่อไปจากค่าฟังก์ชันสารสนเทศของข้อสอบที่สูงที่สุด ยุติการทดสอบเมื่อความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบเท่ากับหรือต่ำกว่า 0.3

จากผลงานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ พบว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบมีคุณภาพไม่แตกต่างกับการทดสอบแบบประเพณีนิยม แต่ใช้จำนวนข้อสอบและเวลาในการทดสอบน้อยกว่าซึ่งทำให้ผู้เข้าสอบไม่เบื่อหน่ายในการสอบ และเป็นการประหยัดทรัพยากรที่ใช้ในการทดสอบ จะเห็นได้ว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มีประโยชน์ในหลาย ๆ ด้าน และมีความเป็นไปได้ที่จะนำเอาการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (CAT) มาใช้แทนการทดสอบแบบประเพณีนิยม ซึ่งในปัจจุบันได้มีการนำเทคโนโลยีมาใช้ในวงการศึกษากันมากขึ้น เช่น การนำคอมพิวเตอร์เข้ามาช่วยในการจัดการเรียนการสอน และการนำคอมพิวเตอร์มาใช้ในการทดสอบซึ่งมีแนวโน้มจะได้รับความนิยมกันอย่างแพร่หลายในประเทศไทย เนื่องจากมีการพัฒนาโปรแกรมคอมพิวเตอร์เพื่อใช้ในการวิเคราะห์ข้อมูลและประมวลผล ทำให้ได้รับความสะดวกและรวดเร็วเป็นอย่างมาก โดยเฉพาะการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ซึ่งมีข้อดีหลายประการดังที่ได้กล่าวไว้ข้างต้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

วิธีการดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง เพื่อศึกษาผลที่เกิดขึ้นจากตัวแปรอิสระ 3 ตัวแปร คือ เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบ ที่มีต่อความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (CAT) โดยมีวิธีดำเนินการวิจัยดังนี้

1. ประชากรและกลุ่มตัวอย่างที่ใช้ในการวิจัย

ประชากร

ประชากรที่ใช้ในการวิจัยครั้งนี้ ได้แก่ นักเรียนชั้นประถมศึกษาปีที่ 6 ที่กำลังศึกษาในภาคเรียนที่ 2 ประจำปีการศึกษา 2549 ในโรงเรียนประถมศึกษาขนาดใหญ่พิเศษ สังกัดสำนักงานเขตพื้นที่การศึกษานครราชสีมา เขต 1 สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ จำนวน 4 โรงเรียน ซึ่งมีจำนวนนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 1,982 คน (ข้อมูล ณ วันที่ 10 มิถุนายน 2549 จาก http://www.edkorat1.info)

กลุ่มตัวอย่างและการสุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ ผู้วิจัยคำนวณขนาดของกลุ่มตัวอย่าง โดยใช้แนวคิดของ Hair (1998) ซึ่งเสนอว่า จำนวนหน่วยตัวอย่างที่นำมาศึกษาควรมีประมาณ 20 เท่าของตัวแปรอิสระ สำหรับในการวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง ซึ่งมีเงื่อนไขที่ใช้ในการทดลองจัดเป็นกลุ่มได้ 12 กลุ่ม นั่นหมายความว่า ขนาดของกลุ่มตัวอย่างที่นำมาใช้ในแต่ละกลุ่มควรมีอย่างน้อย 20 คน ดังนั้น ผู้วิจัยจึงคำนวณขนาดของกลุ่มตัวอย่างขั้นต่ำได้ 240 คน นอกจากนี้ในการตรวจสอบความตรงตามเกณฑ์สัมพันธ์ เพื่อให้สัมประสิทธิ์ความตรงที่ได้จากกลุ่มตัวอย่างมีความน่าเชื่อถือ จึงควรใช้กลุ่มตัวอย่างที่มีขนาดใหญ่ ถ้ากลุ่มตัวอย่างมีขนาด 200 คนขึ้นไปในแต่ละกลุ่มของการทดลอง ค่าสัมประสิทธิ์ความตรงที่ได้จากกลุ่มตัวอย่างสามารถสะท้อนค่าสัมประสิทธิ์ความตรงที่แท้จริงของประชากรได้ถูกต้องด้วยความน่าจะเป็นอย่างน้อยร้อยละ 90 ขึ้นไป (ศิริชัย กาญจนวาสี, 2548) และผู้วิจัยได้ดำเนินการสุ่มอย่างง่าย (Simple random sampling) โดยสุ่มโรงเรียนประถมศึกษาขนาดใหญ่พิเศษ 3 โรงเรียน แต่เพื่อให้งานวิจัยมีความแกร่ง (Robustness) ของข้อมูลและป้องกันความไม่สมบูรณ์ครบถ้วนของข้อมูล ประกอบกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบได้นำเอาทฤษฎีการตอบสนองข้อสอบ (IRT) มาประยุกต์ใช้ และจากการศึกษางานวิจัยที่เกี่ยวกับ

ทฤษฎีการตอบสนองข้อสอบทั้งการสร้างสถานการณ์จำลองข้อมูลและการเก็บรวบรวมข้อมูลจริงจากกลุ่มตัวอย่าง พบว่า งานวิจัยส่วนมากกำหนดขนาดของกลุ่มตัวอย่างโดยรวมประมาณ 1,000 คนขึ้นไปแล้วแต่เรื่องที่นำมาศึกษานั้น (ต่าย เซี่ยงฉี, 2534 ; รังสรรค์ มณีเล็ก, 2540 ; เกียรติศักดิ์ ส่องแสง, 2547 ; Fan, 1988) ดังนั้นผู้วิจัยจึงเพิ่มขนาดของกลุ่มตัวอย่างเป็น 1,498 คน และสามารถเก็บข้อมูลจริงได้ 924 คน คิดเป็นร้อยละ 61.68 ของจำนวนกลุ่มตัวอย่างที่ผู้วิจัยกำหนดไว้ ดังนั้นกลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2549 จำนวน 924 คน

นอกจากนี้ผู้วิจัยสุ่มโรงเรียนประถมศึกษาขนาดใหญ่พิเศษในจังหวัดนครราชสีมา เขต 1 จำนวน 3 โรงเรียน เนื่องจากเป็นโรงเรียนในสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานที่มีความพร้อมในด้านเครื่องคอมพิวเตอร์ที่เพียงพอสำหรับใช้ในการเก็บรวบรวมข้อมูล และบุคลากรในโรงเรียนยินดีให้ความร่วมมือในการทำวิจัย ซึ่งมีรายชื่อโรงเรียน จำนวนนักเรียนและจำนวนนักเรียนที่เป็นกลุ่มตัวอย่างที่สามารถเก็บข้อมูลได้จริง จำแนกเป็นรายโรงเรียนได้ ดังตารางที่ 2

ตารางที่ 2 จำนวนกลุ่มตัวอย่าง จำแนกเป็นรายโรงเรียน

ลำดับที่	โรงเรียน	จำนวนกลุ่มตัวอย่าง (คน)	จำนวนกลุ่มตัวอย่างที่สามารถเก็บข้อมูลได้จริง (คน)
1	วัดสระแก้ว	409	315
2	สุขานารี	508	408
3	อนุบาลนครราชสีมา	582	201
	รวม	1,498	924

2. เครื่องมือที่ใช้ในการวิจัย

2.1. เครื่องมือที่ใช้ในการวิจัยครั้งนี้ ประกอบด้วย

1. แบบสอบวัดความสามารถทางคณิตศาสตร์ ตามกลุ่มสาระการเรียนรู้คณิตศาสตร์ โดยวัดความเข้าใจทางคณิตศาสตร์สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 เป็นข้อสอบเลือกตอบ 4 ตัวเลือก จำนวน 40 ข้อ เพื่อตรวจสอบความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

2. คลังข้อสอบ เป็นกลุ่มข้อสอบที่ใช้วัดความเข้าใจในวิชาคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 6 ซึ่งพัฒนาโดยรังสรรค์ มณีเล็ก (2540) จำนวน 200 ข้อและเกียรติศักดิ์ ส่องแสง (2547) จำนวน 44 ข้อ รวมเป็น 244 ข้อ ลักษณะของข้อสอบเป็นแบบเลือกตอบ 4 ตัวเลือก มีค่าอำนาจจำแนกอยู่ระหว่าง 0.41 ถึง 2.34 ค่าความยากง่ายอยู่ในช่วง -2.23 ถึง 2.72 และค่าการเดาอยู่ในช่วง 0.06 ถึง 0.30

3. โปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ผู้วิจัยได้ใช้โปรแกรมการทดสอบแบบปรับเหมาะตามระดับความสามารถของผู้สอบ ซึ่งพัฒนาโดยเกียรติศักดิ์ ส่องแสง (2547) ซึ่งได้กำหนดรูปแบบทางแยกแปรผัน การประมาณค่าความสามารถของผู้สอบโดยวิธีของเบส์ โดยประกอบด้วย 2 โปรแกรมย่อย คือ โปรแกรมย่อยสำหรับการพิมพ์เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบ และ โปรแกรมย่อยสำหรับการทดสอบ ซึ่งพัฒนาให้สามารถใช้ได้บนระบบปฏิบัติการ Windows 98 ขึ้นไป และได้เพิ่มเงื่อนไขเพื่อให้สอดคล้องกับงานวิจัย มีทั้งหมด 12 เงื่อนไข ซึ่งได้จากตัวแปรอิสระ 3 ตัว คือ เกณฑ์การคัดเลือกข้อสอบขั้นแรก 2 วิธี (สารสนเทศของฟิชเชอร์ สารสนเทศของคูลเบค-ไลเบลอร์) อัตราการใช้ข้อสอบซ้ำ 3 ระดับ (10%, 20%, 30%) และ เกณฑ์ยุติการทดสอบ 2 เกณฑ์ (SEE $\leq$ 0.30, SEE $\leq$0.45) เงื่อนไขของการทดสอบทั้ง 12 เงื่อนไข ดังตารางที่ 3

ตารางที่ 3 เงื่อนไขในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

เงื่อนไข	เกณฑ์การคัดเลือกข้อสอบขั้นแรก	อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)	เกณฑ์ยุติการทดสอบ
1	สารสนเทศของฟิชเชอร์	10	SEE$\leq$0.30
2	สารสนเทศของฟิชเชอร์	10	SEE$\leq$0.45
3	สารสนเทศของฟิชเชอร์	20	SEE$\leq$0.30
4	สารสนเทศของฟิชเชอร์	20	SEE$\leq$0.45
5	สารสนเทศของฟิชเชอร์	30	SEE$\leq$0.30
6	สารสนเทศของฟิชเชอร์	30	SEE$\leq$0.45
7	สารสนเทศของคูลเบค-ไลเบลอร์	10	SEE$\leq$0.30
8	สารสนเทศของคูลเบค-ไลเบลอร์	10	SEE$\leq$0.45
9	สารสนเทศของคูลเบค-ไลเบลอร์	20	SEE$\leq$0.30
10	สารสนเทศของคูลเบค-ไลเบลอร์	20	SEE$\leq$0.45
11	สารสนเทศของคูลเบค-ไลเบลอร์	30	SEE$\leq$0.30
12	สารสนเทศของคูลเบค-ไลเบลอร์	30	SEE$\leq$0.45

2.2. วิธีดำเนินการพัฒนาและตรวจสอบคุณภาพแบบสอบวัดความสามารถทางคณิตศาสตร์

แบบวัดที่สร้างขึ้นเป็นแบบสอบเลือกตอบ (Multiple choices) 4 ตัวเลือก ในสาระการเรียนรู้คณิตศาสตร์พื้นฐาน ตามหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 โดยมีขั้นตอน ดังนี้

1. ศึกษาหลักสูตร สาระการเรียนรู้คณิตศาสตร์พื้นฐาน มาตรฐานการเรียนรู้ช่วงชั้น และผลการเรียนรู้ที่คาดหวัง ชั้นประถมศึกษาปีที่ 6 เรื่อง จำนวนนับและการบวก การลบ การคูณ การหารจำนวนนับ สมการและการแก้สมการ ตัวประกอบของจำนวนนับ มุมและส่วนของเส้นตรง เส้นขนาน ทิศและแผนผัง เศษส่วนและการบวก การลบ การคูณ การหารเศษส่วน จากแผนการจัดการเรียนรู้ และเอกสารที่เกี่ยวข้อง

2. ศึกษาเนื้อหาและผลการเรียนรู้ที่คาดหวัง จากหนังสือเรียนสาระการเรียนรู้พื้นฐานคณิตศาสตร์ กลุ่มสาระการเรียนรู้คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 และ หนังสือการจัดสาระการเรียนรู้ กลุ่มสาระการเรียนรู้คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 1- 6 ตามหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 ของกรมวิชาการ กระทรวงศึกษาธิการ เรื่อง จำนวนนับและการบวก การลบ การคูณ การหารจำนวนนับ สมการและการแก้สมการ ตัวประกอบของจำนวนนับ มุมและส่วนของเส้นตรง เส้นขนาน ทิศและแผนผัง เศษส่วนและการบวก การลบ การคูณ การหารเศษส่วน

3. กำหนดลักษณะของเครื่องมือวัด ซึ่งผู้วิจัยเลือกใช้รูปแบบของข้อสอบแบบเลือกตอบ 4 ตัวเลือก (Multiple choice)

4. สร้างตารางวิเคราะห์ข้อสอบ โดยให้น้ำหนักจำนวนข้อตามความสำคัญของเนื้อหา ปริมาณเนื้อหาและจำนวนชั่วโมงที่ทำการสอน ดังตารางต่อไปนี้

ตารางที่ 4 ตารางวิเคราะห์ข้อสอบ โดยให้น้ำหนักจำนวนข้อตามความสำคัญของเนื้อหา ปริมาณเนื้อหาและจำนวนชั่วโมงที่ทำการสอน

	เนื้อหา	จำนวนชั่วโมงที่ทำการสอน (ชั่วโมง)	น้ำหนักรายข้อ		
			คิดเป็น %	จำนวนข้อ	ข้อที่
1.	จำนวนนับและการบวก การลบ การคูณ การหารจำนวนนับ	15	17.24	24	1 - 2, 20 - 27, 47 - 58, 115 - 116
2.	สมการและการแก้สมการ	15	17.24	24	117 - 140
3.	ตัวประกอบของจำนวนนับ	18	20.69	29	59 - 87
4.	มุมและส่วนของเส้นตรง	5	5.75	8	99 – 104, 112 - 113
5.	เส้นขนาน	5	5.75	8	105 – 111, 114
6.	ทิศและแผนผัง	7	8.05	11	88 - 98
7.	เศษส่วนและการบวก การลบ การคูณ การหารเศษส่วน	22	25.29	36	3 – 19, 8 - 46
	รวม	87	100	140	

5. ศึกษาข้อสอบคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 6 จากหลายแหล่ง เช่น จากข้อสอบเข้าโรงเรียนต่าง ๆ จากแบบฝึกหัด จากวิทยานิพนธ์ เพื่อนำมาเป็นแนวทางในการพัฒนาข้อสอบ

6. พัฒนาข้อสอบวัดความสามารถทางคณิตศาสตร์ โดยวัดความเข้าใจเกี่ยวกับคณิตศาสตร์สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 เรื่อง จำนวนนับและการบวก การลบ การคูณ การหารจำนวนนับ สมการและการแก้สมการ ตัวประกอบของจำนวนนับ มุมและส่วนของเส้นตรง เส้นขนาน ทิศและแผนผัง เศษส่วนและการบวก การลบ การคูณ การหารเศษส่วน ได้ 140 ข้อ ซึ่งจัดได้ตามสาระและมาตรฐานการเรียนรู้ กลุ่มสาระการเรียนรู้คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 หลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 ดังตารางต่อไปนี้

ตารางที่ 5 จำนวนข้อสอบ ตามสาระและมาตรฐานการเรียนรู้ กลุ่มสาระการเรียนรู้คณิตศาสตร์

สาระหลัก	มาตรฐานการเรียนรู้	จำนวนข้อ
สาระที่ 1 : จำนวนและการดำเนินการ	มาตรฐาน ค 1.1 : เข้าใจความหลากหลายของการแสดงจำนวนและการใช้จำนวนในชีวิตจริง **(ข้อที่ 1 - 19)**	19
	มาตรฐาน ค 1.2 : เข้าใจถึงผลที่เกิดขึ้นจากการดำเนินการของจำนวนและความสัมพันธ์ระหว่างการดำเนินการต่าง ๆ และสามารถใช้การดำเนินการในการแก้ปัญหาได้ **(ข้อที่ 20 - 46)**	27
	มาตรฐาน ค 1.3 : ใช้การประมาณค่าในการคำนวณและแก้ปัญหาได้ **(ข้อที่ 47 - 49)**	3
	มาตรฐาน ค 1.4 : เข้าใจระบบจำนวนและสามารถนำสมบัติเกี่ยวกับจำนวนไปใช้ได้ **(ข้อที่ 50 - 87)**	38
สาระที่ 2 : การวัด	มาตรฐาน ค 2.1 : เข้าใจพื้นฐานเกี่ยวกับการวัด **(ข้อที่ 88 - 98)**	11
สาระที่ 3 : เรขาคณิต	มาตรฐาน ค 3.1 : อธิบายและวิเคราะห์รูปเรขาคณิตสองมิติและสามมิติได้ **(ข้อที่ 99 - 111)**	13
	มาตรฐาน ค 3.2 : ใช้การนึกภาพ ใช้เหตุผลเกี่ยวกับปริภูมิ และใช้แบบจำลองทางเรขาคณิตในการแก้ปัญหาได้ **(ข้อที่ 112 - 114)**	3
สาระที่ 4 : พีชคณิต	มาตรฐาน ค 4.1 : อธิบายและวิเคราะห์แบบรูป ความสัมพันธ์ และฟังก์ชันต่าง ๆ ได้ **(ข้อที่ 115 - 116)**	2
	มาตรฐาน ค 4.2 : ใช้นิพจน์ สมการ อสมการ กราฟ และแบบจำลองทางคณิตศาสตร์อื่น ๆ แทนสถานการณ์ต่าง ๆ ตลอดจนแปลความหมายและนำไปใช้แก้ปัญหาได้ **(ข้อที่ 117 - 140)**	24
	รวม	140

7. การวิเคราะห์คุณภาพของแบบวัดความสามารถทางคณิตศาสตร์

7.1. นำแบบสอบวัดความสามารถทางคณิตศาสตร์ที่สร้างขึ้นไปให้อาจารย์ที่ปรึกษาวิทยานิพนธ์ตรวจสอบความเหมาะสม

7.2. นำแบบสอบวัดความสามารถทางคณิตศาสตร์มาปรับปรุงตามคำแนะนำของอาจารย์ที่ปรึกษาวิทยานิพนธ์ เกี่ยวกับโครงสร้างภาษาที่ใช้เขียนเป็นข้อคำถาม

7.3. นำแบบสอบวัดความสามารถทางคณิตศาสตร์ฉบับปรับปรุงแล้วไปให้ผู้เชี่ยวชาญ 12 ท่าน พิจารณาความสอดคล้องของแบบวัดกับสาระการเรียนรู้และผลการเรียนรู้ที่คาดหวัง เพื่อตรวจสอบความตรงเชิงเนื้อหา (Content validity) โดยหาค่าดัชนีความสอดคล้องของข้อสอบ (Item Objective Congruence; IOC) ข้อสอบที่ใช้ได้มีค่า IOC ร้อยละ 80 (IOC $\geq$ 0.80) ซึ่งแสดงว่าข้อคำถามนั้นสอดคล้องกับจุดมุ่งหมาย/ เนื้อหาที่มุ่งวัด (ศิริชัย กาญจนวาสี, 2544) โดยกำหนดคุณสมบัติของผู้เชี่ยวชาญไว้ ดังนี้ (รายชื่อผู้เชี่ยวชาญแสดงในภาคผนวก ก)

1) ทำงานเกี่ยวข้องกับการเรียนการสอนคณิตศาสตร์ ไม่น้อยกว่า 10 ปี หรือ
2) มีคุณวุฒิทางการศึกษาไม่ต่ำกว่าระดับปริญญาโทและทำงานเกี่ยวข้องกับการเรียนการสอนคณิตศาสตร์ ไม่น้อยกว่า 2 ปี หรือ
3) มีความรู้และประสบการณ์ทางด้านการวัดและประเมินผลการศึกษา

7.4. นำผลการตัดสินของผู้เชี่ยวชาญมาคำนวณค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับลักษณะที่มุ่งวัด (Item Objective Congruence: IOC) เป็นรายข้อ ซึ่งมีค่า +1, 0, -1 พร้อมทั้งนำมาปรับปรุงตามคำแนะนำของผู้เชี่ยวชาญ จำนวน 12 ข้อ ตัดทิ้ง 60 ข้อ ทำให้เหลือข้อสอบจำนวน 80 ข้อรายละเอียดการปรับปรุง แก้ไข แสดงในภาคผนวก ข

7.5. นำแบบสอบวัดความสามารถทางคณิตศาสตร์ที่ปรับปรุงแล้วทั้ง 80 ข้อ ไปทดลองใช้กับนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนเมืองนครราชสีมา จำนวน 428 คน ซึ่งมีลักษณะใกล้เคียงกับกลุ่มตัวอย่าง

7.6. นำผลการทดสอบมาวิเคราะห์หาคุณภาพแบบสอบโดยใช้โปรแกรม Loma IAS Version 11.0 ได้ค่าอำนาจจำแนกเฉลี่ย เท่ากับ 0.610 ค่าความยากเฉลี่ย เท่ากับ 1.346 ค่าการเดาเฉลี่ย เท่ากับ 0.207 และค่าฟังก์ชันสารสนเทศของแบบสอบ เท่ากับ 2.630 และได้หาค่าความเที่ยงของแบบสอบวัดความสามารถทางคณิตศาสตร์ทั้งฉบับ 80 ข้อ (KR 20) เท่ากับ 0.925 รวมทั้งจำแนกความเที่ยงเป็นรายเนื้อหาโดยใช้โปรแกรม TAP (Test Analysis Program) version 6.63 ได้ดังตารางที่ 6

ตารางที่ 6 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และค่าความเที่ยงของแบบสอบวัดความสามารถทางคณิตศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 6 จำแนกตามความเที่ยงเป็นรายเนื้อหา

	เนื้อหา	จำนวนข้อสอบ	*M*	*SD*	ค่าความเที่ยง KR 20
1.	จำนวนนับและการบวก การลบ การคูณ การหารจำนวนนับ	14	9.374	2.872	0.725
2.	สมการและการแก้สมการ	16	7.755	2.993	0.639
3.	ตัวประกอบของจำนวนนับ	16	7.661	3.417	0.741
4.	มุมและส่วนของเส้นตรง	4	1.914	1.219	0.541
5.	เส้นขนาน	4	2.093	1.217	0.460
6.	ทิศและแผนผัง	6	3.806	1.597	0.581
7.	เศษส่วนและการบวก การลบ การคูณ การหารเศษส่วน	20	9.012	4.010	0.773
	รวม	80	41.614	14.115	0.925

7.7. ตรวจสอบคุณภาพข้อสอบเป็นรายข้อจากโปรแกรมวิเคราะห์ข้อสอบ Loma IAS Version 11.0 และพิจารณาเกณฑ์การคัดเลือกข้อสอบ คือ ค่าอำนาจจำแนก (a) ค่าความยาก (b) ค่าการเดา (c) ตามโปรแกรมวิเคราะห์ข้อสอบ Loma IAS Version 11.0 ซึ่งพัฒนาโดยคณิต ไข่มุกด์ มีดังนี้ $0.5 \le a \le 2.5$, $-3.0 \le b \le 3.0$ และ $c \le 0.3$ จากนั้นคัดเลือกข้อสอบได้ข้อสอบทั้งสิ้นจำนวน 40 ข้อ มีค่าอำนาจจำแนกเฉลี่ย เท่ากับ 0.706 ค่าความยากเฉลี่ย เท่ากับ 1.358 และค่าการเดาเฉลี่ย เท่ากับ 0.201 และ ค่าความเที่ยง (KR20) เท่ากับ 0.871 โดยครอบคลุมเนื้อหาเรื่อง จำนวนนับและการบวก การลบ การคูณ การหารจำนวนนับ สมการและการแก้สมการ ตัวประกอบของจำนวนนับ มุมและส่วนของเส้นตรง เส้นขนาน ทิศและแผนผัง เศษส่วนและการบวก การลบ การคูณ การหารเศษส่วน ที่นักเรียนได้ศึกษาตามสาระและมาตรฐานการเรียนรู้ กลุ่มสาระการเรียนรู้คณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 หลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 รายละเอียดแบบสอบวัดความสามารถทางคณิตศาสตร์ แสดงในภาคผนวก ค

2.3. พิจารณาคุณภาพของคลังข้อสอบ

จากการศึกษางานวิจัยของรังสรรค์ มณีเล็ก (2540) และ เกียรติศักดิ์ ส่องแสง (2547) ซึ่งมีการสร้างและพัฒนาคลังข้อสอบอย่างเป็นระบบและมีคุณภาพ โดยพิจารณาจาก การตรวจสอบความตรงตามเนื้อหาและความตรงเชิงโครงสร้างจากผู้เชี่ยวชาญด้านการสอนคณิตศาสตร์และผู้เชี่ยวชาญด้านการวัดผล รวมทั้งได้มีการตรวจความเป็นเอกมิติ (Unidimensionality) ด้วยวิธีการ

วิเคราะห์องค์ประกอบ (Factor Analysis) เพื่อคำนวณค่าไอเกน (Eigen value) สำหรับศึกษาอัตราส่วนระหว่างค่าไอเกนขององค์ประกอบแรกกับองค์ประกอบถัดไป ถ้ามีอัตราส่วนที่สูงแสดงถึงเครื่องมือหรือแบบสอบวัดคุณลักษณะเด่นเดียว (Single dominant factor) ผลปรากฏว่าข้อสอบที่ถูกจัดเก็บเข้าคลังข้อสอบมีความเป็นเอกมิติ ซึ่งตรงตามข้อตกลงเบื้องของทฤษฎีการตอบสนองข้อสอบ ถึงแม้ว่าข้อสอบจำนวน 44 ข้อซึ่งพัฒนาโดยเกียรติศักดิ์ ส่องแสง (2547) ไม่ได้ปรับเทียบคะแนนให้ข้อสอบอยู่บนสเกลเดียวกัน แต่มีค่าพารามิเตอร์ของข้อสอบโดยเฉลี่ยใกล้เคียงกัน คือ ค่าอำนาจจำแนกประมาณ 1.0 ซึ่งอยู่ในเกณฑ์ดี ค่าความยากง่ายประมาณ 0.8 ซึ่งอยู่ในเกณฑ์ปานกลางค่อนไปทางยาก และค่าการเดาประมาณ 0.2 ซึ่งอยู่ในเกณฑ์ดี รวมทั้งข้อสอบมีความเป็นเอกมิติกับข้อสอบจำนวน 200 ข้อซึ่งพัฒนาโดยรังสรรค์ มณีเล็ก (2540) ผู้วิจัยมีความเห็นว่าคลังข้อสอบควรมีข้อสอบมาก ๆ เพื่อที่คัดเลือกข้อสอบให้ครอบคลุมระดับความสามารถของผู้สอบทุกระดับได้

คลังข้อสอบที่ผู้วิจัยใช้ในการศึกษาวิจัยในครั้งนี้ เป็นกลุ่มข้อสอบที่ใช้วัดความรู้ความเข้าใจในวิชาคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 6 จำนวนทั้งสิ้น 244 ข้อ ซึ่งมีค่าอำนาจจำแนกเฉลี่ย 1.006 ค่าความยากง่ายเฉลี่ย 0.783 และค่าการเดาเฉลี่ย 0.205 และมีโครงสร้างของคลังข้อสอบวิชาคณิตศาสตร์ เรื่อง เศษส่วน ระดับชั้นประถมศึกษาปีที่ 6 ดังตารางที่ 7

ตารางที่ 7 โครงสร้างของข้อสอบวัดความเข้าใจ เรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 จำแนกตามเนื้อหาย่อย

เนื้อหาย่อย	ข้อที่	จำนวนข้อ
1. เศษส่วนของจำนวนนับ	1, 2, 32, 33, 66, 115, 169, 237	8
2. การเปรียบเทียบเศษส่วน	3, 4, 5, 25, 26, 27, 34, 36, 67, 92, 93, 94, 116, 117, 143, 144, 170, 171, 201, 202, 203, 204, 213, 214, 215, 216, 217, 218, 219, 220, 223, 231, 235, 236	34
3. การบวกเศษส่วน	6, 7, 37, 38, 39, 68, 95, 118, 145, 146, 147, 172, 173, 174	14
4. การลบเศษส่วน	8, 9, 10, 40, 41, 42, 43, 69, 70, 96, 119, 120, 121, 175, 176	15
5. การคูณเศษส่วน	11, 20, 21, 71, 72, 73, 97, 122, 123, 148, 149, 150, 177, 178, 209, 212	16
6. การหารเศษส่วน	12, 44, 45, 74, 98, 124, 125, 151, 152, 153, 179, 180, 206, 211, 224	15

ตารางที่ 7 (ต่อ)

เนื้อหาย่อย	ข้อที่	จำนวนข้อ
7. การทำให้เศษเกินเป็นจำนวนคละ	13, 15, 28, 29, 30, 31, 46, 75, 76, 99, 126, 154, 156, 182, 183, 207, 208, 234	18
8. ชนิดของเศษส่วน	77, 100, 155, 184, 221, 222	6
9. การนำเศษส่วนไปใช้ในชีวิตประจำวัน	18, 19, 35, 47, 48, 49, 50, 78, 79, 80, 101, 102, 103, 104, 127, 128, 157, 158, 159, 228, 238, 239, 240, 241, 242, 243, 244	27
10. เศษซ้อน	22, 51, 53, 81, 82, 83, 105, 106, 107, 129, 130, 131, 160, 161, 189, 190, 197	18
11. เศษส่วนที่มีค่าเท่ากัน	14, 16, 23, 24, 54, 55, 56,57, 58, 59, 65, 84, 85, 86, 108, 109, 110, 111, 132, 133, 134, 136, 162, 163, 164, 166, 181, 185, 186, 187, 188, 191, 193, 195, 196, 205, 210, 225, 226227, 229, 230, 232	43
12. ความหมายของเศษส่วน	17, 91, 135, 137, 165, 167, 192, 194, 200	9
13. เศษส่วนอย่างต่ำ	60, 61, 87, 88, 112, 113, 138, 139, 168, 233	10
14. การบวก ลบ คูณ หารระคน	62, 63, 64, 89, 90, 114, 140, 142, 198, 199	11

2.4. การพัฒนาโปรแกรมสำหรับการทดสอบ

โปรแกรมสำหรับการทดสอบที่ใช้ในการวิจัยนี้ คือ โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ ซึ่งผู้วิจัยได้นำโปรแกรมที่พัฒนาโดยเกียรติศักดิ์ ส่องแสง (2547) มาปรับให้เข้ากับเงื่อนไขที่ใช้ในการวิจัยโดยเพิ่มตัวแปรอิสระ 3 ตัวแปร ได้แก่ เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบ

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ดำเนินการทดสอบโดยให้ผู้สอบนั่งสอบหน้าจอคอมพิวเตอร์และตอบข้อสอบครั้งละ 1 ข้อ โดยคอมพิวเตอร์สุ่มข้อสอบจากคลังขึ้นมาให้ผู้สอบได้ตอบ ข้อสอบที่ถูกเลือกขึ้นมานั้นจะมีความเหมาะสมกับความสามารถของผู้สอบแต่ละบุคคล ในการดำเนินการทดสอบลักษณะนี้จำเป็นต้องมีโปรแกรมสั่งให้คอมพิวเตอร์จัดกระทำกับข้อมูลในลักษณะต่าง ๆ ตามตัวแปรอิสระที่ต้องการศึกษา ซึ่งผู้วิจัยได้นำโปรแกรมทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ ซึ่งพัฒนาโดยเกียรติศักดิ์ ส่องแสง (2547) มาพัฒนาต่อโดยใช้โปรแกรม Borland Delphi 7 ในการพัฒนาโปรแกรมมีขั้นตอนดำเนินการดังนี้

การทดสอบในรอบที่ 1 เมื่อผู้สอบเริ่มทำข้อสอบข้อแรกตามเกณฑ์การคัดเลือกข้อสอบขั้นแรก เมื่อผู้สอบตอบข้อสอบแล้วโปรแกรมจะทำการประมาณค่าความสามารถ และค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า เพื่อคัดเลือกข้อสอบข้อต่อไปที่มีความเหมาะสมกับความสามารถของผู้สอบ กระบวนการนี้จะดำเนินต่อไปเรื่อยจนกระทั่งค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าน้อยกว่าหรือเท่ากับ 0.45 จากนั้นโปรแกรมจะทำการบันทึกข้อมูลที่เป็นสารสนเทศในการทำข้อสอบของผู้สอบ แล้วโปรแกรมก็ดำเนินการทดสอบต่อไปเรื่อยจนกระทั่งค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าน้อยกว่าหรือเท่ากับ 0.3 การทดสอบก็จะยุติ ดังแสดงในแผนภาพที่ 14 และ แผนภาพที่ 15

การทดสอบในรอบที่ 2 ซึ่งเป็นการทดสอบที่มีอัตราการใช้ข้อสอบซ้ำ เริ่มต้นทำการอ่านข้อมูลเดิม ได้แก่ ข้อ และจำนวนข้อของการทดสอบในรอบที่ 1 เพื่อทำการคำนวณจำนวนข้อที่จะใช้ซ้ำ (10%, 20%, 30%) โดยผู้วิจัยทำการเขียนโปรแกรมเพื่ออ่านข้อมูลในรอบที่ 1 ซึ่งเริ่มต้นการทดสอบโดยการคัดเลือกข้อสอบจากกลุ่มข้อสอบที่ผู้สอบคนนั้นได้สอบมาแล้วในรอบที่ 1 ตามจำนวนข้อที่ได้คำนวณไว้ ทำการประมาณค่าความสามารถและความคลาดเคลื่อนมาตรฐานในการประมาณค่า จากนั้นทำการคัดเลือกข้อสอบภายในคลังข้อสอบที่มีความเหมาะสมกับความสามารถของผู้สอบ กระบวนการนี้จะดำเนินต่อไปเรื่อยจนกระทั่งค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าน้อยกว่าหรือเท่ากับ 0.45 จากนั้นโปรแกรมจะทำการบันทึกข้อมูลที่เป็นสารสนเทศในการทำข้อสอบของผู้สอบ แล้วโปรแกรมก็ดำเนินการทดสอบต่อไปเรื่อยจนกระทั่งค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าน้อยกว่าหรือเท่ากับ 0.3 การทดสอบก็จะยุติ ดังแสดงในแผนภาพที่ 16 และ แผนภาพที่ 17

นอกจากนี้ผู้วิจัยได้นำโปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ไปให้ผู้เชี่ยวชาญทางด้านวัดและประเมินผลทางการศึกษาและผู้มีประสบการณ์ในการทำวิจัยเกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ หรือผู้มีความสามารถในการเขียนโปรแกรม ทำการตรวจสอบความถูกต้องของสูตรที่ใช้ในโปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์และตรวจสอบประสิทธิภาพการทำงานของโปรแกรม (รายชื่อผู้เชี่ยวชาญ แสดงในภาคผนวก ก)

รวมทั้งได้นำโปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ไปทดลองใช้กับนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนวัดโสมนัส จำนวน 30 คน เพื่อตรวจสอบการทำงานของโปรแกรมและข้อบกพร่องในขณะดำเนินการเก็บรวบรวมข้อมูล รวมทั้งเตรียมความพร้อมของผู้วิจัยก่อนเก็บข้อมูลจริง จากการสังเกตของผู้วิจัย พบว่า นักเรียนจำนวนหนึ่งที่ได้ทำข้อสอบจำนวนมาก (มากกว่า 40 ข้อ) ถามผู้วิจัยว่า

"คุณครูครับจะต้องทำข้อสอบอีกกี่ข้อ แล้วจะเลิกสอบอย่างไงครับ"

นักเรียนคนที่ 1

"ผมยังสอบไม่เสร็จเลย ทำไมคนนี้เขาสอบเสร็จแล้ว ผมจะต้องทำข้อสอบอีกกี่ข้อ"

นักเรียนคนที่ 2

"ข้อสอบมีทั้งหมดกี่ข้อ แล้วเมื่อไรจะเลิกสอบครับ"

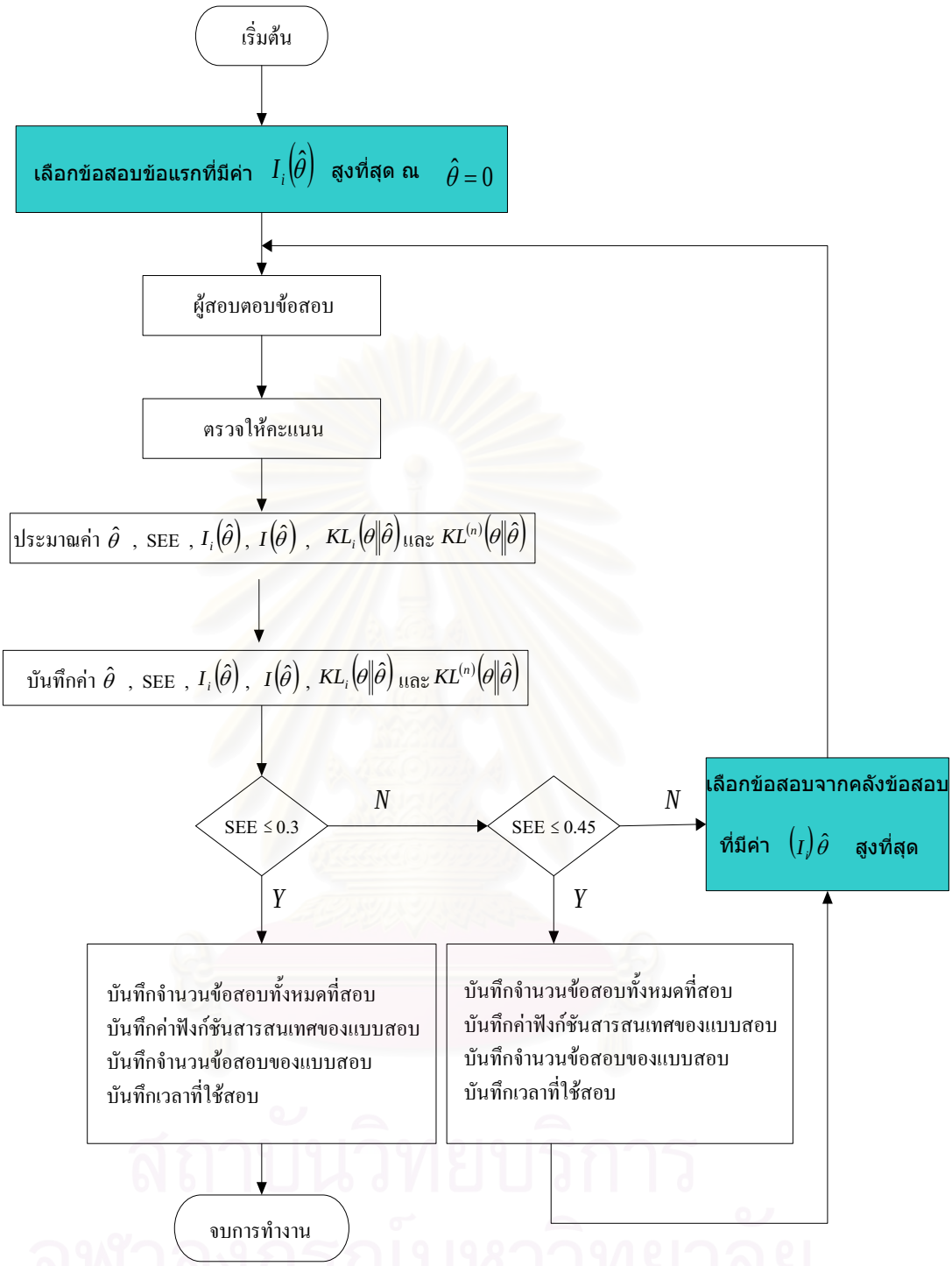
นักเรียนคนที่ 3

"ข้อสอบมีทั้งหมดกี่ข้อ ผมจะต้องทำอีกนานไหม"

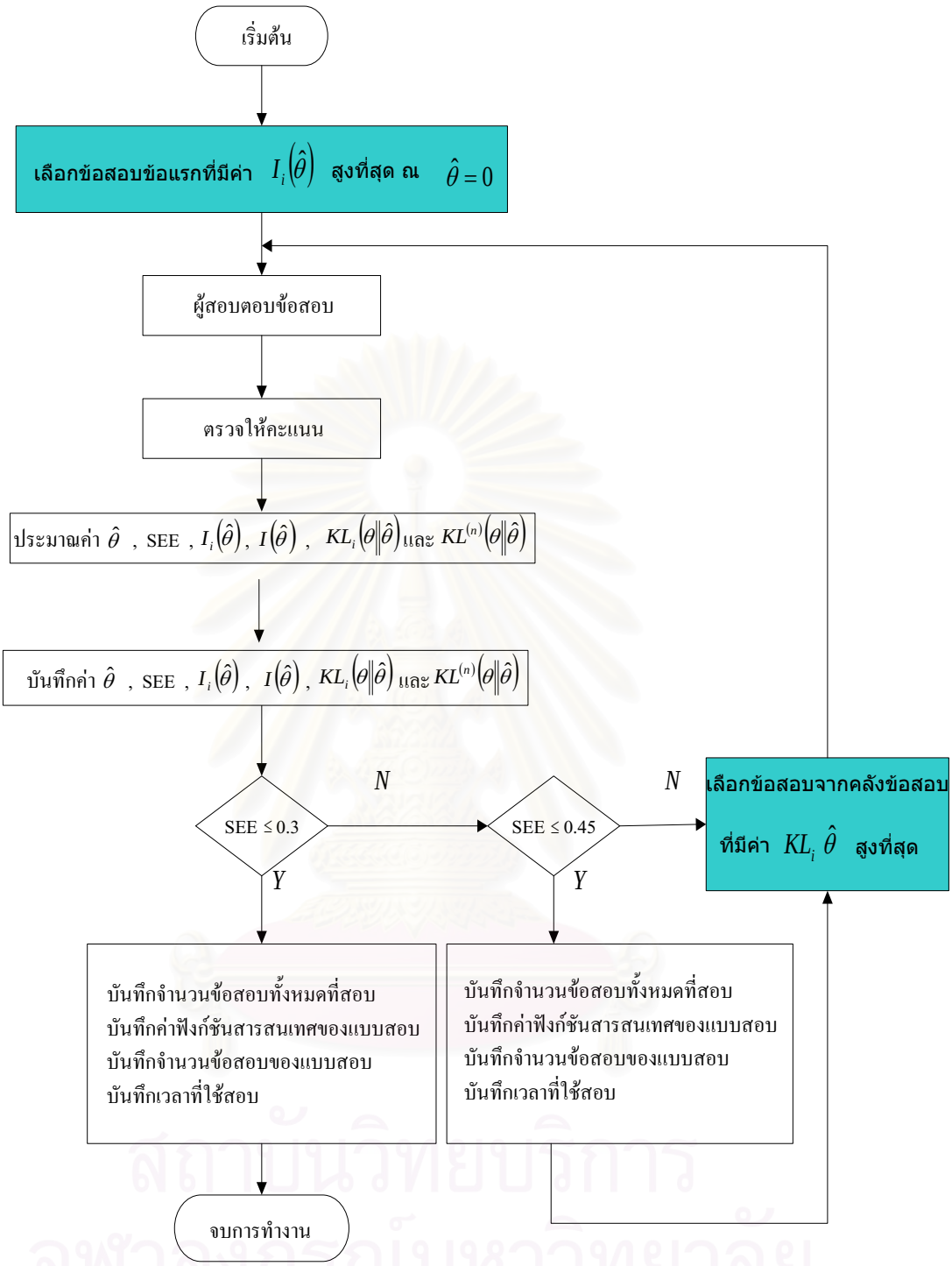
นักเรียนคนที่ 4

หลังจากนั้นนักเรียนเหล่านี้ก็ตอบข้อสอบอย่างรวดเร็วโดยไม่ได้อ่านข้อคำถาม นั่นหมายความว่านักเรียนเกิดความเบื่อหน่ายในการทำข้อสอบจำนวนมาก และจากการสอบถามอาจารย์ที่สอนวิชาคณิตศาสตร์ในโรงเรียน ทำให้ทราบว่า อาจารย์ในโรงเรียนส่วนมากออกข้อสอบวิชาคณิตศาสตร์ฉบับละประมาณ 40 ข้อ รวมทั้งแบบทดสอบวิชาคณิตศาสตร์ระดับเขตพื้นที่การศึกษามีข้อสอบฉบับละ 40 ข้อ และแบบสอบวัดผลสัมฤทธิ์ทางการเรียนในการประเมินคุณภาพการศึกษาระดับชาติ มีข้อสอบคณิตศาสตร์ฉบับละ 40 ข้อเช่นกัน ดังนั้นผู้วิจัยจึงได้กำหนดจำนวนข้อสอบที่ใช้ในโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ไว้สูงสุดจำนวน 40 ข้อ ถ้านักเรียนคนใดที่ต้องทำข้อสอบถึง 40 ข้อแล้วค่าความคลาดเคลื่อนมาตรฐานยังไม่น้อยกว่าหรือเท่ากับ 0.30 จะถูกคัดชื่อออกจากการประมาณค่าความสามารถของผู้สอบ เนื่องจากข้อสอบชุดนี้ไม่เหมาะกับความสามารถของผู้สอบดังกล่าว

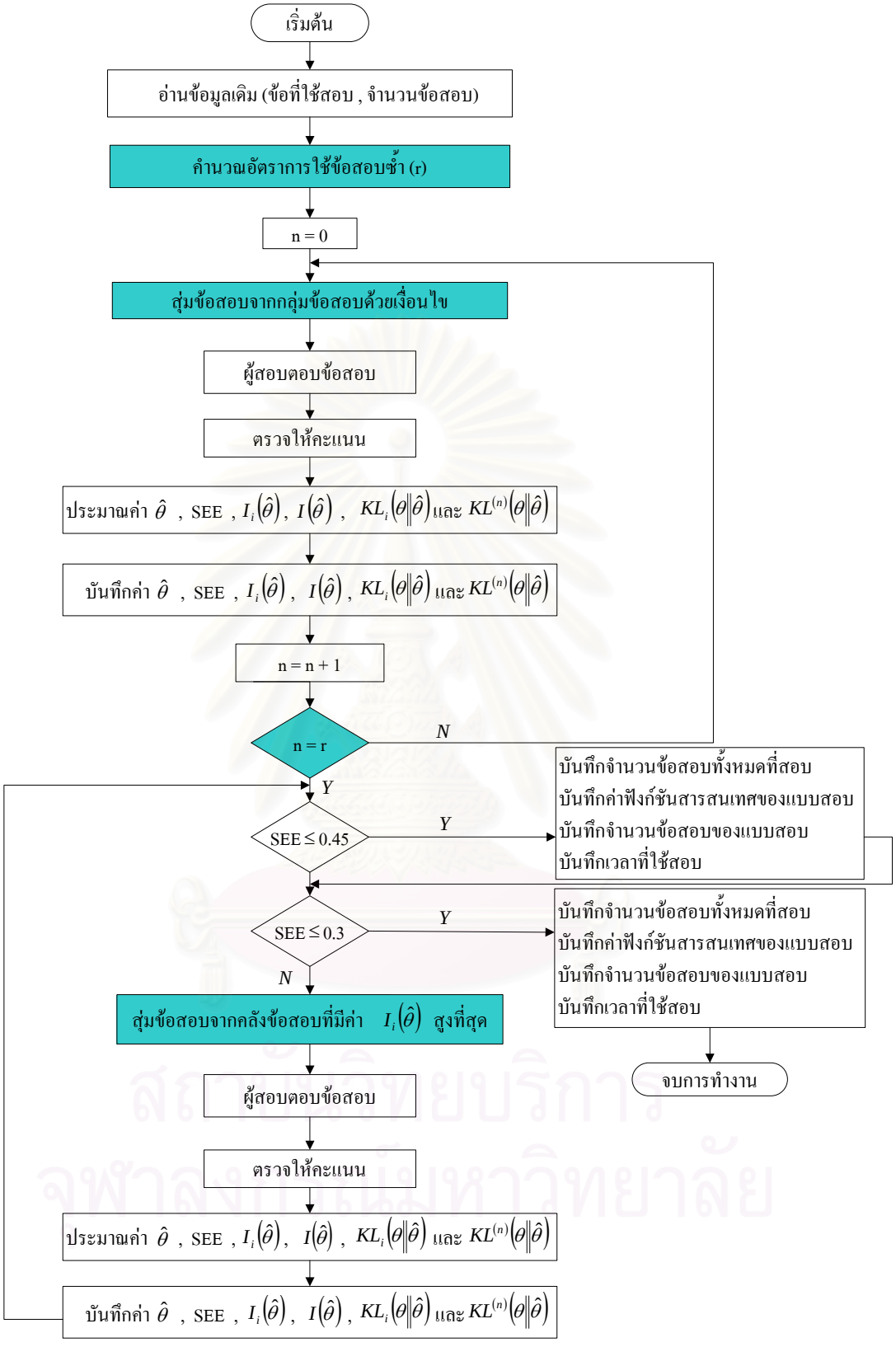
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



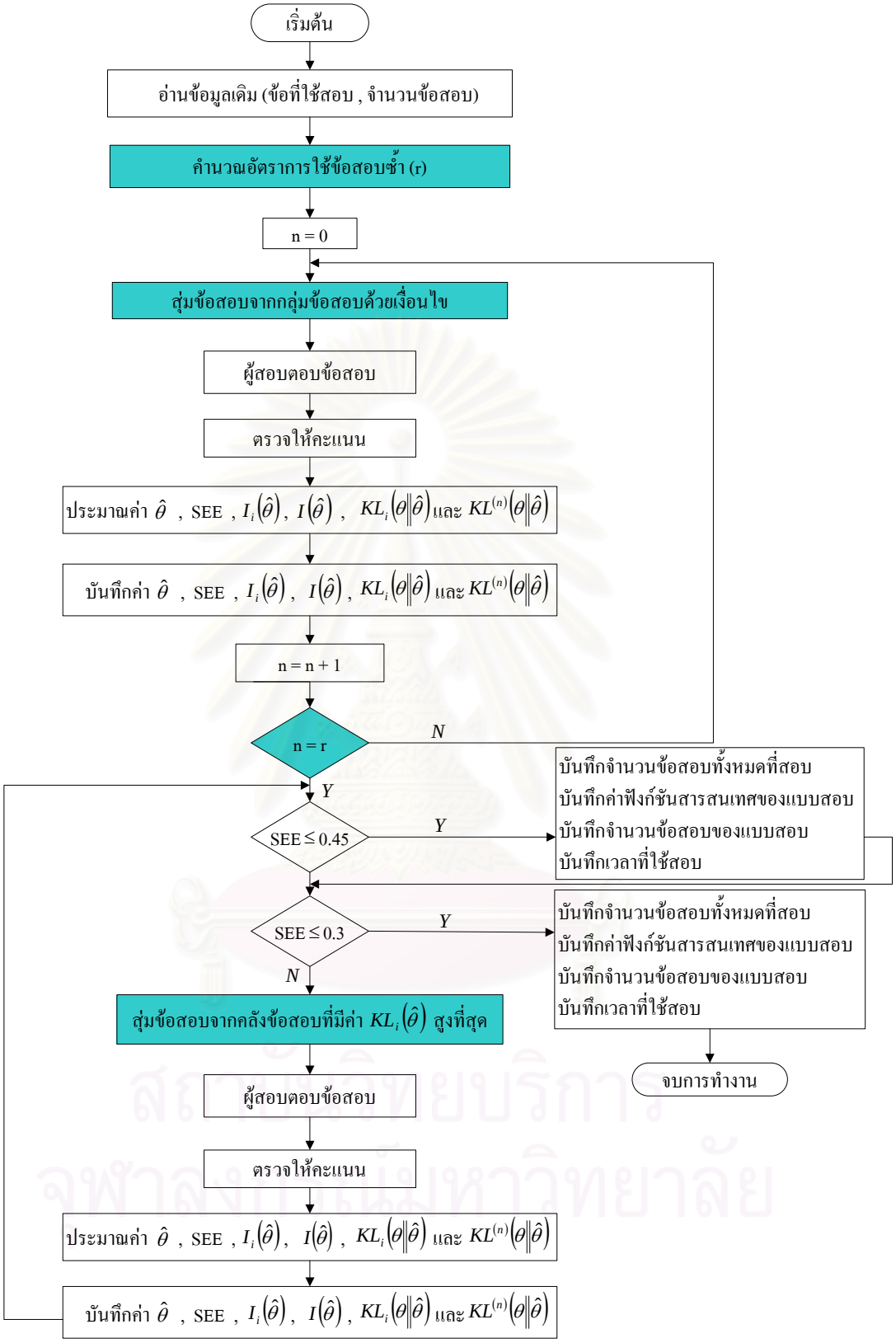
แผนภาพที่ 14 ผังแสดงขั้นตอนการทำงานของคอมพิวเตอร์ที่มีสารสนเทศของฟิชเชอร์เป็นเกณฑ์การคัดเลือกข้อสอบขั้นแรกในการทดสอบรอบที่ 1



แผนภาพที่ 15 ผังแสดงขั้นตอนการทำงานของคอมพิวเตอร์ที่มีสารสนเทศของคูลเบค - ไลเบลอร์เป็นเกณฑ์การคัดเลือกข้อสอบขั้นแรกในการทดสอบรอบที่ 1



แผนภาพที่ 16 ผังแสดงขั้นตอนการทำงานของคอมพิวเตอร์ที่มีสารสนเทศของฟิชเชอร์เป็นเกณฑ์การคัดเลือกข้อสอบขั้นแรกในการทดสอบรอบที่ 2



แผนภาพที่ 17 ผังแสดงขั้นตอนการทำงานของคอมพิวเตอร์ที่มีสารสนเทศของคูลเบค - ไลเบลอร์เป็นเกณฑ์การคัดเลือกข้อสอบขั้นแรกในการทดสอบรอบที่ 2

3. วิธีการเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลเพื่อศึกษาผลที่เกิดขึ้นต่อค่าความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์นั้น ได้แบ่งออกเป็น 2 ขั้นตอน ขั้นตอนแรก เป็นการทดสอบความสามารถทางคณิตศาสตร์ด้วยแบบสอบวัดความสามารถทางคณิตศาสตร์ ขั้นตอนที่สอง เป็นการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง

ในการดำเนินการทดสอบครั้งนี้ ผู้วิจัยไม่สามารถดำเนินการทดสอบด้วยตนเองทุกกลุ่มตัวอย่าง เนื่องด้วยเวลาและความสะดวกของโรงเรียนแตกต่างกัน จึงได้รับการอนุเคราะห์จากทางโรงเรียนอนุบาลนครราชสีมา ให้อาจารย์ปานทิพย์ พิมพ์เกาะ หัวหน้าสายชั้นประถมศึกษาปีที่ 6 เป็นผู้รับทำหน้าที่จัดดำเนินการทดสอบด้วยแบบสอบวัดความสามารถทางคณิตศาสตร์ และอาจารย์จรัส นวลดี อาจารย์สอนคอมพิวเตอร์ชั้นประถมศึกษาปีที่ 6 เป็นผู้รับทำหน้าที่จัดดำเนินการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้งให้ พร้อมทั้งให้ผู้วิจัยอธิบายวิธีดำเนินการทดสอบโดยใช้คอมพิวเตอร์ และลักษณะการทำงานของโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบให้อาจารย์จรัส นวลดีได้มีความเข้าใจตรงกับผู้วิจัย สำหรับโรงเรียนวัดสระแก้ว และ โรงเรียนสุขานารี ผู้วิจัยสามารถดำเนินการทดสอบด้วยตนเอง ซึ่งการเก็บรวบรวมข้อมูลในการวิจัยครั้งนี้ ดำเนินการระหว่างเดือนกันยายน 2549 ถึงเดือนมกราคม 2550 โดยมีขั้นตอนการเก็บรวบรวมข้อมูลดังนี้

1. ทำหนังสือขอความร่วมมือในการเก็บรวบรวมข้อมูล จากภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ถึงผู้บริหารสถานศึกษา เพื่อขอความช่วยเหลือและความสะดวกในการเก็บรวบรวมข้อมูลจากโรงเรียนที่เป็นกลุ่มตัวอย่าง

2. ทำหนังสือขอความร่วมมือในการเก็บรวบรวมข้อมูล จากและภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ถึงผู้อำนวยการสำนักงานเขตพื้นที่การศึกษานครราชสีมา เขต 1 เพื่อขอความร่วมมือไปยังโรงเรียนในสังกัดที่เป็นกลุ่มตัวอย่าง

3. นำหนังสือขอความร่วมมือในการเก็บรวบรวมข้อมูลจากภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยและผู้อำนวยการสำนักงานเขตพื้นที่การศึกษานครราชสีมา เขต 1 ไปติดต่อโรงเรียนที่ใช้เป็นกลุ่มตัวอย่าง เพื่อขอความร่วมมือในการทดสอบ กำหนดวัน เวลา และชั้นที่จะทำการทดสอบ

4. จัดเตรียมแบบสอบวัดความสามารถทางคณิตศาสตร์ โดยจัดแบบสอบให้มีจำนวนเพียงพอกับจำนวนนักเรียนในแต่ละห้องของโรงเรียนที่เป็นกลุ่มตัวอย่าง แล้วบรรจุแบบสอบไว้ในซองเอกสารสีน้ำตาล แยกเป็นห้อง ๆ เพื่อความสะดวกในการเก็บรวบรวมข้อมูล

5. นำแบบสอบวัดความสามารถทางคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 มาทำการทดสอบนักเรียนที่เป็นกลุ่มตัวอย่าง โดยมีคุณครูประจำชั้นเป็นผู้คุมสอบและผู้วิจัยสำรวจความเรียบร้อยในขณะการดำเนินการทดสอบ ซึ่งมีขั้นตอนในการดำเนินการสอบดังนี้

5.1. คุณครูประจำชั้นให้คำชี้แจงเกี่ยวกับการทดสอบว่าเป็นข้อสอบวิชาคณิตศาสตร์ให้เลือกตอบ 40 ข้อ ให้ระยะเวลาในการทำข้อสอบ 45 นาที

5.2. นักเรียนลงมือทำข้อสอบ คุณครูประจำชั้นนั่งคุมสอบอยู่ในห้อง

5.3. เมื่อนักเรียนสอบเสร็จ คุณครูรวบรวมแบบทดสอบและกระดาษคำตอบ และจัดเรียงตามเลขที่และส่งให้ผู้วิจัย

6. นำผลการตอบแบบสอบวัดความสามารถทางคณิตศาสตร์มาตรวจให้คะแนน และทำการประมาณค่าความสามารถของผู้สอบด้วยโปรแกรม Loma IAS Version 11.0 ซึ่งพัฒนาโดย คณิต ไข่มุกด์

7. เก็บคะแนนค่าความสามารถของผู้สอบในข้อ 6 ไว้เพื่อนำไปใช้เป็นคะแนนเกณฑ์ในการศึกษาความตรงตามสภาพจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์

8. จัดเตรียมแผ่นโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบซึ่งมี 6 โปรแกรม 12 เงื่อนไขที่ใช้ในการทำวิจัยนี้ให้พร้อมสำหรับติดตั้งโปรแกรมที่เครื่องคอมพิวเตอร์ของโรงเรียนที่เป็นกลุ่มตัวอย่าง

9. นำโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบทั้ง 6 โปรแกรม 12 เงื่อนไขที่ใช้ในการทำวิจัยนี้มาติดตั้งโปรแกรมที่เครื่องคอมพิวเตอร์ ซึ่งดำเนินการสุ่มอย่างง่ายโดยสุ่มเงื่อนไขทั้ง 12 เงื่อนไขให้กับคอมพิวเตอร์แต่ละเครื่อง ผู้วิจัยดำเนินการทดสอบด้วยคอมพิวเตอร์ ซึ่งมีขั้นตอนในการดำเนินการสอบดังนี้

9.1. ผู้วิจัยชี้แจงเกี่ยวกับวิธีการตอบข้อสอบให้ผู้สอบจนเป็นที่เข้าใจแล้วให้ผู้สอบเริ่มสอบ

9.2. นักเรียนลงมือทำข้อสอบ โดยมีผู้วิจัยและผู้ช่วยในการเก็บรวบรวมข้อมูลดูแลความเรียบร้อยในห้องสอบ

9.3. เมื่อนักเรียนสอบเสร็จ ผลการตอบของผู้สอบแต่ละคนจะถูกบันทึกไว้ใน Hard Disk ของเครื่องคอมพิวเตอร์

9.4. ให้นักเรียนทำแบบประเมินความคิดเห็นและความพึงพอใจในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

10. ผู้วิจัยดำเนินการ Write ผลการตอบของผู้สอบจาก Hard Disk ให้บรรจุไว้ในแผ่น CD

11. ดำเนินการทดสอบจำนวน 2 ครั้ง โดยเว้นระยะห่างในการทดสอบแต่ละครั้งประมาณ 2 สัปดาห์

12. ตรวจสอบความสมบูรณ์ครบถ้วนของข้อมูล จากการดำเนินการสอบโดยใช้แบบสอบวัดความสามารถทางคณิตศาสตร์ จำนวน 1 ครั้งและการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ จำนวน 2 ครั้ง ซึ่งในการวิจัยครั้งนี้มีการทดสอบทั้งหมด จำนวน 3 ครั้ง กับนักเรียนที่เป็นกลุ่มตัวอย่างจำนวน 1,498 คน หลังจากตรวจสอบความครบถ้วนของข้อมูล ปรากฏว่า มีจำนวนนักเรียนที่เป็นกลุ่มตัวอย่างมีข้อมูลของการทดสอบไม่ครบทั้ง 3 ครั้ง รวมทั้งนักเรียนที่ถูกคัดชื่อออกจากจากการประมาณค่าความสามารถของผู้สอบ เนื่องจากข้อสอบชุดนี้ไม่เหมาะกับความสามารถของผู้สอบ จึงตัดออกไม่นำมาวิเคราะห์จำนวน 574 คน คิดเป็นร้อยละ 38.32 ของจำนวนกลุ่มตัวอย่าง จึงเหลือกลุ่มตัวอย่างจำนวน 924 คน คิดเป็นร้อยละ 61.68 ของจำนวนกลุ่มตัวอย่างทั้งหมด

13. นำข้อมูลของผู้สอบแต่ละคนที่บันทึกไว้ใน CD คือ คะแนนความสามารถ มาวิเคราะห์ทางสถิติ เพื่อศึกษาผลของตัวแปรอิสระและตัวแปรตาม ซึ่งได้จำนวนข้อมูลจากกลุ่มตัวอย่างที่สามารถนำมาวิเคราะห์ข้อมูลกับเงื่อนไขทั้ง 12 เงื่อนไข ดังตารางที่ 8 (การวิจัยในครั้งนี้ได้ออกแบบโปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ที่สามารถให้ผลการทดสอบได้ดังนี้ ผู้สอบ 1 คน จะได้ผลการทดสอบคือคะแนนความสามารถทั้ง 2 เงื่อนไขในการทดสอบ 1 ครั้ง)

ตารางที่ 8 เงื่อนไขกับจำนวนข้อมูลที่ได้จากกลุ่มตัวอย่างซึ่งนำมาวิเคราะห์ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

แผ่นที่	เงื่อนไข	เกณฑ์การคัดเลือกข้อสอบขั้นแรก	อัตราการใช้ข้อสอบซ้ำ (%)	เกณฑ์ยุติการทดสอบ (SEE ≤)	จำนวนกลุ่มตัวอย่าง	จำนวนข้อมูลจากกลุ่มตัวอย่างที่นำมาวิเคราะห์
1	1	สารสนเทศของฟิชเชอร์	10	0.30	142	142
	2	สารสนเทศของฟิชเชอร์	10	0.45		142
2	3	สารสนเทศของฟิชเชอร์	20	0.30	138	138
	4	สารสนเทศของฟิชเชอร์	20	0.45		138
3	5	สารสนเทศของฟิชเชอร์	30	0.30	150	150
	6	สารสนเทศของฟิชเชอร์	30	0.45		150
4	7	สารสนเทศของคูลเบค-ไลเบลอร์	10	0.30	194	194
	8	สารสนเทศของคูลเบค-ไลเบลอร์	10	0.45		194
5	9	สารสนเทศของคูลเบค-ไลเบลอร์	20	0.30	154	154
	10	สารสนเทศของคูลเบค-ไลเบลอร์	20	0.45		154
6	11	สารสนเทศของคูลเบค-ไลเบลอร์	30	0.30	146	146
	12	สารสนเทศของคูลเบค-ไลเบลอร์	30	0.45		146
รวม					924	1,848

4. การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลสำหรับการวิจัยครั้งนี้แบ่งการวิเคราะห์ออกเป็น 2 ตอน ซึ่งมีแนวทางในการวิเคราะห์ข้อมูลและสถิติที่ใช้ ดังนี้

ตอนที่ 1 วิเคราะห์ข้อมูลเบื้องต้นด้วยค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย (Mean) และส่วนเบี่ยงเบนมาตรฐาน (Standard deviation) ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำนวน 2 ครั้ง โดยใช้โปรแกรม SPSS for Windows version 15.0

ตอนที่ 2 วิเคราะห์ค่าของตัวแปรตาม คือ ค่าความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ เพื่อศึกษาผลที่เกิดขึ้นจากตัวแปรอิสระ ได้แก่ เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบ โดยมีแนวทางในการดำเนินการดังนี้

1) หาค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จากสูตร (Glass and Stanley, 1970; ศิริชัย กาญจนวาสี, 2548)

$$r_{xy} = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left(n\sum x^2 - \left(\sum x\right)^2\right)\left(n\sum y^2 - \left(\sum y^2\right)\right)}}$$

เมื่อ r_{xy} แทน ความสัมพันธ์ระหว่างค่าความสามารถจากแบบสอบวัดความสามารถทางคณิตศาสตร์กับค่าความสามารถเฉลี่ยจากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

n แทน จำนวนผู้สอบ

x แทน ค่าความสามารถโดยประมาณจากแบบสอบวัดความสามารถทางคณิตศาสตร์

y แทน ค่าความสามารถเฉลี่ยจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

2) แปลงค่าความตรงตามสภาพให้เป็นคะแนนฟิชเชอร์ (Fishe's Z) และแปลงคะแนนฟิชเชอร์ กลับเป็นค่าความตรงตามสภาพ จากสูตร (Glass and Stanley, 1970; ศิริชัย กาญจนวาสี, 2548)

$$Z_r = 0.5[In(1+r) - In(1-r)]$$

$$r = \frac{\left(e^{2z} - 1\right)}{\left(e^{2z} + 1\right)}$$

เมื่อ Z_r แทน ค่าคะแนนฟิชเชอร์ซี

r แทน ค่าความตรงตามสภาพ

3) ทดสอบความแตกต่างของค่าความตรงตามตามสภาพ โดยใช้ค่าไคสแควร์ (Chi-Square) จากสูตร (Glass and Stanley, 1970; ศิริชัย กาญจนวาสี, 2548)

$$\sum_{i=1}^{j}\left(Z_{rj}-Z_{r}\right)^{2}\left(n_{j}-3\right)\rightarrow\chi_{j-1}^{2};\alpha$$

$$Z_{r}=\frac{\sum_{i=1}^{j}(n_{j}-3)Z_{rj}}{\sum_{i=1}^{j}(n_{j}-3)}$$

เมื่อ Z_{rj} แทน ค่าฟิชเชอร์ซี (Fisher's Z) ของกลุ่มผู้สอบj

n_{j} แทน จำนวนผู้สอบกลุ่มที่ j

4) เปรียบเทียบความแตกต่างของค่าความตรงตามสภาพเป็นรายคู่จากสูตร (Glass and Stanley, 1970; ศิริชัย กาญจนวาสี, 2548)

$$Z=\frac{Z_{r1}-Z_{r2}}{\sqrt{\frac{1}{n_{1}-3}+\frac{1}{n_{2}-3}}}$$

เมื่อ Z แทน สถิติทดสอบซี (Z test)

Z_{r1},Z_{r2} แทน ค่าฟิชเชอร์ซี ซึ่งแปลงมาจากค่าสหสัมพันธ์ r_{1} และ r_{2} ตามลำดับ

n แทน ขนาดของกลุ่มตัวอย่าง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่ต่างกัน ดังนั้นผู้วิจัยจึงนำเสนอผลการวิเคราะห์ข้อมูลเพื่อตอบวัตถุประสงค์ของการวิจัย ดังกล่าว โดยแบ่งออกเป็น 2 ตอน เพื่อเป็นการตรวจสอบความเป็นตัวแทนของข้อมูลและเป็นการตอบวัตถุประสงค์ของการวิจัย ดังนี้

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลเบื้องต้นในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง

ตอนที่ 2 ผลการเปรียบเทียบค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์

2.1 ผลการคำนวณค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ เมื่อเกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่ต่างกัน และทำการทดสอบความแตกต่างของค่าความตรงตามสภาพของการทดสอบ

2.2 ผลการทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ จำแนกตามเกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่ต่างกัน

2.3 ผลการคำนวณค่าความตรงตามสภาพและผลการทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกในแต่ละระดับของตัวแปรอิสระ 3 ตัวแปร คือ เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบ

เพื่อให้การนำเสนอผลการวิเคราะห์ข้อมูล และการทำความเข้าใจเกี่ยวกับผลการวิเคราะห์ข้อมูลมีความสะดวกยิ่งขึ้น ผู้วิจัยจึงกำหนดสัญลักษณ์และความหมายที่ใช้แทนค่าสถิติและตัวแปรต่างๆ ในการนำเสนอ ดังนี้

สัญลักษณ์ที่ใช้แทนค่าสถิติ

M	หมายถึง	ค่ามัชฌิมเลขคณิต หรือค่าเฉลี่ย (Mean)
SD	หมายถึง	ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation)
r	หมายถึง	ค่าความตรงตามสภาพของการทดสอบ

Z_r หมายถึง คะแนนฟิชเชอร์ซี

Z หมายถึง สถิติทดสอบซี (Z test)

χ^2 หมายถึง ค่าสถิติไคสแควร์ (Chi-Square)

p หมายถึง ระดับนัยสำคัญทางสถิติ

SEE หมายถึง ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลเบื้องต้นในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง

ผลการวิเคราะห์ข้อมูลเบื้องต้น โดยการวิเคราะห์ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบ จำนวนข้อสอบ และค่าฟังก์ชันสารสนเทศของแบบสอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง ได้ผลการวิเคราะห์ดังตารางที่ 9

ตารางที่ 9 จำนวนข้อมูลของผู้สอบ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของค่าความสามารถของผู้สอบ จำนวนข้อสอบ และค่าฟังก์ชันสารสนเทศของแบบสอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง

อัตราการใช้ข้อสอบซ้ำ (%)	เกณฑ์การคัดเลือกข้อสอบขั้นแรก	เกณฑ์ยุติการทดสอบ (SEE≤)	ครั้งที่	จำนวนข้อมูลของผู้สอบ	ค่าความสามารถ		จำนวนข้อสอบ		ค่าฟังก์ชันสารสนเทศของแบบสอบ	
					M	*SD*	*M*	*SD*	*M*	*SD*
10	วิธีสารสนเทศของฟิชเชอร์	0.30	1	142	0.2655	0.6712	19.7800	7.2150	11.3162	2.7957
			2	142	0.1844	0.6786	23.2900	6.9700	11.4031	2.7920
		0.45	1	142	0.2890	0.5789	8.0800	2.5000	5.1347	1.6766
			2	142	0.2054	0.6379	10.0400	2.6920	5.3092	1.5810
	วิธีสารสนเทศของคูลเบค-ไลเบลอร์	0.30	1	194	0.3933	0.7632	15.7800	3.3650	12.7044	2.3660
			2	194	0.3661	0.8919	19.9300	6.8520	12.8097	2.0705
		0.45	1	194	0.3768	0.6375	7.2600	0.9250	5.3852	1.5294
			2	194	0.3992	0.7895	8.6600	1.8760	5.7169	1.2044

ตารางที่ 9 (ต่อ)

อัตราการใช้ข้อสอบซ้ำ (%)	เกณฑ์การคัดเลือกข้อสอบขั้นแรก	เกณฑ์ยุติการทดสอบ (SEE≤)	ครั้งที่	จำนวนข้อมูลของผู้สอบ	ค่าความสามารถ		จำนวนข้อสอบ		ค่าฟังก์ชันสารสนเทศของแบบสอบ	
					M	*SD*	*M*	*SD*	*M*	*SD*
20	วิธีสารสนเทศของฟิชเชอร์	0.30	1	138	0.2520	0.5900	20.1700	7.2570	11.6798	2.3927
			2	138	0.1226	0.5930	22.8000	7.3090	11.6488	2.4964
		0.45	1	138	0.2622	0.5667	8.5300	3.0880	5.2930	1.7095
			2	138	0.1557	0.5629	10.0700	2.6350	5.3948	1.5466
	วิธีสารสนเทศของคูลเบค-ไลเบลอร์	0.30	1	154	0.3106	0.7648	15.9700	3.8880	12.4019	2.0656
			2	154	0.2617	0.9133	20.1800	7.4700	12.2737	2.1666
		0.45	1	154	0.3431	0.6621	7.2400	1.0360	5.3108	1.4070
			2	154	0.3594	0.8482	8.3400	1.7350	5.6612	1.4502
30	วิธีสารสนเทศของฟิชเชอร์	0.30	1	150	0.3301	0.7312	19.6800	6.9310	11.3393	2.8534
			2	150	0.2496	0.7034	23.2500	7.7480	11.8339	3.2173
		0.45	1	150	0.3425	0.6452	8.3200	3.1070	5.4972	1.7195
			2	150	0.2417	0.6732	10.3700	3.3490	5.5364	1.7591
	วิธีสารสนเทศของคูลเบค-ไลเบลอร์	0.30	1	146	0.3604	0.7880	15.9900	4.0760	12.7008	2.3090
			2	146	0.3349	0.9289	19.8500	7.0520	12.5771	2.3972
		0.45	1	146	0.3912	0.6889	7.4600	1.2660	5.4950	1.5601
			2	146	0.3644	0.8594	8.6200	1.6740	5.7848	1.5999
		รวม		1,848						

จากตารางที่ 9 การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์จำนวน 2 ครั้ง เมื่ออัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 พบว่า ผู้สอบมีค่าความสามารถเฉลี่ยอยู่ระหว่าง 0.1844 ถึง 0.2655 จำนวนข้อสอบเฉลี่ยประมาณ 20 ถึง 23 ข้อ และค่าฟังก์ชันสารสนเทศของแบบสอบเฉลี่ยอยู่ระหว่าง 11.3162 ถึง 11.4031

เมื่ออัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานใน

การประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 พบว่า ผู้สอบมีค่าความสามารถเฉลี่ยอยู่ระหว่าง 0.2054 ถึง 0.2890 จำนวนข้อสอบเฉลี่ยประมาณ 8 ถึง 10 ข้อ และค่าฟังก์ชันสารสนเทศของแบบสอบเฉลี่ยอยู่ระหว่าง 5.1347 ถึง 5.3092

เมื่ออัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค - ไลเบลอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 พบว่า ผู้สอบมีค่าความสามารถเฉลี่ยอยู่ระหว่าง 0.3661 ถึง 0.3933 จำนวนข้อสอบเฉลี่ยประมาณ 16 ถึง 20 ข้อ และค่าฟังก์ชันสารสนเทศของแบบสอบเฉลี่ยอยู่ระหว่าง 12.7044 ถึง 12.8097

เมื่ออัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค - ไลเบลอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 พบว่า ผู้สอบมีค่าความสามารถเฉลี่ยอยู่ระหว่าง 0.3768 ถึง 0.3992 จำนวนข้อสอบเฉลี่ยประมาณ 7 ถึง 9 ข้อ และค่าฟังก์ชันสารสนเทศของแบบสอบเฉลี่ยอยู่ระหว่าง 5.3852 ถึง 5.7169

เมื่ออัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 พบว่า ผู้สอบมีค่าความสามารถเฉลี่ยของผู้สอบอยู่ระหว่าง 0.1226 ถึง 0.2520 จำนวนข้อสอบเฉลี่ยประมาณ 20 ถึง 23 ข้อ และค่าฟังก์ชันสารสนเทศของแบบสอบเฉลี่ยอยู่ระหว่าง 11.6488 ถึง 11.6798

เมื่ออัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 พบว่า ผู้สอบมีค่าความสามารถเฉลี่ยอยู่ระหว่าง 0.1557 ถึง 0.2622 จำนวนข้อสอบเฉลี่ยประมาณ 9 ถึง 10 ข้อ และค่าฟังก์ชันสารสนเทศของแบบสอบเฉลี่ยอยู่ระหว่าง 5.2930 ถึง 5.3948

เมื่ออัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค - ไลเบลอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 พบว่า ผู้สอบมีค่าความสามารถเฉลี่ยอยู่ระหว่าง 0.2617 ถึง 0.3106 จำนวนข้อสอบเฉลี่ยประมาณ 16 ถึง 20 ข้อ และค่าฟังก์ชันสารสนเทศของแบบสอบเฉลี่ยอยู่ระหว่าง 12.2737 ถึง 12.4019

เมื่ออัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค - ไลเบลอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 พบว่า ผู้สอบมีค่า

ความสามารถเฉลี่ยอยู่ระหว่าง 0.3431 ถึง 0.3594 จำนวนข้อสอบเฉลี่ยประมาณ 7 ถึง 8 ข้อ และค่าฟังก์ชันสารสนเทศของแบบสอบเฉลี่ยอยู่ระหว่าง 5.3108 ถึง 5.6612

เมื่ออัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 พบว่า ผู้สอบมีค่าความสามารถเฉลี่ยของผู้สอบอยู่ระหว่าง 0.2496 ถึง 0.3301 จำนวนข้อสอบเฉลี่ยประมาณ 20 ถึง 23 ข้อ และค่าฟังก์ชันสารสนเทศของแบบสอบเฉลี่ยอยู่ระหว่าง 11.3393 ถึง 11.8339

เมื่ออัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 พบว่า ผู้สอบมีค่าความสามารถเฉลี่ยอยู่ระหว่าง 0.2417 ถึง 0.3425 จำนวนข้อสอบเฉลี่ยประมาณ 8 ถึง 10 ข้อ และค่าฟังก์ชันสารสนเทศของแบบสอบเฉลี่ยอยู่ระหว่าง 5.4972 ถึง 5.5364

เมื่ออัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค - ไลเบลอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 พบว่า ผู้สอบมีค่าความสามารถเฉลี่ยอยู่ระหว่าง 0.3349 ถึง 0.3604 จำนวนข้อสอบเฉลี่ยประมาณ 16 ถึง 20 ข้อ และค่าฟังก์ชันสารสนเทศของแบบสอบเฉลี่ยอยู่ระหว่าง 12.5771 ถึง 12.7008

เมื่ออัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค - ไลเบลอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 พบว่า ผู้สอบมีค่าความสามารถเฉลี่ยอยู่ระหว่าง 0.3644 ถึง 0.3912 จำนวนข้อสอบเฉลี่ยประมาณ 7 ถึง 9 ข้อ และค่าฟังก์ชันสารสนเทศของแบบสอบเฉลี่ยอยู่ระหว่าง 5.4950 ถึง 5.7848

จากที่กล่าวมาข้างต้นจะเห็นได้ว่า ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จำนวน 2 ครั้ง เมื่อพิจารณาค่าความสามารถในอัตราการใช้ข้อสอบซ้ำ10 เปอร์เซ็นต์ 20 เปอร์เซ็นต์ และ 30 เปอร์เซ็นต์ พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์จะให้การประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยเฉลี่ยสูงกว่าเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ครั้งที่ 1 จะมีแนวโน้มให้ค่าความสามารถของผู้สอบโดยเฉลี่ยมากกว่าครั้งที่ 2 ส่วนเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 ผู้วิจัยมีความเห็นว่าค่าความสามารถของผู้สอบไม่ต่างกันมากในทุกอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ 20 เปอร์เซ็นต์ และ 30 เปอร์เซ็นต์

นอกจากนี้เป็นที่น่าสังเกตว่า เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ ในอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ มีแนวโน้มในการประมาณค่าความสามารถของผู้สอบสูงกว่าในอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ และ 20 เปอร์เซ็นต์

เมื่อพิจารณาจำนวนข้อสอบในอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ 20 เปอร์เซ็นต์ และ 30 เปอร์เซ็นต์ พบว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ครั้งที่ 2 ใช้จำนวนข้อสอบโดยเฉลี่ยมากกว่าครั้งที่ 1 และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ใช้จำนวนข้อสอบโดยเฉลี่ยมากกว่าเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ประมาณ 50 เปอร์เซ็นต์ ส่วนเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ จะใช้จำนวนข้อสอบโดยเฉลี่ยมากกว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ ในจำนวนไม่มาก

เมื่อพิจารณาค่าฟังก์ชันสารสนเทศของแบบสอบในอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ 20 เปอร์เซ็นต์ และ 30 เปอร์เซ็นต์ พบว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ครั้งที่ 2 มีแนวโน้มในการให้ค่าฟังก์ชันสารสนเทศของแบบสอบโดยเฉลี่ยสูงกว่าครั้งที่ 1 และเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ ให้ค่าฟังก์ชันสารสนเทศโดยเฉลี่ยสูงกว่าเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ ส่วนเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ให้ค่าฟังก์ชันสารสนเทศของแบบสอบโดยเฉลี่ยสูงกว่าเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ประมาณ 50 เปอร์เซ็นต์

จากที่กล่าวมาทั้งหมดสามารถสรุปได้ว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค - ไลเบลอร์และใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ที่อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ จะให้ค่าฟังก์ชันสารสนเทศของแบบสอบสูงที่สุด

การรายงานข้อมูลเกี่ยวกับแบบสอบวัดความสามารถทางคณิตศาสตร์จำนวน 40 ข้อ ซึ่งเป็นเกณฑ์ในการตรวจสอบความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ ได้แสดงในตารางที่ 10

ตารางที่ 10 คะแนน ค่าความสามารถ และส่วนเบี่ยงเบนมาตรฐานจากแบบสอบวัดความสามารถทางคณิตศาสตร์ของนักเรียนที่เป็นกลุ่มตัวอย่าง

รายการ	*M*	*SD*
คะแนนของผู้สอบ	24.7300	8.1140
ค่าความสามารถของผู้สอบ	0.0330	1.6010

จากตารางที่ 10 แบบสอบวัดความสามารถทางคณิตศาสตร์มีคะแนนเต็ม 40 คะแนน ผู้สอบที่เป็นกลุ่มตัวอย่างมีคะแนนเฉลี่ย 24.7300 คะแนน (SD = 8.1140) และค่าความสามารถเฉลี่ยเท่ากับ 0.0330 (SD_{θ} = 1.6010)

ตอนที่ 2 ผลการเปรียบเทียบความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์

2.1 ผลการคำนวณค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ เมื่อเกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่ต่างกัน

การคำนวณค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ โดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าความสามารถเฉลี่ยที่ได้จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ กับค่าความสามารถที่ได้จากแบบสอบวัดความสามารถทางคณิตศาสตร์ที่ผู้วิจัยสร้างขึ้นเพื่อตรวจสอบความตรงตามสภาพ รวมทั้งศึกษาความแตกต่างของค่าความตรงตามสภาพโดยใช้ค่าไคสแควร์ (Chi-Square) ได้ผลวิเคราะห์ดังตารางที่ 11

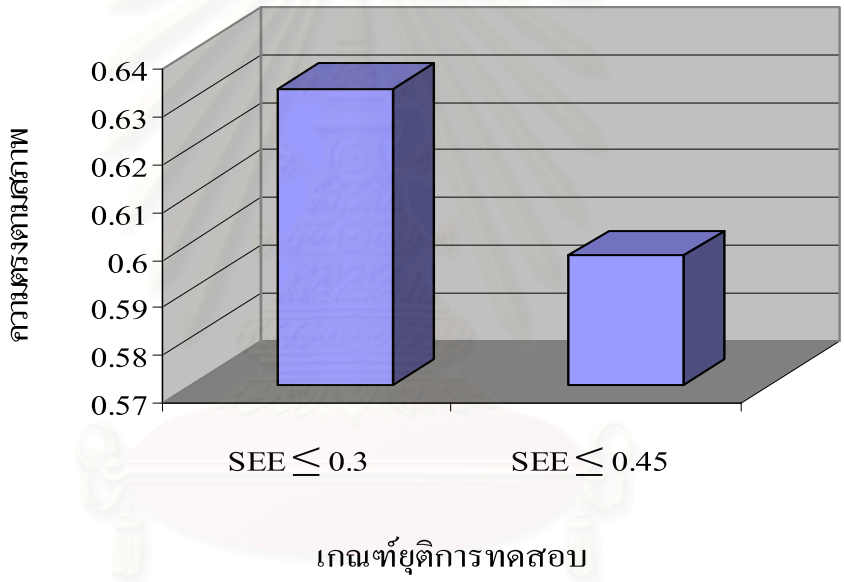
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 11 ค่าความตรงตามสภาพและค่าไคสแควร์ (Chi-Square) สำหรับการทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์

อัตราการใช้ข้อสอบซ้ำ (%)	เกณฑ์การคัดเลือกข้อสอบขั้นแรก	เกณฑ์ยุติการทดสอบ	จำนวนข้อมูลของผู้สอบ	ความตรงตามสภาพ			χ^2
				r	Z_r		
10%	วิธีสารสนเทศฟิชเชอร์	SEE ≤ 0.30	142	0.6320	0.7447	0.2200	10.3408*
		SEE ≤ 0.45	142	0.5970	0.6885		
	วิธีสารสนเทศคูลเบค-ไลเบลอร์	SEE ≤ 0.30	194	0.7670	1.0130	1.8660	
		SEE ≤ 0.45	194	0.7030	0.8732		
20%	วิธีสารสนเทศฟิชเชอร์	SEE ≤ 0.30	138	0.5970	0.6885	0.0377	7.5940
		SEE ≤ 0.45	138	0.6120	0.7121		
	วิธีสารสนเทศคูลเบค-ไลเบลอร์	SEE ≤ 0.30	154	0.7460	0.9639	0.5165	
		SEE ≤ 0.45	154	0.7070	0.8812		
30%	วิธีสารสนเทศฟิชเชอร์	SEE ≤ 0.30	150	0.7020	0.8712	0.0137	1.0794
		SEE ≤ 0.45	150	0.6950	0.8575		
	วิธีสารสนเทศคูลเบค-ไลเบลอร์	SEE ≤ 0.30	146	0.7480	0.9684	0.5195	
		SEE ≤ 0.45	146	0.7080	0.8832		
รวม							$\chi^2 = 20.7556$*

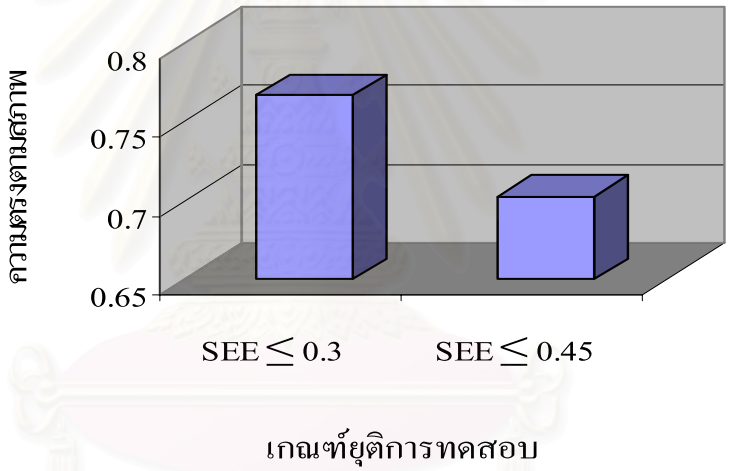
* $p < .05$

จากตารางที่ 11 พบว่า เมื่ออัตราการใช้ข้อสอบซ้ำ 10% เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เท่ากับ 0.6320 0.5970 ตามลำดับ เมื่อแปลงค่าความตรงตามสภาพของการทดสอบให้เป็นคะแนนฟิชเชอร์ซี (Z_r) จากสูตรของ Glass and Stanley (1970) ได้ค่าเท่ากับ 0.7447 0.6885 ตามลำดับ และทำการทดสอบความแตกต่างของค่าความตรงตามสภาพดังกล่าว พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 18



แผนภาพที่ 18 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์ยุติการทดสอบและเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์และอัตราการใช้สอบซ้ำ 10%

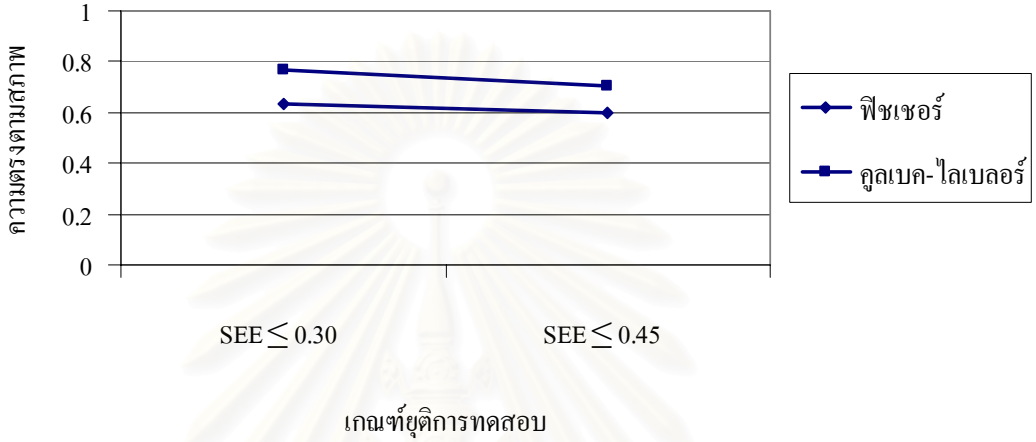
เมื่ออัตราการใช้ข้อสอบซ้ำ 10% เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เท่ากับ 0.7670 0.7030 ตามลำดับ เมื่อแปลงค่าความตรงตามสภาพของการทดสอบให้เป็นคะแนนฟิชเชอร์ซี (Z_r) จากสูตรของ Glass and Stanley (1970) ได้ค่าเท่ากับ 1.0130 0.8732 ตามลำดับ และทำการทดสอบความแตกต่างของค่าความตรงตามสภาพดังกล่าว พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 19



แผนภาพที่ 19 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์ยุติการทดสอบและเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์และอัตราการใช้สอบซ้ำ 10%

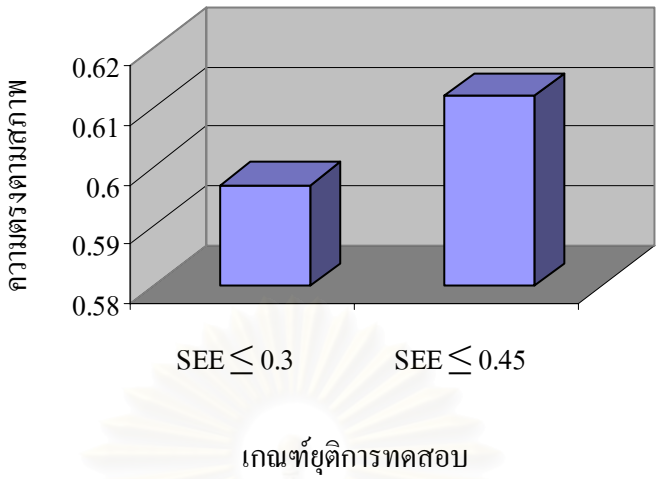
เมื่ออัตราการใช้ข้อสอบซ้ำ 10% และใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 และเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 พบว่า เกณฑ์การคัดเลือกข้อสอบที่ขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ เกณฑ์ยุติการ

ทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 และวิธีสารสนเทศของคูลเบค-ไลเบลอร์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังแสดงในแผนภาพที่ 20



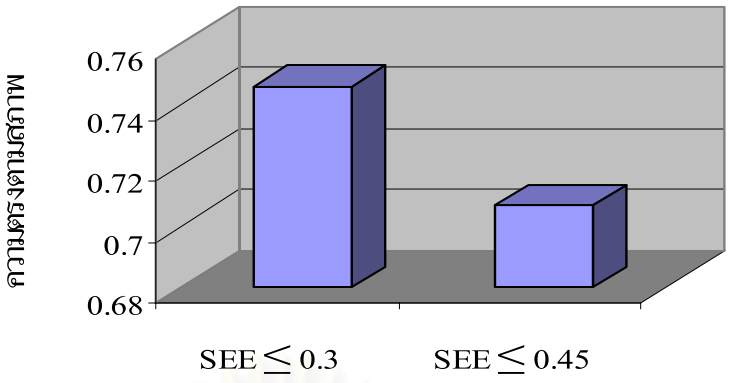
แผนภาพที่ 20 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์ยุติการทดสอบและเกณฑ์การคัดเลือกข้อสอบขั้นแรกและอัตราการใช้สอบซ้ำ 10%

เมื่ออัตราการใช้ข้อสอบซ้ำ 20% เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เท่ากับ 0.5970 0.6120 ตามลำดับ เมื่อแปลงค่าความตรงตามสภาพของการทดสอบให้เป็นคะแนนฟิชเชอร์ซี (Z_r) จากสูตรของ Glass and Stanley (1970) ได้ค่าเท่ากับ 0.6885 0.7121 ตามลำดับ และทำการทดสอบความแตกต่างของค่าความตรงตามสภาพดังกล่าว พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 21



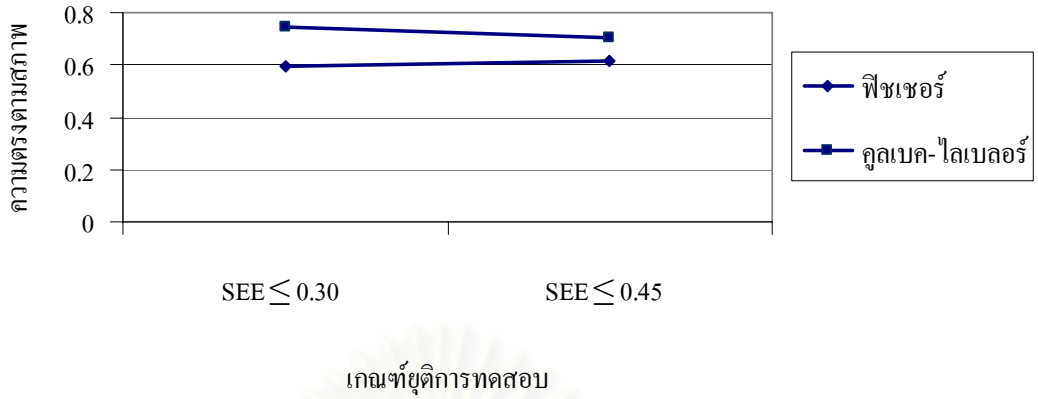
แผนภาพที่ 21 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์ยุติการทดสอบและเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์และอัตราการใช้สอบซ้ำ 20%

เมื่ออัตราการใช้ข้อสอบซ้ำ 20% เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เท่ากับ 0.7460 0.7070 ตามลำดับ เมื่อแปลงค่าความตรงตามสภาพของการทดสอบให้เป็นคะแนนฟิชเชอร์ซี (Z_r) จากสูตรของ Glass and Stanley (1970) ได้ค่าเท่ากับ 0.9639 0.8812 ตามลำดับ และทำการทดสอบความแตกต่างของค่าความตรงตามสภาพดังกล่าว พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 22



แผนภาพที่ 22 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์ยุติการทดสอบและเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ และอัตราการใช้สอบซ้ำ 20%

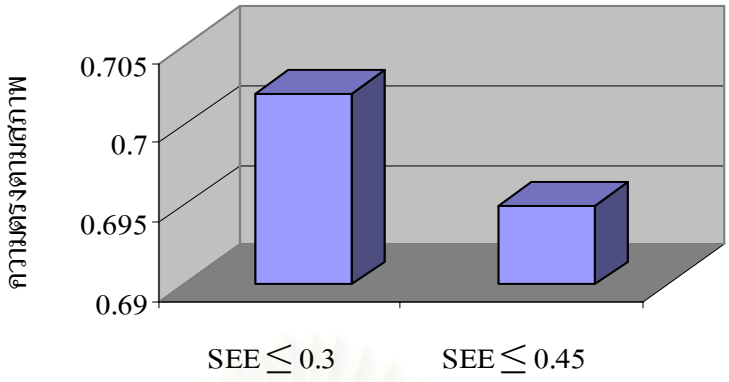
เมื่ออัตราการใช้ข้อสอบซ้ำ 20% และใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 และเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 พบว่า เกณฑ์การคัดเลือกข้อสอบที่ขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 และวิธีสารสนเทศของคูลเบค-ไลเบลอร์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 23



แผนภาพที่ 23 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์ยุติการทดสอบและเกณฑ์การคัดเลือกข้อสอบขั้นแรกและอัตราการใช้สอบซ้ำ 20%

เมื่ออัตราการใช้ข้อสอบซ้ำ 30% เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เท่ากับ 0.7020 0.6950 ตามลำดับ เมื่อแปลงค่าความตรงตามสภาพของการทดสอบให้เป็นคะแนนฟิชเชอร์ซี (Z_r) จากสูตรของ Glass and Stanley (1970) ได้ค่าเท่ากับ 0.8712 0.8575 ตามลำดับ และทำการทดสอบความแตกต่างของค่าความตรงตามสภาพดังกล่าว พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 24

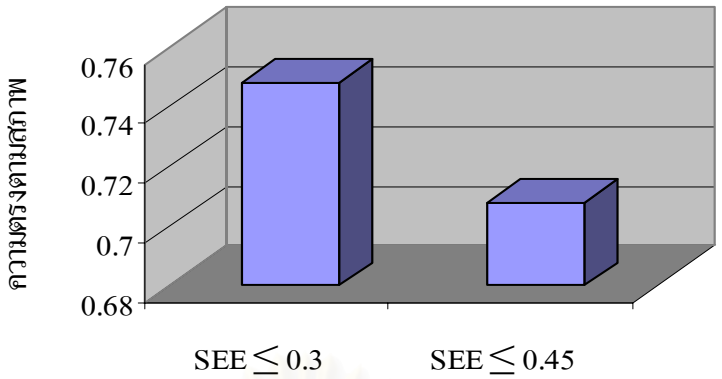
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



แผนภาพที่ 24 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์ยุติการทดสอบและเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์และอัตราการใช้สอบซ้ำ 30%

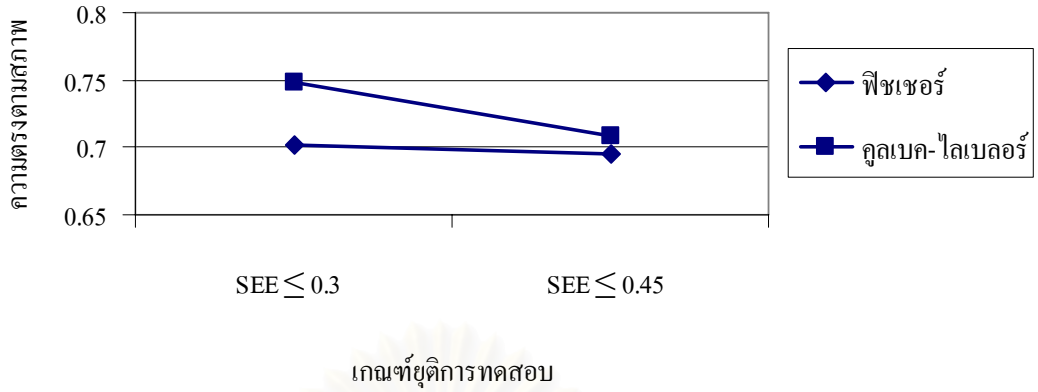
เมื่ออัตราการใช้ข้อสอบซ้ำ 30% เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เท่ากับ 0.7480 0.7080 ตามลำดับ เมื่อแปลงค่าความตรงตามสภาพของการทดสอบให้เป็นคะแนนฟิชเชอร์ซี (Z_r) จากสูตรของ Glass and Stanley (1970) ได้ค่าเท่ากับ 0.9684 0.8832 ตามลำดับ และทำการทดสอบความแตกต่างของค่าความตรงตามสภาพดังกล่าว พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 25

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



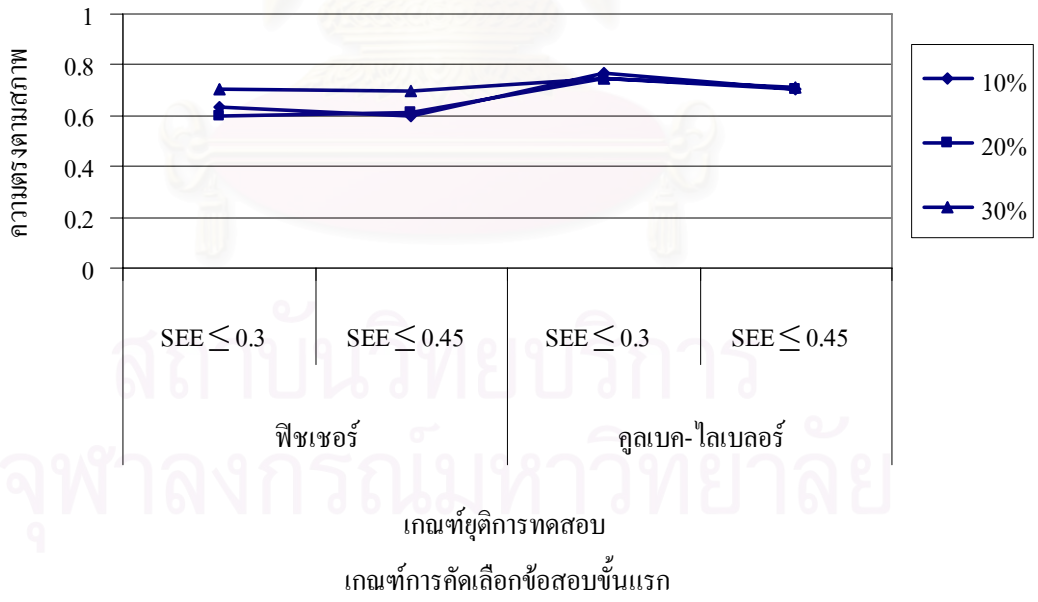
แผนภาพที่ 25 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์ยุติการทดสอบและเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์และอัตราการใช้สอบซ้ำ 30%

เมื่ออัตราการใช้ข้อสอบซ้ำ 30% และใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 และเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 พบว่า เกณฑ์การคัดเลือกข้อสอบที่ขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 และวิธีสารสนเทศของคูลเบค-ไลเบลอร์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังแสดงในแผนภาพที่ 26



แผนภาพที่ 26 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์ยุติการทดสอบและเกณฑ์การคัดเลือกข้อสอบขั้นแรกและอัตราการใช้สอบซ้ำ 30%

เมื่อทำการทดสอบภาพรวมของค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่า มีอย่างน้อย 1 คู่ ที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังแสดงในแผนภาพที่ 27



แผนภาพที่ 27 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำและเกณฑ์ยุติการทดสอบ

2.2 ผลการทดสอบความแตกต่างของค่าความตรงสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

จากผลการทดสอบภาพรวมของค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่า มีอย่างน้อย 1 คู่ ที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ผู้วิจัยจึงนำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ มาทดสอบความแตกต่างเป็นรายคู่ จำแนกตามเกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบ ซึ่งผลการทดสอบปรากฏดังตารางที่ 12 – 41

2.2.1 การทดสอบความแตกต่างของเกณฑ์ยุติการทดสอบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบจำแนกตามเกณฑ์การคัดเลือกข้อสอบขั้นแรก ดังแสดงในตารางที่ 12 – 13

ตารางที่ 12 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์

เกณฑ์ยุติการทดสอบ	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
SEE ≤ 0.30	430	0.6250	0.7788	0.2762
SEE ≤ 0.45	430	0.6410	0.7599	

จากตารางที่ 12 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 13 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์

เกณฑ์ยุติการทดสอบ	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
SEE ≤ 0.30	494	0.7540	0.9822	1.6138
SEE ≤ 0.45	494	0.7060	0.8792	

จากตารางที่ 13 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

2.2.2 การทดสอบความแตกต่างของเกณฑ์การคัดเลือกข้อสอบขั้นแรก จำแนกตามเกณฑ์ยุติการทดสอบ ดังแสดงในตารางที่ 14 – 15

ตารางที่ 14 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30

เกณฑ์การคัดเลือกข้อสอบขั้นแรก	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
วิธีสารสนเทศของฟิชเชอร์	430	0.6250	0.7788	3.0739**
วิธีสารสนเทศของคูลเบค-ไลเบลอร์	494	0.7540	0.9822	

** $p < .01$

จากตารางที่ 14 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพของการทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .01

ตารางที่ 15 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45

เกณฑ์การคัดเลือกข้อสอบขั้นแรก	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
วิธีสารสนเทศของฟิชเชอร์	430	0.6410	0.7599	1.8029
วิธีสารสนเทศของคูลเบค-ไลเบลอร์	494	0.7060	0.8792	

จากตารางที่ 15 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2.2.3 การทดสอบความแตกต่างของอัตราการใช้ข้อสอบซ้ำ จำแนกตามเกณฑ์การคัดเลือกข้อสอบขั้นแรก ดังแสดงในตารางที่ 16 – 17

ตารางที่ 16 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์

อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)	จำนวนข้อมูลของผู้สอบ	ความตรงสภาพ		อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)		
		r	Z_r	10%	20%	30%
10%	284	0.6140	0.7153		0.1871	1.7556
20%	276	0.6040	0.6994			1.9321
30%	300	0.6970	0.8614			

จากตารางที่ 16 นำค่าความตรงตามสภาพของการทดสอบแบบปรับกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศฟิชเชอร์ มาทำการทดสอบความแตกต่าง พบว่า อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ และอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 17 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์

อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)	จำนวนข้อมูลของผู้สอบ	ความตรงสภาพ		อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)		
		r	Z_r	10%	20%	30%
10%	388	0.7350	0.9395		0.2792	0.2216
20%	308	0.7250	0.9181			0.0512
30%	292	0.7270	0.9223			

จากตารางที่ 17 นำค่าความตรงตามสภาพของการทดสอบแบบปรับกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ มาทำการทดสอบความแตกต่าง พบว่า อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ และอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

2.2.4 การทดสอบความแตกต่างของอัตราการใช้ข้อสอบซ้ำ จำแนกตามเกณฑ์ยุติการทดสอบ ดังแสดงในตารางที่ 18 – 19

ตารางที่ 18 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30

อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)	จำนวนข้อมูลของผู้สอบ	ความตรงสภาพ		อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)		
		r	Z_r	10%	20%	30%
10%	336	0.7230	0.9139		0.6282	0.0787
20%	292	0.6980	0.8634			0.6851
30%	296	0.7260	0.9202			

จากตารางที่ 18 นำค่าความตรงตามสภาพของการทดสอบแบบปรับกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 มาทำการทดสอบความแตกต่าง พบว่า อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ และอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 19 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45

อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)	จำนวนข้อมูลของผู้สอบ	ความตรงสภาพ		อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)		
		r	Z_r	10%	20%	30%
10%	336	0.6690	0.8089		0.1592	0.7291
20%	292	0.6760	0.8217			0.5500
30%	296	0.7000	0.8673			

จากตารางที่ 19 นำค่าความตรงตามสภาพของการทดสอบแบบปรับกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 มาทำการทดสอบความแตกต่าง พบว่า อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ และอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

2.2.5 การทดสอบความแตกต่างของเกณฑ์การคัดเลือกข้อสอบขั้นแรก จำแนกตามอัตราการใช้ข้อสอบซ้ำ ดังแสดงในตารางที่ 20 – 22

ตารางที่ 20 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำ 10 %

เกณฑ์การคัดเลือกข้อสอบขั้นแรก	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
วิธีสารสนเทศของฟิชเชอร์	284	0.6140	0.7153	2.8575**
วิธีสารสนเทศของคูลเบค-ไลเบลอร์	388	0.7350	0.9395	

** $p < .01$

จากตารางที่ 20 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และวิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพของการทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ตารางที่ 21 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำ 20 %

เกณฑ์การคัดเลือกข้อสอบขั้นแรก	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
วิธีสารสนเทศของฟิชเชอร์	276	0.6040	0.6994	2.6249**
วิธีสารสนเทศของคูลเบค-ไลเบลอร์	308	0.7250	0.9181	

** p < .01

จากตารางที่ 21 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และวิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพของการทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ตารางที่ 22 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำ 30 %

เกณฑ์การคัดเลือกข้อสอบขั้นแรก	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
วิธีสารสนเทศของฟิชเชอร์	300	0.6970	0.8614	0.7370
วิธีสารสนเทศของคูลเบค-ไลเบลอร์	292	0.7270	0.9223	

จากตารางที่ 22 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และวิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

2.2.6 การทดสอบความแตกต่างของเกณฑ์ยุติการทดสอบ จำแนกตามอัตราการใช้ข้อสอบซ้ำ ดังแสดงในตารางที่ 23 – 25

ตารางที่ 23 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำ 10%

เกณฑ์ยุติการทดสอบ	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
SEE ≤ 0.30	336	0.7230	0.9139	1.3549
SEE ≤ 0.45	336	0.6690	0.8089	

จากตารางที่ 23 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 24 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำ 20%

เกณฑ์ยุติการทดสอบ	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
SEE ≤ 0.30	292	0.6980	0.8634	0.5013
SEE ≤ 0.45	292	0.6760	0.8217	

จากตารางที่ 24 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 25 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำ 30%

เกณฑ์ยุติการทดสอบ	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
SEE ≤ 0.30	296	0.7260	0.9202	0.6403
SEE ≤ 0.45	296	0.7000	0.8673	

จากตารางที่ 25 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2.2.7 การทดสอบความแตกต่างของอัตราการใช้ข้อสอบซ้ำ จำแนกตามเกณฑ์การคัดเลือกข้อสอบขั้นแรก และเกณฑ์ยุติการทดสอบ ดังแสดงในตารางที่ 26 – 29

ตารางที่ 26 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30

อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)	จำนวนข้อมูลของผู้สอบ	ความตรงสภาพ		อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)		
		r	Z_r	10%	20%	30%
10%	142	0.6320	0.7447		0.4651	1.0692
20%	138	0.5970	0.6885			1.5443
30%	150	0.7020	0.8712			

จากตารางที่ 26 นำค่าความตรงตามสภาพของการทดสอบแบบปรับกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 มาทำการทดสอบความแตกต่าง พบว่า อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ และอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 27 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45

อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)	จำนวนข้อมูลของผู้สอบ	ความตรงสภาพ		อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)		
		r	Z_r	10%	20%	30%
10%	142	0.5970	0.6885		0.1953	1.4285
20%	138	0.6120	0.7121			1.2197
30%	150	0.6950	0.8575			

จากตารางที่ 27 นำค่าความตรงตามสภาพของการทดสอบแบบปรับกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 มาทำการทดสอบความแตกต่าง พบว่า อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ และอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ มีค่าความตรงตามสภาพของการไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 28 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ และ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30

อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)	จำนวนข้อมูลของผู้สอบ	ความตรงสภาพ		อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)		
		r	Z_r	10%	20%	30%
10%	194	0.7670	1.0130		0.4509	0.4033
20%	154	0.7460	0.9639			0.0386
30%	146	0.7480	0.9684			

จากตารางที่ 28 นำค่าความตรงตามสภาพของการทดสอบแบบปรับกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 มาทำการทดสอบความแตกต่าง พบว่า อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ และอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 29 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ และ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45

อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)	จำนวนข้อมูลของผู้สอบ	ความตรงสภาพ		อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)		
		r	Z_r	10%	20%	30%
10%	194	0.7030	0.8732		0.0735	0.0904
20%	154	0.7070	0.8812			0.0171
30%	146	0.7080	0.8832			

จากตารางที่ 29 นำค่าความตรงตามสภาพของการทดสอบแบบปรับกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 มาทำการทดสอบความแตกต่าง พบว่า อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ และอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

2.2.8 การทดสอบความแตกต่างของเกณฑ์ยุติการทดสอบ จำแนกตามเกณฑ์การคัดเลือกข้อสอบขั้นแรกและอัตราการใช้ข้อสอบซ้ำ ดังแสดงในตารางที่ 30– 35

ตารางที่ 30 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์และอัตราการใช้ข้อสอบซ้ำ 10%

เกณฑ์ยุติการทดสอบ	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
SEE ≤ 0.30	142	0.6320	0.7447	0.4685
SEE ≤ 0.45	142	0.5970	0.6885	

จากตารางที่ 30 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และอัตราการใช้ข้อสอบซ้ำ 10% มาทำการทดสอบความแตกต่าง พบว่าเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 31 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์และอัตราการใช้ข้อสอบซ้ำ 20%

เกณฑ์ยุติการทดสอบ	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
SEE ≤ 0.30	138	0.5970	0.6885	0.1939
SEE ≤ 0.45	138	0.6120	0.7121	

* $p < .05$

จากตารางที่ 31 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และอัตราการใช้ข้อสอบซ้ำ 20% มาทำการทดสอบความแตกต่าง พบว่าเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ

ของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 32 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์และอัตราการใช้ข้อสอบซ้ำ 30%

เกณฑ์ยุติการทดสอบ	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
SEE ≤ 0.30	150	0.7020	0.8712	0.1175
SEE ≤ 0.45	150	0.6950	0.8575	

จากตารางที่ 32 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และอัตราการใช้ข้อสอบซ้ำ 30% มาทำการทดสอบความแตกต่าง พบว่าเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 33 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์และอัตราการใช้ข้อสอบซ้ำ 10%

เกณฑ์ยุติการทดสอบ	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
SEE ≤ 0.30	194	0.7670	1.0130	1.3662
SEE ≤ 0.45	194	0.7030	0.8732	

จากตารางที่ 33 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ และอัตราการใช้ข้อสอบซ้ำ 10% มาทำการทดสอบความแตกต่าง

พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่าง กันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 34 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์และอัตราการใช้ข้อสอบซ้ำ 20%

เกณฑ์ยุติการทดสอบ	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
SEE ≤ 0.30	154	0.7460	0.9639	0.7186
SEE ≤ 0.45	154	0.7070	0.8812	

จากตารางที่ 34 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ และอัตราการใช้ข้อสอบซ้ำ 20% มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 35 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์และอัตราการใช้ข้อสอบซ้ำ 30%

เกณฑ์ยุติการทดสอบ	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
SEE ≤ 0.30	146	0.7480	0.9684	0.7204
SEE ≤ 0.45	146	0.7080	0.8832	

จากตารางที่ 35 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ และอัตราการใช้ข้อสอบซ้ำ 30% มาทำการทดสอบความแตกต่าง

พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

2.2.9 การทดสอบความแตกต่างของเกณฑ์การคัดเลือกข้อสอบขั้นแรก จำแนกตามเกณฑ์ยุติการทดสอบและอัตราการใช้ข้อสอบซ้ำ ดังแสดงในตารางที่ 36 – 41

ตารางที่ 36 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30และอัตราการใช้ข้อสอบซ้ำ 10%

เกณฑ์การคัดเลือกข้อสอบขั้นแรก	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
วิธีสารสนเทศของฟิชเชอร์	139	0.6320	0.7447	2.4065*
วิธีสารสนเทศของคูลเบค-ไลเบลอร์	139	0.7670	1.0130	

* $p < .05$

จากตารางที่ 36 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และอัตราการใช้ข้อสอบซ้ำ 10% มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพของการทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 37 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30และอัตราการใช้ข้อสอบซ้ำ 20%

เกณฑ์การคัดเลือกข้อสอบขั้นแรก	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
วิธีสารสนเทศของฟิชเชอร์	138	0.5970	0.6885	2.3495*
วิธีสารสนเทศของคูลเบค-ไลเบลอร์	154	0.7460	0.9639	

* $p < .05$

จากตารางที่ 37 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และอัตราการใช้ข้อสอบซ้ำ 20% มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพของการทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตารางที่ 38 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30และอัตราการใช้ข้อสอบซ้ำ 30%

เกณฑ์การคัดเลือกข้อสอบขั้นแรก	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
วิธีสารสนเทศของฟิชเชอร์	150	0.7020	0.8712	0.8275
วิธีสารสนเทศของคูลเบค-ไลเบลอร์	147	0.7480	0.9684	

* $p < .05$

จากตารางที่ 38 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และอัตราการใช้ข้อสอบซ้ำ 30% มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์การคัดเลือกข้อสอบขั้น

แรกด้วยวิธีสารสนเทศของฟิชเชอร์ และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 39 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 และอัตราการใช้ข้อสอบซ้ำ 10%

เกณฑ์การคัดเลือกข้อสอบขั้นแรก	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
วิธีสารสนเทศของฟิชเชอร์	142	0.5970	0.6885	1.6567
วิธีสารสนเทศของคูลเบค-ไลเบลอร์	194	0.7030	0.8732	

จากตารางที่ 39 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 และอัตราการใช้ข้อสอบซ้ำ 10% มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 40 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 และอัตราการใช้ข้อสอบซ้ำ 20%

เกณฑ์การคัดเลือกข้อสอบขั้นแรก	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
วิธีสารสนเทศของฟิชเชอร์	138	0.6120	0.7121	1.4276
วิธีสารสนเทศของคูลเบค-ไลเบลอร์	154	0.7070	0.8812	

จากตารางที่ 40 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 และอัตราการใช้ข้อสอบซ้ำ 20% มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ตารางที่ 41 การทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 และอัตราการใช้ข้อสอบซ้ำ 30%

เกณฑ์การคัดเลือกข้อสอบขั้นแรก	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
วิธีสารสนเทศของฟิชเชอร์	150	0.6950	0.8575	0.2188
วิธีสารสนเทศของคูลเบค-ไลเบลอร์	146	0.7080	0.8832	

จากตารางที่ 41 นำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 และอัตราการใช้ข้อสอบซ้ำ 30% มาทำการทดสอบความแตกต่าง พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

2.3 ผลการคำนวณค่าความตรงตามสภาพและผลการทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบ

การคำนวณค่าความตรงตามสภาพของการทดสอบ พิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าความสามารถเฉลี่ยที่ได้จากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์กับค่าความสามารถที่ได้จากแบบสอบวัดความสามารถทางคณิตศาสตร์ และนำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วย

คอมพิวเตอร์มาทดสอบความแตกต่างเป็นรายคู่ในแต่ละระดับของตัวแปรอิสระ 3 ตัวแปร คือ เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำและเกณฑ์ยุติการทดสอบ ได้ผลการคำนวณค่าความตรงตามสภาพและผลการทดสอบความแตกต่างดังตารางที่ 42 - 44

ตารางที่ 42 ค่าความตรงตามสภาพและการทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์การคัดเลือกข้อสอบขั้นแรก

เกณฑ์การคัดเลือกข้อสอบขั้นแรก	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
วิธีสารสนเทศของฟิชเชอร์	860	0.6460	0.7684	3.3866**
วิธีสารสนเทศของคูลเบค-ไลเบลอร์	988	0.7290	0.9266	

** $p < .01$

จากตารางที่ 42 เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกเป็นเกณฑ์ในการศึกษาความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพ เท่ากับ 0.6460 0.7290 ตามลำดับ และนำมาทดสอบความแตกต่างเป็นรายคู่ จึงแปลงค่าความตรงตามสภาพของการทดสอบให้เป็นคะแนนฟิชเชอร์ซี (Z_r) จากสูตรของ Glass and Stanley (1970) ได้ค่าเท่ากับ 0.7684 0.9266 ตามลำดับ และเมื่อนำค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มาทดสอบความแตกต่างเป็นรายคู่จำแนกตามเกณฑ์การคัดเลือกข้อสอบขั้นแรก พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ มีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แตกต่างจากเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อใช้อัตราการใช้ข้อสอบซ้ำเป็นเกณฑ์ในการศึกษาความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ได้ผลการคำนวณและการทดสอบความแตกต่างของค่าความตรงตามสภาพดังตารางที่ 43

ตารางที่ 43 ค่าความตรงตามสภาพและการทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามอัตราการใช้ข้อสอบซ้ำ

อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)	จำนวนข้อมูลของผู้สอบ	ความตรงสภาพ		อัตราการใช้ข้อสอบซ้ำ (เปอร์เซ็นต์)		
		r	Z_r	10%	20%	30%
10%	672	0.6960	0.8595		0.3368	0.5610
20%	584	0.6860	0.8404			0.8688
30%	592	0.7120	0.8912			

จากตารางที่ 43 พบว่า เมื่ออัตราการใช้ข้อสอบซ้ำเป็น 10%, 20% และ 30% มีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ เท่ากับ 0.6960 0.6860 และ0.7120 ตามลำดับ และนำมาทดสอบความแตกต่างเป็นรายคู่ จึงแปลงค่าความตรงตามสภาพของการทดสอบให้เป็นคะแนนฟิชเชอร์ซี (Z_r) จากสูตรของ Glass and Stanley (1970) ได้ค่าเท่ากับ 0.8595 0.8404 และ 0.8912 ตามลำดับ พบว่า อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ และอัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ มีค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

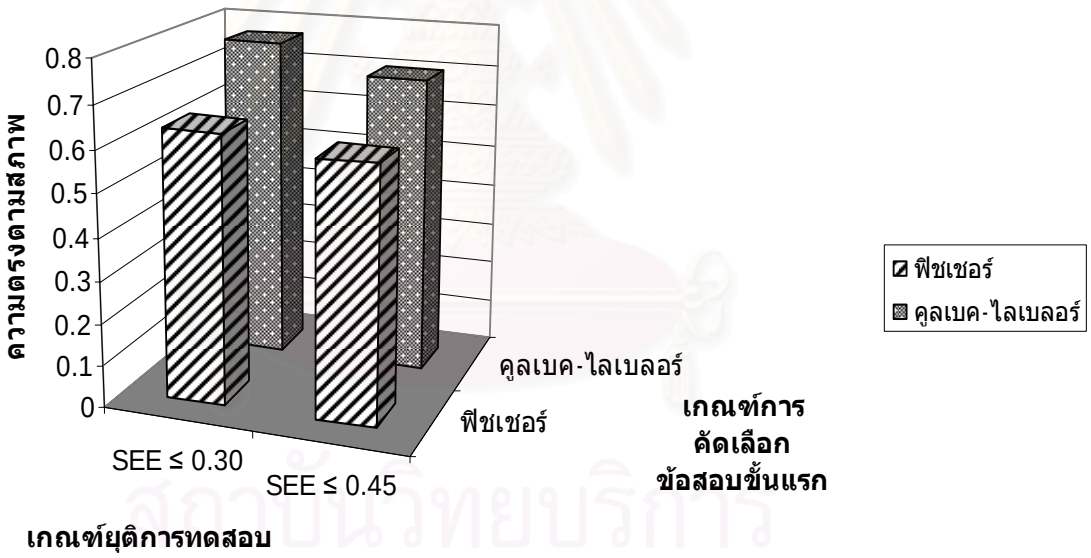
เมื่อใช้เกณฑ์ยุติการทดสอบเป็นเกณฑ์ในการศึกษาความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ได้ผลการคำนวณและการทดสอบความแตกต่างของค่าความตรงตามสภาพดังตารางที่ 44

ตารางที่ 44 ค่าความตรงตามสภาพและการทดสอบความแตกต่างของค่าความตรงตามสภาพในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์ยุติการทดสอบ

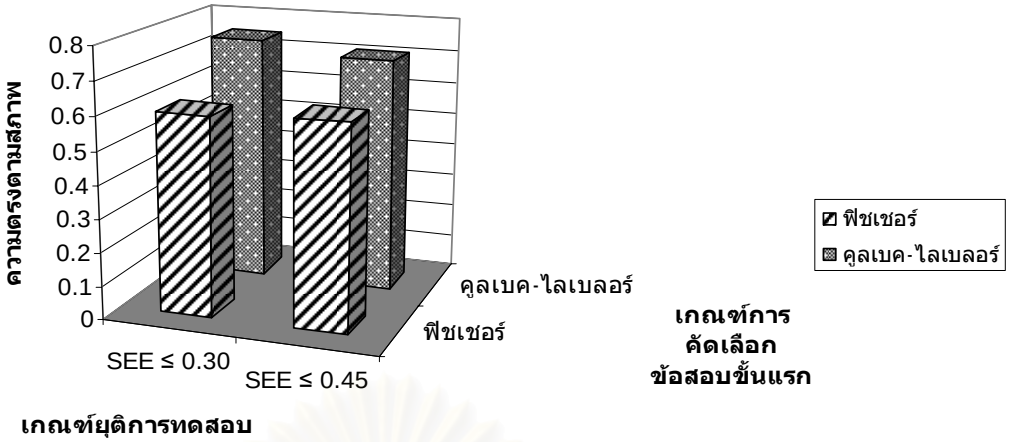
เกณฑ์ยุติการทดสอบ	จำนวนข้อมูลของผู้สอบ	ค่าความตรงตามสภาพ		Z
		r	Z_r	
SEE ≤ 0.30	924	0.7160	0.8994	1.4292
SEE ≤ 0.45	924	0.6820	0.8328	

จากตารางที่ 44 พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เท่ากับ 0.7160 0.6820 ตามลำดับ เมื่อแปลงค่าความตรงตามสภาพของการทดสอบให้เป็นคะแนนฟิชเชอร์ ซี (Z_r) จากสูตรของ Glass and Stanley (1970) ได้ค่าเท่ากับ 0.8994 0.8328 ตามลำดับ เพื่อนำมาทดสอบความแตกต่างเป็นรายคู่ พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 มีค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ไม่แตกต่างอย่างมีนัยสำคัญทางสถิติ จากเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45

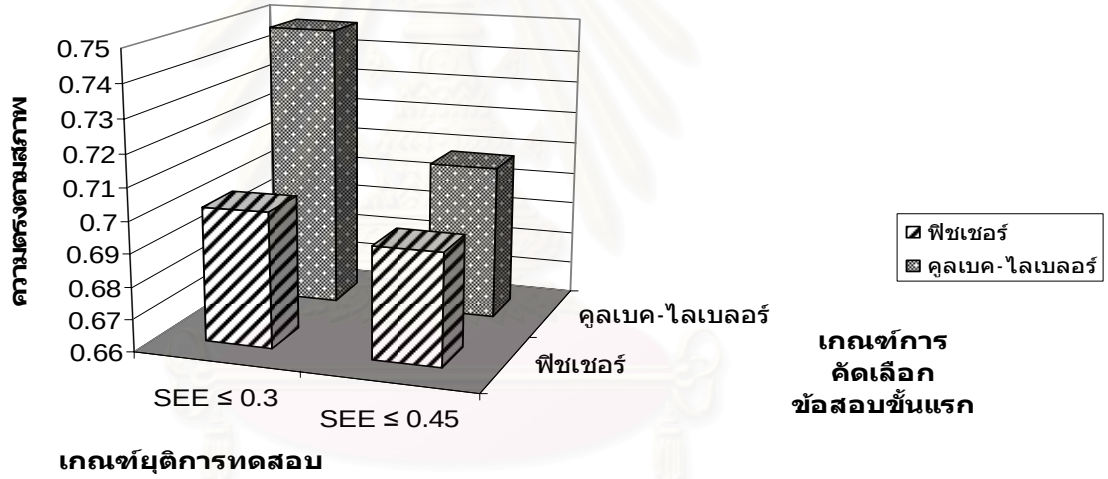
จากผลการเปรียบเทียบความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ที่กล่าวมาทั้งหมด สามารถสรุปได้ดังแผนภาพที่ 28 -31



แผนภาพที่ 28 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์ยุติการทดสอบและเกณฑ์การคัดเลือกข้อสอบขั้นแรกและอัตราการใช้สอบซ้ำ 10%



แผนภาพที่ 29 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์ยุติการทดสอบและเกณฑ์การคัดเลือกข้อสอบขั้นแรกและอัตราการใช้สอบซ้ำ 20%



แผนภาพที่ 30 ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามเกณฑ์ยุติการทดสอบและเกณฑ์การคัดเลือกข้อสอบขั้นแรกและอัตราการใช้สอบซ้ำ 30%

จากแผนภาพที่ 28 – 30 จะเห็นได้ว่า เมื่ออัตราการใช้ข้อสอบซ้ำ 10% 20% และ 30% พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ ให้ค่าความตรงตามสภาพของการทดสอบสูงกว่าวิธีสารสนเทศของฟิชเชอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ให้ค่าความตรงตามสภาพของการทดสอบสูงกว่า 0.45 แต่เมื่ออัตราการใช้ข้อสอบซ้ำ 20% เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ให้ค่าความตรงตามสภาพของการทดสอบสูงกว่า 0.30

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่ต่างกัน โดยใช้โปรแกรมการทดสอบแบบปรับกับระดับความสามารถของผู้สอบ ซึ่งพัฒนาโดย เกียรติศักดิ์ ส่องแสง (2547)

สมมติฐานการวิจัย คือ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกต่างกัน อัตราการใช้ข้อสอบซ้ำต่างกัน และเกณฑ์ยุติการทดสอบต่างกัน น่าจะทำให้ค่าความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แตกต่างกัน

ประชากรในการวิจัยครั้งนี้ คือ นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2549 ในโรงเรียนขนาดใหญ่พิเศษ สังกัดสำนักงานเขตพื้นที่การศึกษานครราชสีมา เขต 1 ซึ่งมีอยู่จำนวน 4 โรงเรียน จำนวนนักเรียน 1,982 คน และผู้วิจัยดำเนินการสุ่มโรงเรียนด้วยวิธีการสุ่มอย่างง่าย (Simple random sampling) ได้กลุ่มตัวอย่างในการวิจัยครั้งนี้ คือ นักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 3 โรงเรียน จำนวนนักเรียน 924 คน

เครื่องมือที่ใช้ในการวิจัย ประกอบด้วย 3 ประเภท ได้แก่ แบบสอบวัดความสามารถทางคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 6 ที่ผู้วิจัยพัฒนาขึ้นเอง และคลังข้อสอบวิชาคณิตศาสตร์ ใช้วัดความเข้าใจ เรื่อง เศษส่วน ระดับชั้นประถมศึกษาปีที่ 6 พัฒนาโดย รังสรรค์ มณีเล็ก (2540) และเกียรติศักดิ์ ส่องแสง (2547) และโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับระดับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ซึ่งพัฒนาโดย เกียรติศักดิ์ ส่องแสง (2547)

การเก็บรวบรวมข้อมูล โดยนำแบบสอบวัดความสามารถทางคณิตศาสตร์ทดสอบนักเรียนที่เป็นกลุ่มตัวอย่าง และทำการประมาณค่าความสามารถของผู้สอบด้วยโปรแกรม Loma IAS Version 11.0 จากนั้นดำเนินการสุ่มอย่างง่ายโดยสุ่มเงื่อนไขทั้ง 12 เงื่อนไขที่ใช้ในการวิจัยครั้งนี้ให้กับคอมพิวเตอร์แต่ละเครื่อง และทำการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำนวน 2 ครั้ง

การวิจัยในครั้งนี้ผู้วิจัยทำการวิเคราะห์ข้อมูลเบื้องต้น โดยการวิเคราะห์ค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าความสามารถของผู้สอบ จำนวนข้อสอบ และค่าฟังก์ชันสารสนเทศของแบบสอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ด้วยโปรแกรม SPSS for Windows version 15.0 และวิเคราะห์ค่าความตรงตามสภาพของการทดสอบ โดยการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าความสามารถเฉลี่ยที่ได้จากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์กับค่าความสามารถที่ได้จากการทดสอบด้วยแบบสอบวัดความสามารถทางคณิตศาสตร์ จำแนกตามเกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบ และเปรียบเทียบค่าความตรงตามสภาพของการทดสอบเป็นรายคู่ สถิติที่ใช้ในการคำนวณค่าความตรงตามสภาพและเปรียบเทียบค่าความตรงตามสภาพ ได้แก่ ค่าสหสัมพันธ์ ค่าไคสแควร์(Chi-Square) และ สถิติทดสอบซี (Z test)

สรุปผลการวิจัย

จากผลการวิเคราะห์ข้อมูล สามารถสรุปผลการวิจัยที่ได้จากการวิเคราะห์ค่าสถิติพื้นฐาน และสรุปผลการวิจัยตามวัตถุประสงค์ของการวิจัย ได้ดังนี้

1. ผลการวิเคราะห์ข้อมูลเบื้องต้นในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง

ผลการวิเคราะห์ค่าสถิติพื้นฐานโดยภาพรวม ของค่าความสามารถของผู้สอบ จำนวนข้อสอบ และค่าฟังก์ชันสารสนเทศของแบบสอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง พบว่า

ในทุกอัตราการใช้ข้อสอบซ้ำ10 เปอร์เซ็นต์ 20 เปอร์เซ็นต์ และ 30 เปอร์เซ็นต์

1. เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์จะให้การประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยเฉลี่ยสูงกว่าเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์
2. เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์จะใช้จำนวนข้อสอบโดยเฉลี่ยมากกว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ในจำนวนไม่มาก
3. เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ ให้ค่าฟังก์ชันสารสนเทศโดยเฉลี่ยสูงกว่าเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์

4. เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 ผู้วิจัยมีความเห็นว่าค่าความสามารถของผู้สอบโดยเฉลี่ยไม่ต่างกันมาก

5. เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ใช้จำนวนข้อสอบโดยเฉลี่ยมากกว่าเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ประมาณ 50 เปอร์เซ็นต์

6. เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ให้ค่าฟังก์ชันสารสนเทศของแบบสอบโดยเฉลี่ยสูงกว่าเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ประมาณ 50 เปอร์เซ็นต์

7. เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค - ไลเบลอร์และใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ที่อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ จะให้ค่าฟังก์ชันสารสนเทศของแบบสอบโดยเฉลี่ยสูงที่สุด

2. ผลการเปรียบเทียบค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์

จากการเปรียบเทียบความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่แตกต่างกัน พบว่า

1. เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกต่างกันในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่า ค่าความตรงตามสภาพของการทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ระหว่างเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ กับ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ โดยที่เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพสูงกว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์

2. เมื่อใช้อัตราการใช้ข้อสอบซ้ำที่ต่างกัน (10%, 20%, 30%)ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่า ค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

3. เมื่อใช้เกณฑ์ยุติการทดสอบที่ต่างกัน (SEE ≤ 0.30, SEE ≤ 0.45) ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่า ค่าความตรงตามสภาพของการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

4. เมื่ออัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ และ 20 เปอร์เซ็นต์ พบว่า ความตรงตามสภาพของการทดสอบระหว่างเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .01

5. ในอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ 20 เปอร์เซ็นต์ และ 30 เปอร์เซ็นต์ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพของการทดสอบสูงกว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์

6. ในอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ 20 เปอร์เซ็นต์ และ 30 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 มีค่าความตรงตามสภาพของการทดสอบสูงกว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 แต่เมื่ออัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ และเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 มีค่าความตรงตามสภาพของการทดสอบสูงกว่า 0.30

7. เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 พบว่า ความตรงตามสภาพของการทดสอบระหว่างเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .01

8. เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และอัตราการใช้ข้อสอบซ้ำ10 เปอร์เซ็นต์ และ 20 เปอร์เซ็นต์ พบว่า ความตรงตามสภาพของการทดสอบระหว่างเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และ วิธีสารสนเทศของคูลเบค-ไลเบลอร์ แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .05

9. เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่แตกต่างกัน พบว่า ค่าความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยที่เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อน

มาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ให้ค่าความตรงตามสภาพของการทดสอบสูงที่สุด

อภิปรายผล

การอภิปรายผลในงานวิจัยนี้นำเสนอ 2 ประเด็นหลัก คือ การอภิปรายผลที่ได้จากการวิเคราะห์ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบ จำนวนข้อสอบ และค่าฟังก์ชันสารสนเทศของแบบสอบในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง และการอภิปรายผลตามสมมติฐานการวิจัย มีรายละเอียด ดังนี้

1. การอภิปรายผลที่ได้จากการวิเคราะห์ค่าสถิติพื้นฐานของค่าความสามารถของผู้สอบ จำนวนข้อสอบ และค่าฟังก์ชันสารสนเทศของแบบสอบในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง

ในอัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ 20 เปอร์เซ็นต์ และ 30 เปอร์เซ็นต์ เป็นที่สังเกตได้ว่าเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ จะให้ค่าความสามารถของผู้สอบโดยเฉลี่ยสูงกว่า และให้ค่าฟังก์ชันสารสนเทศของแบบสอบโดยเฉลี่ยสูงกว่าเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ แต่เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ จะใช้จำนวนข้อสอบโดยเฉลี่ยมากกว่า วิธีสารสนเทศของคูลเบค-ไลเบลอร์ในจำนวนไม่มาก ซึ่งอาจเป็นไปได้ว่า การวิจัยในครั้งนี้มีการกำหนดจำนวนข้อสอบที่ใช้ในการทดสอบสูงสุดจำนวน 40 ข้อ จึงทำให้เกณฑ์การคัดเลือกข้อสอบขั้นแรกทั้ง 2 วิธี มีการใช้จำนวนข้อสอบไม่แตกต่างกันมากนัก ส่วนเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 จะใช้จำนวนข้อสอบโดยเฉลี่ยมากกว่าและให้ค่าฟังก์ชันสารสนเทศของแบบสอบโดยเฉลี่ยสูงกว่าเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ประมาณ 50 เปอร์เซ็นต์ ซึ่งสอดคล้องกับทฤษฎีการทดสอบที่ว่าถ้าจำนวนข้อสอบมีมากขึ้น ค่าความเที่ยงและความตรงก็จะเพิ่มขึ้น นั่นคือ ค่าความคลาดเคลื่อนมาตรฐานในการวัดก็จะลดลง (Gulliksen, 1967 ; Lord and Novick, 1968 อ้างถึงใน ต่าย เซี่ยงฉี, 2534)

2. การอภิปรายผลตามสมมติฐานการวิจัย

จากผลการวิจัยเกี่ยวกับการเปรียบเทียบความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบที่แตกต่างกัน มีประเด็นที่น่านำมาอภิปรายได้ดังนี้

1. การเปรียบเทียบความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกต่างกัน พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ มีค่าความตรงตามสภาพสูงกว่าเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ อาจเนื่องมาจากสูตรในการคำนวณเกณฑ์การคัดเลือกข้อสอบขั้นแรกแตกต่างกัน โดยสูตรการคำนวณเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ จะเป็นการคำนวณโดยการคัดเลือกข้อสอบที่มีค่าสารสนเทศที่สูงที่สุดเพียงข้อเดียว (The item information at point) ส่วนสูตรการคำนวณเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ จะเป็นการคำนวณโดยการคัดเลือกข้อสอบที่มีเกณฑ์สารสนเทศแบบช่วง (Interval information criterion) ซึ่งเริ่มต้นโดยการกำหนดช่วงในการอินทิเกรตทั้งหมดก่อน สำหรับงานวิจัยนี้ ผู้วิจัยกำหนดช่วงในการอินทิเกรตเป็นช่วงเดียวกับการสร้างสถานการณ์ในการจำลองข้อมูลในงานวิจัยของ Chen, Ankenmann and Chang (2000) ซึ่งมีช่วงในการอินทิเกรตดังนี้

$$\left(\hat{\theta}_l, \hat{\theta}_u\right) = \left(\hat{\theta} - \frac{3}{\sqrt{n}+1}, \hat{\theta} + \frac{3}{\sqrt{n}+1}\right)$$

และในการคัดเลือกข้อสอบข้อต่อไปโดยใช้เกณฑ์การคัดเลือกข้อสอบด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์จะเป็นการลดความกว้างของช่วงให้แคบลง โดยการคัดเลือกข้อสอบด้วยพื้นที่มากที่สุด (Maximum area) และข้อสอบข้อนั้นจะถูกคัดเลือกเป็นข้อต่อไป เพื่อให้ได้ข้อสอบที่เหมาะกับความสามารถของผู้สอบ รวมทั้งการที่ผู้สอบทำข้อสอบไปเรื่อย ๆ จนกระทั่งถึงเกณฑ์ยุติการทดสอบ จะเป็นการเพิ่มจำนวนข้อสอบที่ถูกใช้ และทำให้ช่วงของการทดสอบแคบลงเรื่อย ๆ จนกระทั่งลู่เข้าสู่ระดับความสามารถที่แท้จริงของผู้สอบได้ (Chen, Ankenmann and Chang, 2000) ซึ่งผลการวิจัยนี้สอดคล้องกับข้อค้นพบของ Chang and Ying (1996) พบว่า วิธีสารสนเทศของคูลเบค-ไลเบลอร์ (KL) มีประสิทธิภาพดีกว่าวิธีสารสนเทศของฟิชเชอร์ (FI) และChen, Ankenmann and Chang (2000) พบว่า วิธีสารสนเทศของฟิชเชอร์ที่มีการถ่วงน้ำหนัก (FII) วิธีสารสนเทศของฟิชเชอร์ที่มีการแจกแจงภายหลัง (FIP) วิธีสารสนเทศของคูลเบค-ไลเบลอร์ (KL) และวิธีสารสนเทศของคูลเบค-ไลเบลอร์ที่มีการแจกแจงภายหลัง (KLP) มีประสิทธิภาพและความแม่นยำในการประมาณค่าคุณลักษณะที่อยู่ปลายสุด (θ = -3,-2) มากกว่า วิธีสารสนเทศของฟิชเชอร์

(FI) รวมทั้ง Chang and Ying (1996) ให้ข้อเสนอแนะว่า ควรใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ (KL) กับคลังข้อสอบที่มีจำนวนข้อสอบไม่มาก และควรใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ (FI) กับคลังข้อสอบที่มีจำนวนข้อสอบมาก ซึ่งสอดคล้องกับข้อแนะนำของรังสรรค์ มณีเล็ก (2540) ได้แนะนำว่า เกณฑ์การคัดเลือกข้อสอบด้วยค่าสารสนเทศของข้อสอบสูงสุด (FI) เหมาะกับคลังข้อสอบขนาดใหญ่

2. การเปรียบเทียบความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่ออัตราการใช้ข้อสอบซ้ำต่างกัน พบว่า อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ 20 เปอร์เซ็นต์ และ 30 เปอร์เซ็นต์ มีค่าความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบไม่แตกต่างกัน ซึ่งผลการวิจัยนี้สอดคล้องกับข้อค้นพบของ Kalohn and Sparay (1998) พบว่า วิธีการคัดเลือกข้อสอบโดยมีอัตราการใช้ข้อสอบซ้ำ 10 และ 20 เปอร์เซ็นต์ มีความถูกต้องเหมือนกัน และผู้สอบที่มีความสามารถระดับสูงควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 10 เปอร์เซ็นต์ ของจำนวนข้อสอบที่ผู้สอบได้รับในการทดสอบครั้งที่ผ่านมา แต่ผลการวิจัยนี้ไม่สอดคล้องกับข้อค้นพบของ กนกวรรณ รัตนธน (2544) พบว่า อัตราการใช้ข้อสอบซ้ำที่ต่างกันส่งผลต่อความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบ และ Revuelta and Ponsoda (1998) พบว่า การที่ผู้สอบทำการทดสอบหลาย ๆ ครั้งด้วยการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คลังข้อสอบเดิม ส่งผลให้ความตรงตามสภาพของการทดสอบลดลง ทั้งนี้อาจเนื่องมาจาก การวิจัยในครั้งนี้เก็บรวบรวมข้อมูลกับผู้สอบคนเดิมซ้ำแค่ 2 ครั้ง อาจทำให้ไม่เห็นความแตกต่างว่า อัตราการใช้ข้อสอบซ้ำส่งผลต่อการประมาณค่าความสามารถของผู้สอบก็เป็นได้

3. การเปรียบเทียบความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์ยุติการทดสอบต่างกัน พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 มีความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบไม่แตกต่างกัน ซึ่งผลการวิจัยนี้ไม่สอดคล้องกับข้อค้นพบของ รังสรรค์ มณีเล็ก (2544) พบว่า เมื่อใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่างกันในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ทำให้ค่าความตรงตามสภาพของการทดสอบต่างกัน ทั้งนี้อาจมีผลมาจากโปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ที่ออกแบบการเก็บรวบรวมข้อมูลที่ใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบเท่ากับหรือน้อยกว่า 0.45 และ 0.30 ทำการเก็บรวบรวมข้อมูลและบันทึกข้อมูลที่เป็นสารสนเทศจากผู้สอบคนเดียวกัน ซึ่งในความเป็นไปได้ ความตรงตามสภาพของการทดสอบเกี่ยวข้องโดยตรงกับความถูกต้องแม่นยำในการวัด หรืออาจจะ

แปรเปลี่ยนไปตามค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ อย่างไรก็ตามหากพิจารณาถึงความสัมพันธ์ระหว่างความเที่ยงกับความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบตามสูตร $r_{tt} = 1 - \overline{SEE}^2$ เมื่อ $\overline{SEE}$ แทน ค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ (Warm, 1978 อ้างถึงใน ต่าย เซี่ยงฉี่, 2534 ; รังสรรค์ มณีเล็ก, 2540) จะพบว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบที่ 0.45 นี้จะทำให้ข้อสอบมีความเที่ยงในการวัดเท่ากับ 0.8127 นับว่าสูงน่าจะใช้แทนค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบที่ 0.30 ซึ่งทำให้ข้อสอบมีความเที่ยงในการวัดเท่ากับ 0.9133 นับว่าสูงมากได้

4. การเปรียบเทียบความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรก อัตราการใช้ข้อสอบซ้ำ และเกณฑ์ยุติการทดสอบต่างกัน พบว่า ค่าความตรงตามสภาพของการทดสอบต่างกัน โดยเมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 จะมีค่าความตรงตามสภาพของการทดสอบสูงสุด ซึ่งผลการวิจัยนี้ตรงกับสมมติฐานที่กำหนดไว้ และสอดคล้องกับข้อค้นพบของ Chang and Ying (1996) พบว่า เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค- ไลเบลอร์ (KL) มีประสิทธิภาพดีกว่าวิธีสารสนเทศของฟิชเชอร์ (FI) อาจเนื่องจากเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค- ไลเบลอร์ (KL) เป็นการคัดเลือกข้อสอบที่มีเกณฑ์สารสนเทศแบบช่วง (Interval information criterion) น่าจะเหมาะกับคลังข้อสอบที่มีจำนวนข้อสอบไม่มาก ซึ่งเหมือนกับเหตุผลในข้อที่ 1 และจากข้อค้นพบของ Stocking and Lewis (1998) พบว่าในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบแต่ละครั้งควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 20 เปอร์เซ็นต์ของจำนวนข้อสอบที่ผู้สอบได้รับในการสอบครั้งที่ผ่านมา รวมทั้ง Kalohn and Sparay (1998) พบว่า ผู้สอบที่มีความสามารถสูงควรมีอัตราการใช้ข้อสอบซ้ำไม่เกิน 10 เปอร์เซ็นต์ของจำนวนข้อสอบที่ผู้สอบได้รับในการทดสอบครั้งที่ผ่านมาซึ่งสอดคล้องกับงานวิจัยของกนกวรรณ รัตนธน (2544) พบว่า ผู้สอบที่มีความสามารถสูง และมีอัตราการใช้ข้อสอบซ้ำที่สูงขึ้นส่งผลให้ความตรงตามสภาพลดลง ซึ่งในการทดสอบแต่ละครั้งมีผู้สอบที่มีความสามารถหลายระดับ (สูง ปานกลาง ต่ำ) ดังนั้นถ้ามีการใช้ข้อสอบซ้ำควรกำหนดอัตราการใช้ข้อสอบซ้ำไม่เกิน10 เปอร์เซ็นต์ของจำนวนข้อสอบที่ผู้สอบได้รับในการทดสอบครั้งที่ผ่านมา รวมทั้งข้อค้นพบของรังสรรค์ มณีเล็ก (2540) พบว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ให้ค่าความตรงตามสภาพของการทดสอบสูงที่สุด เนื่องจากความตรงของแบบทดสอบจะขึ้นอยู่กับความถูกต้องแม่นยำในการวัด

ซึ่งในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบนั้น สามารถกำหนดค่าความถูกต้องในการวัดให้เป็นค่าเดียวกันสำหรับผู้สอบทุกคนได้ โดยใช้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถเป็นตัวกำหนด ถ้ากำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถลดต่ำลง ค่าความตรงก็น่าจะเพิ่มขึ้น

5. จากผลการวิจัยที่ พบว่า เมื่ออัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ และเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ ที่ใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 มีค่าความตรงตามสภาพสูงกว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ซึ่งไม่น่าจะเป็นไปได้ เนื่องจากเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 จะมีการใช้จำนวนข้อสอบที่มากกว่าเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ทำให้ค่าฟังก์ชันสารสนเทศของแบบสอบที่ใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 มีค่าสูงกว่าด้วย เป็นเพราะกำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่ำกว่า นั่นหมายความว่า ความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบที่ใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 สูงกว่า 0.45 ด้วย และความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบเกี่ยวข้องโดยตรงกับความตรงตามสภาพ ดังนั้น เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ควรจะมีความตรงตามสภาพสูงกว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 แต่ผลการวิจัยดังกล่าวอาจมีผลมาจากโปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ที่ออกแบบการเก็บรวบรวมข้อมูลที่ใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบเท่ากับหรือน้อยกว่า 0.45 และ 0.30 ทำการเก็บรวบรวมข้อมูลและบันทึกข้อมูลที่เป็นสารสนเทศจากผู้สอบคนเดียวกัน และผู้สอบที่เป็นกลุ่มตัวอย่างในกลุ่มเงื่อนไขที่ใช้เกณฑ์การคัดเลือกของสอบขั้นแรกด้วยวิธีของฟิชเชอร์ อัตราการใช้ข้อสอบซ้ำ 20% กับเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 และ 0.45 อาจตั้งใจทำข้อสอบในช่วงแรกที่มีการบันทึกผลในเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 แต่ช่วงหลังผู้สอบเริ่มทำข้อสอบในจำนวนมาก

ขึ้น และอาจเป็นเพราะผู้สอบบางคนมีความกังวลเรื่องสายตา (จากผลการประเมิน ภาคผนวก ซ) หรือผู้สอบอาจเริ่มเบื่อหน่ายในการทำข้อสอบ จึงไม่ตั้งใจทำข้อสอบ อาจส่งผลทำให้ความตรงตามสภาพที่ใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 มีค่าสูงกว่า เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ก็เป็นได้

6. ข้อสังเกตในขณะดำเนินการทำวิจัย ผู้วิจัยพบว่า เมื่ออัตราการใช้ข้อสอบซ้ำ 10, 20 และ 30 เปอร์เซ็นต์ เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ และเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ผู้สอบจะถูกคัดชื่อออกจากกลุ่มตัวอย่างที่นำผลมาวิเคราะห์ข้อมูลมาก เนื่องจากข้อสอบที่ถูกคัดเลือกมาจากเกณฑ์ดังกล่าวชุดนี้อาจไม่เหมาะกับความสามารถของผู้สอบ จึงคัดเลือกข้อสอบที่มีคุณลักษณะใกล้เคียงขึ้นมาแทน นอกจากนี้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 จะใช้ข้อสอบในจำนวนมากกับผู้สอบที่มีระดับความสามารถสูงมากและต่ำมากจนหาข้อสอบที่ยุติการทดสอบไม่ได้ ในขณะเดียวกันเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค-ไลเบลอร์ไม่ค่อยพบปัญหาดังกล่าว ทั้งนี้อาจเนื่องมาจากหลักการของเกณฑ์การคัดเลือกข้อสอบขั้นแรกของทั้ง 2 วิธีต่างกัน ขนาดของคลังข้อสอบและค่าพารามิเตอร์ของข้อสอบที่อยู่ในคลังข้อสอบที่นำมาคำนวณกับเกณฑ์การคัดเลือกข้อสอบข้อแรกและข้อต่อไปอาจไม่เหมาะกับความสามารถของผู้สอบก็เป็นได้

ข้อเสนอแนะ

จากผลการวิจัยที่ได้ ผู้วิจัยมีข้อเสนอแนะสำหรับผู้ที่จะนำผลการวิจัยนี้ไปใช้ และผู้ที่ต้องการศึกษาเกี่ยวกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ต่อไป ดังนี้

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

การพิจารณาตัวแปรเพื่อที่จะนำไปใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ให้เกิดประโยชน์และมีประสิทธิภาพสูงสุดนั้น ผู้วิจัยขอเสนอผลของตัวแปรที่ได้จากข้อค้นพบในการวิจัยครั้งนี้

1. จากผลการวิจัยแสดงให้เห็นว่า ค่าความตรงตามสภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบแตกต่างกัน เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกต่างกัน โดยที่เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ มีค่าความตรงตามสภาพสูงกว่าเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ ดังนั้นในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ควรใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ หากมีอัตราการใช้ข้อสอบซ้ำก็ควรไม่เกิน 10 เปอร์เซ็นต์สำหรับการทดสอบโดยทั่วไป ซึ่งมีผู้เข้าสอบที่มีระดับความสามารถที่แตกต่างกัน และควรใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 จึงทำให้การทดสอบมีความถูกต้องและมีความตรงตามสภาพมากที่สุด

2. โรงเรียนในเขตพื้นที่การศึกษาเดียวกันหรือจังหวัดเดียวกันน่าจะร่วมกันสร้างคลังข้อสอบขึ้นและนำการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบมาใช้ เพื่อความสะดวกในการทดสอบ เนื่องจากเป็นการทดสอบรายบุคคลและการทดสอบแต่ละครั้งไม่จำเป็นต้องสอบพร้อมกันทุกคน เมื่อผู้สอบที่ต้องการจะสอบก็สามารถทำการทดสอบได้ด้วยตนเอง ผลการทดสอบแต่ละครั้งจะถูกบันทึกไว้ และจำนวนเครื่องคอมพิวเตอร์ไม่เป็นปัญหาสำหรับการทดสอบลักษณะนี้ คลังข้อสอบที่สร้างขึ้นถ้ามีจำนวนข้อสอบในคลังข้อสอบไม่มากควรใช้วิธีสารสนเทศของคูลเบค-ไลเบลอร์เป็นเกณฑ์ในการคัดเลือกข้อสอบข้อแรกและข้อต่อไป แต่ถ้ามีจำนวนข้อสอบในคลังข้อสอบมากควรใช้วิธีสารสนเทศของฟิชเชอร์เป็นเกณฑ์ในการคัดเลือกข้อสอบข้อแรกและข้อต่อไป สำหรับเกณฑ์ยุติการทดสอบถ้าเป็นการทดสอบเพื่อติดตามความก้าวหน้าทางการเรียนของนักเรียนน่าจะใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 เนื่องจากใช้จำนวนข้อสอบน้อยกว่าและค่าความเที่ยงและความตรงตามสภาพก็ไม่ต่างกันมากกับเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 แต่ถ้าเป็นการทดสอบที่ต้องการความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบสูง เช่น การทดสอบเพื่อคัดเลือกบุคคล ควรใช้เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 เนื่องจากให้ค่าความเที่ยงและความตรงตามสภาพสูงกว่า

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

เนื่องจากค่าความตรงตามสภาพในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แตกต่างกัน เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกต่างกัน ดังนั้นผู้วิจัยจึงมีข้อเสนอแนะในการวิจัยครั้งต่อไปดังนี้

1. ควรจะมีการศึกษาเกี่ยวกับผลของตัวแปรอื่น ๆ เช่น การแจกแจงค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบ ขนาดของคลังข้อสอบ หรือจำนวนข้อสอบในแต่ละระดับความสามารถ ว่ามีผลทำให้เกณฑ์การคัดเลือกข้อสอบขั้นแรกแตกต่างกันหรือไม่ อย่างไร

2. ควรมีการศึกษารูปแบบการจัดข้อสอบให้กับผู้สอบในลักษณะของกลุ่มข้อสอบ (Testlet) นั้นคือ ในแต่ละขั้นของการสอบแทนที่จะเลือกข้อสอบในกลุ่มที่จัดให้หมดแล้วจึงเลือกข้อสอบกลุ่มใหม่ให้เหมาะสมกับผู้สอบที่ใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกต่างกัน

3. ควรมีการศึกษาเกี่ยวกับวิธีการควบคุมอัตราการใช้ข้อสอบซ้ำ และวิธีการสร้างความสมดุลของเนื้อหาว่าวิธีใดควรใช้กับเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีไหนถึงจะทำให้การประมาณค่าความสามารถของผู้สอบมีความถูกต้องแม่นยำที่สุด

4. ควรมีการศึกษาเกี่ยวกับเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบที่คำนวณจากการถอดรากที่สองของค่าความแปรแปรวนในการทดสอบกับที่คำนวณจากสูตร $SEE = \frac{1}{\sqrt{I(\theta)}}$ เมื่อใช้เกณฑ์การคัดเลือกข้อสอบที่ขั้นแรกต่างกันว่าสูตรเกณฑ์ยุติการทดสอบสูตรใดส่งผลทำให้ประสิทธิภาพของการทดสอบแตกต่างกันหรือไม่ อย่างไร และมีผลต่อจำนวนข้อสอบที่ใช้ในการทดสอบหรือไม่ อย่างไร

5. ควรมีการศึกษาเกี่ยวกับเกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบกับที่กำหนดให้จำนวนข้อสอบคงที่เป็นเกณฑ์ยุติการทดสอบว่าเกณฑ์ใดมีผลต่อการประมาณค่าความสามารถของผู้สอบแตกต่างกันหรือไม่ อย่างไร เมื่อใช้เกณฑ์การคัดเลือกข้อสอบที่ขั้นแรกต่างกัน

6. จากการที่ผู้วิจัยให้นักเรียนที่เป็นกลุ่มตัวอย่างทำแบบประเมินความพึงพอใจของผู้ใช้โปรแกรม พบว่า มีนักเรียนที่เป็นกลุ่มตัวอย่างประมาณ 4.32% ให้ข้อเสนอแนะเกี่ยวกับการปรับปรุงแก้ไขเพิ่มเติมเกี่ยวกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบว่า ไม่สามารถข้ามข้อสอบที่ทำไม่ได้ไปทำข้อถัดไปได้ ทำให้เสียเวลา ดังนั้นควรมีการศึกษาวิธีที่สามารถทำให้ผู้สอบข้ามข้อสอบไปทำข้ออื่นหรือย้อนกลับไปทวนคำตอบได้

รายการอ้างอิง

ภาษาไทย

กนกวรรณ รัตนธน. (2544). **การเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อความสามารถของผู้สอบ และอัตราการใช้ข้อสอบซ้ำแตกต่างกัน**. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

เกียรติศักดิ์ ส่องแสง. (2547). **ฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์**. ปริญญานิพนธ์การศึกษาศาสตรดุษฎีบัณฑิต.มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

คณะกรรมการการศึกษาแห่งชาติ, สำนักงาน. (2542). **พระราชบัญญัติการศึกษาแห่งชาติ พุทธศักราช 2542.** กรุงเทพมหานคร.

คณิต ไข่มุกด์. (2534). **การพัฒนาเกณฑ์ในการคัดเลือกข้อสอบที่ปรับเหมาะกับความสามารถของผู้สอบ.** วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต. จุฬาลงกรณ์มหาวิทยาลัย.

ตรึงใจ พูลผลอำนวย. (2534). **การพัฒนาแบบสอบเฉพาะบุคคลในวิชาวิทยาศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3.** วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ต่าย เซี่ยงฉี. (2534). **การศึกษาเปรียบเทียบประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบเทเลอร์รูปปิรามิดที่มี รูปแบบ จำนวนขั้น และวิธีการให้คะแนน ที่แตกต่างกันโดยใช้วิธีมอนติคาร์โล.** ปริญญานิพนธ์การศึกษาศาสตรดุษฎีบัณฑิต. มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

นันทิยา พึ่งคำ. (2531). **การเปรียบเทียบคุณภาพของการทดสอบ ซีเอที และแบบสอบประเพณีนิยมในการวัดความสามารถด้านคำศัพท์ภาษาอังกฤษ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3.** วิทยานิพนธ์ปริญญามหาบัณฑิต. จุฬาลงกรณ์มหาวิทยาลัย.

บุญเชิด ภิญโญอนันตพงษ์. (2544). **การประเมินการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ : แนวคิดและวิธีการ.** กรุงเทพมหานคร : โรงพิมพ์อมรินทร์พริ้นติ้งแอนด์พับลิชชิ่ง.

ปราโมทย์ เดชะอำไพ. (2544). **ระเบียบวิธีเชิงตัวเลขในงานวิศวกรรม.** กรุงเทพมหานคร : สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

รังสรรค์ มณีเล็ก. (2540). **ผลของตัวแปรบางตัวต่อความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์.**ปริญญานิพนธ์การศึกษาศาสตรดุษฎีบัณฑิต.มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

วสันต์ ทองไทย. (2539). **การเปรียบเทียบผลการประมาณค่าความสามารถด้วยวิธีการทดสอบแบบเทเลอร์รูปพีระมิดขนาดขั้นคงที่ และรูปพีระมิดข้างตัด : การทดสอบทักษะกระบวนการทางวิทยาศาสตร์.** ปริญญาศึกษาศาสตรมหาบัณฑิต. มหาวิทยาลัยเกษตรศาสตร์.

ศิริชัย กาญจนวาสี. (2538).**การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ**. จุลสารการทดสอบอันดับที่ 1. ศูนย์ทดสอบทางการศึกษาร่วมกับศูนย์ตำราและเอกสารทางวิชาการ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2545). **ทฤษฎีการทดสอบแนวใหม่ (Modern test theories).** พิมพ์ครั้งที่ 2. กรุงเทพฯ : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2548). **ทฤษฎีการทดสอบแบบดั้งเดิม (Classical test theories).** พิมพ์ครั้งที่ 5. กรุงเทพฯ : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

สายชล อบทม. (2539). **การพัฒนาโปรแกรมสำหรับการทดสอบแบบปรับเหมาะโดยคอมพิวเตอร์**. วิทยานิพนธ์ปริญญามหาบัณฑิต. จุฬาลงกรณ์มหาวิทยาลัย.

สำนักงานเขตพื้นที่การศึกษานครราชสีมา เขต 1. (2549). **ข้อมูลสารสนเทศ** [ออนไลน์]. แหล่งที่มา: http://www.edkorat1.info

สุพัฒน์ สุกมลสันต์. (2539). **ธนาคารข้อสอบและการทดสอบปรับเปลี่ยนด้วยคอมพิวเตอร์**. กรุงเทพ เทพมหานคร : บริษัทพิมพ์ดี จำกัด.

โสพส สุขานนท์สวัสดิ์. (2545).**การพัฒนาโปรแกรมสำหรับการทดสอบแบบปรับเหมาะโดยคอมพิวเตอร์**. วิทยานิพนธ์ปริญญาการศึกษามหาบัณฑิต. มหาวิทยาลัยนเรศวร.

อุทุมพร (ทองอุไร) จามรมาน. (2535). **ข้อสอบ: การสร้างและพัฒนา**. กรุงเทพมหานคร : โรงพิมพ์พับบลิชชิ่ง.

เอกรินทร์ สี่มหาศาล. (2546). **กระบวนการจัดทำหลักสูตรสถานศึกษา : แนวคิดสู่ปฏิบัติ.** กรุงเทพมหานคร : บุ๊คพอยท์.

ภาษาอังกฤษ

Blackmore, L.M. (1987). “Computerized, computerized adaptive and pencil-and-paper test administration : A comparative study in high school setting.” **Dissertation Abstracts International**. 47(07) : 2554 – A.

Chang, H.-H., and Ying, Z. (1996). “A global information approach to computerized adaptive testing.” **Applied Psychological Measurement**. 20 : 213-229.

Chang, S. (1998). “A comparative study of item exposure control methods in computerized adaptive testing (Act, Test Security).” **Dissertation Abstracts International**. 59(09): 3418.

Chang, S., and Ankenmann,R.D (2004). “Effects of practical constraints on item selection rules at the early stages of computerized adaptive testing.” **Journal of Educational Measurement**. 41,149-174.

Chang, S., Ankenmann, R. D., and Spray, J.A.(2003). “The relationship between item exposure and test overlap in computerized adaptive testing.” **Journal of Educational Measurement**. 40, 129-145.

Chen, Ankenmann and Chang (2000). “A comparison of item selection rules at the early stages of computerized adaptive testing**.” Applied Psychological Measurement**. 24, 241-255

Chen, S. Y. and Lei, P. W. (2005). “Controlling item exposure and test overlap in computerized adaptive testing.” **Applied Psychological Measurement**. 29(3): 204-217.

De Ayala, R.J., Dodd, G.B., and Koch, R.W. (1990). “A simulation and comparison of flexilevel and Bayesian computerized adaptive testing.” **Journal of Educational Measurement**. 27 : 227-239

Dodd, B.G., Koch, R.W. and De Ayala, R.J. (1993). “Computerized adaptive testing using the partial credit model: Effect of Item pool characteristics and different stopping rules.” **Educational and Psychological Measurement.** 53 : 61-77

Fan, Xitao. (1998). “Item response theory and classical test theory : an empirical comparison of their item/person statistics.” **Educational and Psychological Measurement.** 58 : 357-381

Gerahon, R. C. (2005). “Computer adaptive testing.” **Journal of Applied Measurement**. 6(1), 109-127.

Gifford,J.A. and Swaminathan, H. (1990). “Bias and the effect priors in bayesian estimation of parameters of item response model.” **Applied Psychological Measurement**. 14(1) : 33-34.

Glass, G.V. and Stanley, J.C. (1970). **Statistical methods in education and psychology.** Englewood Cliffs,NJ . : Prentice – Hall.

Gushta, M.M. (2003). **Standard-setting Issues in Computerized-Adaptive Testing**. Paper prepared for presentation at the Annual Conference of the Canadian Society for studies in Education, Halifax, Nova Scotia.

Hair, J.F. JR., Anderson, R. E., Tatham, R. L., and Black, W.C. (1998). **Multivariate data analysis** (5th ed.). Upper Saddle River, NJ : Prentice-Hall,Inc.

Hambleton, R.K. and Swaminathan, H. (1985). **Item response theory : Principles and application**. Boston : Kluner-Nijhoff Publishing.

Hambleton, R.K., Swaminathan,H. and Rogers, H.J. (1991). **Fundamentals of item response theory**. Newbury Park CA : Sage.

Ho, R.G. (1989). "Using MicroCAT in computerized adaptive testing : A comparison of three adaptive testing strategies." **Dissertation Abstracts International**. 50(20) : 421-A.

Hol, A.M., Vorst, H.C.M. and Mellenbergh, G.J. (2005). "A randomized experiment to compare convention, computerized, and computerized adaptive administration of ordinal polytpmous attitude items." **Applied Psychological Measurement**. 29(3) : 159-183.

Hulin, C.L., Darsgow.F and Parsons. (1983). **Item response theory : Application to psychological measurement**, Homewood, IL : dow-Jones.

Kalohn, J.C. and Spray, J.A. (1998). **Test security and item exposure control for computer-based examination : Performance of a computerized classification test for professional certification**.

Leung, C. K., Chang, H. H., and Hua, K. T. (2003). "Incorporation of content balancing requirement in stratification designs for computerized adaptive testing." **Educational and Psychological Measurement**. 63(2) : 257-270.

Leung, C. K., Chang, H. H., and Hua, K. T. (2003). "Computerized adaptive testing: A comparison of three content balancing methods**." Journal of Technology, Learning, and Assessment**, 2(5). Available from http://www.jtla.org

Lord, F.M. and Novick, M.R. (1968). **Statistical theories of mental test score**. Massachusetts : Addison Wesley Publishing Company.

Lord, F.M. and Novick, M.R. (1980). **Applications of item response theory to practical testing problem**. New Jersey : Lawrence Erlbaum Associates.

Revuelta, J. and Ponsoda, V. (1998). "A comparison of item exposure control methods in computerized adaptive testing." **Journal of Educational Measurement**. 34(4) : 311-327

Veekamp, W.J., and Berger,M.P. (1997). “Some new item selection criteria for adaptive testing.” **Journal of Educational and Behavioral Statistics**. 22 : 203-226.

Urry, V.W. (1977). “ Tailored testing : A successful application of latent trait theory.” **Journal of Education Measurement**. 14(2) : 181-196

Weiss, D.J. (1988). “Adaptive testing.” **Educational Research Methodology and Measurement : International Handbook**. New York : Pergamon Press : 372-376.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก

รายชื่อผู้เชี่ยวชาญในการตรวจสอบคุณภาพเครื่องมือ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก
รายชื่อผู้เชี่ยวชาญในการตรวจสอบคุณภาพ
แบบสอบวัดความสามารถทางคณิตศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 6

1. รองศาสตราจารย์ ดร. สิริพร ทิพย์คง	ภาควิชาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์
2. ผู้ช่วยศาสตราจารย์ ดร. สมยศ ชิดมงคล	ภาควิชาหลักสูตร การสอน และ เทคโนโลยีการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
3. ผู้ช่วยศาสตราจารย์ ดร. อวยพร เรืองตระกูล	ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
4. อาจารย์ ดร. กมลวรรณ ตังธนกานนท์	ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
5. อาจารย์ ดร. ยุรวัฒน์ คล้ายมงคล	ภาควิชาหลักสูตร การสอน และ เทคโนโลยีการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
6. นายชยุตม์ ภิรมย์สมบัติ	นักวิชาการสาขาประเมินมาตรฐาน สถานบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี
7. อาจารย์มงคล วงศ์พยัคฆ์	ภาควิชาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์
8. อาจารย์ปิยะ บัณฑิตสกุลชัย	สาขาวิชาวิทยาศาสตร์และคณิศาสตร์ คณะเทคโนโลยีสารสนเทศ วิทยาลัยนอร์ท กรุงเทพ
9. อาจารย์วงเดือน สอนประภา	โรงเรียนเขมะสิริอนุสสรณ์
10. อาจารย์จันทรวรรณ อัตถวิบูลย์กุล	โรงเรียนเขมะสิริอนุสสรณ์
11. อาจารย์ศิริวรรณ แก้วฟอง	โรงเรียนบ้านคลองกรวด
12. อาจารย์เย็นจิต ติ้วสิขเรศ	โรงเรียนปลูกปัญญา

รายชื่อผู้เชี่ยวชาญในการตรวจสอบคุณภาพ โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์

1. ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี

   อาจารย์ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ตรวจสอบความถูกต้องของสูตรเกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์และคูลเบค-ไลเบลอร์

2. ดร. รังสรรค์ มณีเล็ก

   ผู้เชี่ยวชาญด้านวิเคราะห์วิจัยการจัดการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ ผู้มีประสบการณ์ในการทำวิทยานิพนธ์เกี่ยวการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ตรวจสอบความถูกต้องของสูตรการประมาณค่าความสามารถโดยวิธี Bayesian updating

3. ดร. เกียรติศักดิ์ ส่องแสง

   อาจารย์ประจำภาควิชาวิจัยและประเมินผลการศึกษา คณะศึกษาศาสตร์ สถาบันเทคโนโลยีราชมงคล ผู้มีประสบการณ์ในการทำวิทยานิพนธ์เกี่ยวการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ตรวจสอบความถูกต้องและประสิทธิภาพการทำงานของโปรแกรมการทดสอบ

4. อาจารย์กนกวรรณ รัตนธน

   อาจารย์สอนคอมพิวเตอร์โรงเรียนธรรมโชติศึกษาลัย ผู้มีประสบการณ์ในการทำวิทยานิพนธ์เกี่ยวกับตัวแปรอัตราการใช้ข้อสอบซ้ำ และตรวจสอบความถูกต้องของอัตราการใช้ข้อสอบซ้ำและประสิทธิภาพการทำงานของโปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

5. นายโสพส สุขานนท์สวัสดิ์

   เจ้าหน้าที่ระบบงานคอมพิวเตอร์ การไฟฟ้านครหลวงไทย และมีประสบการณ์ในการทำวิทยานิพนธ์เกี่ยวกับการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ตรวจสอบประสิทธิภาพการทำงานของโปรแกรมการทดสอบ

6. นายวรรณแสงชัย แซ่เฮง

   เจ้าหน้าที่ระบบงานคอมพิวเตอร์ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย และมีประสบการณ์ในการเขียนโปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ตรวจสอบประสิทธิภาพการทำงานของโปรแกรมการทดสอบ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

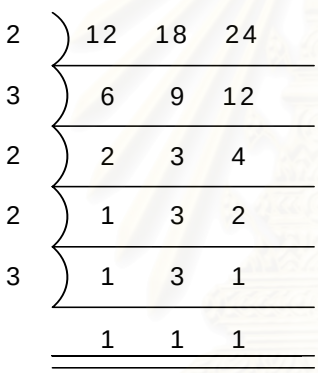
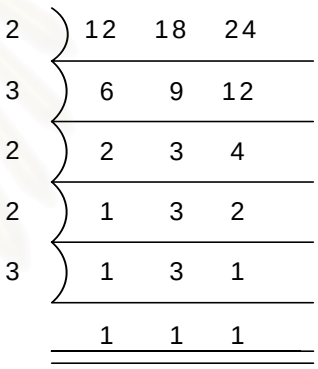
ภาคผนวก ข

การปรับปรุง แก้ไข เครื่องมือที่ใช้ในการวิจัย

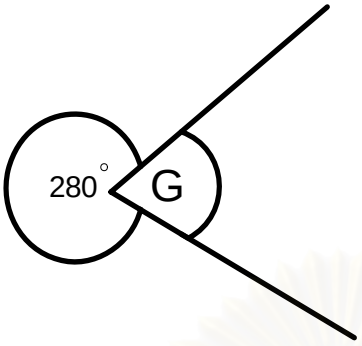
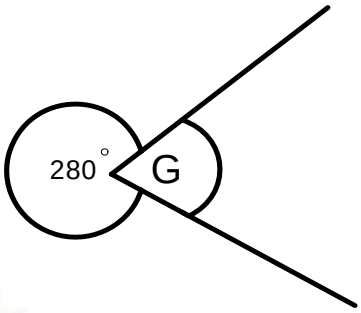
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ข
การปรับปรุง แก้ไข เครื่องมือที่ใช้ในการวิจัย
แบบสอบวัดความสามารถทางคณิตศาสตร์

ข้อที่	ข้อความเดิม	ข้อความหลังปรับปรุง
1	ข้อใดต่อไปนี้เมื่อประมาณเป็นค่าใกล้เคียงจำนวนเต็มร้อยแล้วมีค่าต่างไปจากข้ออื่น ก. 654 ข. 732 ค. 645 ง. 723	ค่าประมาณใกล้เคียงจำนวนเต็มร้อยของจำนวนในข้อใดต่อไปนี้แตกต่างไปจากข้ออื่น ก. 654 ข. 732 ค. 645 ง. 723
4	"เด็กชายมานะมีเงินเก็บ 287 บาท ได้รับรางวัลจากการประกวดคัดลายมือเป็นเงิน 150 บาท เด็กชายมานะต้องเก็บเงินเพิ่มอีกกี่บาท จึงจะซื้อนาฬิการาคา 579 บาทได้" โจทย์ข้อนี้ เขียนเป็นประโยคสัญลักษณ์ได้ดังข้อใดต่อไปนี้ ก. 579 - (287 + 150) = ☐ ข. (287 + 150) - 579 = ☐ ค. 579 - 287 +150 = ☐ ง. 579 - (287 - 150) = ☐	"เด็กชายมานะมีเงินเก็บ 287 บาท ได้รับรางวัลจากการประกวดคัดลายมือเป็นเงิน 150 บาท เด็กชายมานะต้องเก็บเงินเพิ่มอีกกี่บาท จึงจะซื้อนาฬิการาคา 579 บาทได้" เขียนเป็นประโยคสัญลักษณ์ได้ ดังข้อใดต่อไปนี้ ก. 579 - (287 + 150) = ☐ ข. (287 + 150) - 579 = ☐ ค. 579 - 287 +150 = ☐ ง. 579 - (287 - 150) = ☐
5	"มีส้มโอ 986 ผล ครั้งแรกขายไป 189 ครั้งที่สองขายได้มากกว่าครั้งแรก 99 จะเหลือส้มโอกี่ผล โจทย์ข้อนี้ เขียนเป็นประโยคสัญลักษณ์ได้ ดังข้อใดต่อไปนี้ ก. $986-189-99=\square$ ข. $986-189+99=\square$ ค. $986-189-(189+99)=\square$ ง. $986-189-(189-99)=\square$	"มีส้มโอ 986 ผล ครั้งแรกขายไป 189 ครั้งที่สองขายได้มากกว่าครั้งแรก 99 จะเหลือส้มโอกี่ผล เขียนเป็นประโยคสัญลักษณ์ได้ ดังข้อใดต่อไปนี้ ก. $986-189-99=\square$ ข. $986-189+99=\square$ ค. $986-189-(189+99)=\square$ ง. $986-189-(189-99)=\square$

ข้อที่	ข้อความเดิม	ข้อความหลังปรับปรุง
6	"ชัยชนะซื้อข้าวสาร 9 ถุง ราคาถุงละ 115 บาท ให้เงินไป 1,500 บาท จะได้รับเงินทอนกี่บาท" โจทย์ข้อนี้ เขียนเป็นประโยคสัญลักษณ์ได้ ดังข้อใดต่อไปนี้ ก. $1{,}500-(115\times 9)=\square$ ข. $1{,}500-(115\div 9)=\square$ ค. $(1{,}500\div 9)-115=\square$ ง. $(1{,}500\times 9)-115=\square$	"ชัยชนะซื้อข้าวสาร 9 ถุง ราคาถุงละ 115 บาท ให้เงินไป 1,500 บาท จะได้รับเงินทอนกี่บาท" เขียนเป็นประโยคสัญลักษณ์ได้ ดังข้อใดต่อไปนี้ ก. $1{,}500-(115\times 9)=\square$ ข. $1{,}500-(115\div 9)=\square$ ค. $(1{,}500\div 9)-115=\square$ ง. $(1{,}500\times 9)-115=\square$
22	 จากที่กำหนดให้ เป็นวิธีการหาสิ่งใด ก. ห.ร.ม. ข. ค.ร.น. ค. แยกตัวประกอบ ง. ห.ร.ม. และ ค.ร.น.	 จากขั้นตอนการหารที่กำหนดให้ เป็นวิธีการหาสิ่งใด ก. ห.ร.ม. ข. ค.ร.น. ค. แยกตัวประกอบ ง. ห.ร.ม. และ ค.ร.น.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ข้อที่	ข้อความเดิม	ข้อความหลังปรับปรุง
24	 จากรูปที่กำหนดให้ มุม G มีขนาดมุมกี่องศา สามารถหาคำตอบได้ตามข้อใดต่อไปนี้ ก. $280° - 180°$ ข. $280° + 180°$ ค. $360° + 280°$ ง. $360° - 280°$	 จากรูปที่กำหนดให้ สามารถหาค่าของมุม G ได้ตามข้อใดต่อไปนี้ ก. $280° - 180°$ ข. $280° + 180°$ ค. $360° + 280°$ ง. $360° - 280°$
34	ข้อใดต่อไปนี้แปลงเป็นจำนวนคละได้ ก. $\frac{3}{9}$ ข. $\frac{4}{5}$ ค. $\frac{6}{7}$ ง. $\frac{9}{5}$	ข้อใดต่อไปนี้สามารถทำให้เป็นเป็นจำนวนคละได้ ก. $\frac{3}{9}$ ข. $\frac{4}{5}$ ค. $\frac{6}{7}$ ง. $\frac{9}{5}$
36	ข้อใดต่อไปนี้มีค่าเท่ากับ $\frac{A}{B} \times \frac{C}{D} \times \frac{B}{C}$ ก. $\frac{1}{D}$ ข. 1 ค. $\frac{C}{B}$ ง. $\frac{A}{D}$	$\frac{A}{B} \times \frac{C}{D} \times \frac{B}{C}$ มีค่าเท่ากับข้อใดต่อไปนี้ ก. $\frac{1}{D}$ ข. 1 ค. $\frac{C}{B}$ ง. $\frac{A}{D}$

ข้อที่	ข้อความเดิม	ข้อความหลังปรับปรุง
37	ข้อใดต่อไปนี้มีค่าเท่ากับ $\dfrac{\frac{A}{B}\times\frac{B}{C}}{\frac{A}{B}\div\frac{C}{B}}$ ก. 0 ข. 1 ค. $\dfrac{1}{BC}$ ง. $\dfrac{A}{BC}$	$\dfrac{\frac{A}{B}\times\frac{B}{C}}{\frac{A}{B}\div\frac{C}{B}}$ มีค่าเท่ากับข้อใดต่อไปนี้ ก. 0 ข. 1 ค. $\dfrac{1}{BC}$ ง. $\dfrac{A}{BC}$
38	ข้อใดต่อไปนี้ควรเป็นขั้นตอนแรกในการหาคำตอบของ $5\frac{8}{9}+\frac{5}{7}\div1\frac{5}{8}=[\ ]$ ก. หา ค.ร.น. ข. กลับเศษเป็นส่วน ค. เปลี่ยนเครื่องหมายหารเป็นคูณ ง. แปลงจำนวนคละให้เป็นเศษเกิน	ข้อใดต่อไปนี้ควรเป็นขั้นตอนแรกในการหาคำตอบของ $5\frac{8}{9}+\frac{5}{7}\div1\frac{5}{8}=[\ ]$ ก. หา ค.ร.น. ข. กลับตัวเศษเป็นตัวส่วน ค. เปลี่ยนเครื่องหมายหารเป็นเครื่องหมายคูณ ง. แปลงจำนวนคละให้เป็นเศษเกิน
39	“น้ำตาลหนัก 20 กิโลกรัม ตักใส่ถุงเพื่อแบ่งถุงละ $\frac{4}{5}$ กิโลกรัม จะได้กี่ถุง” โจทย์ข้อนี้เขียนเป็นประโยคสัญลักษณ์ได้ดังข้อใดต่อไปนี้ ก. $20+\frac{4}{5}=\square$ ข. $20-\frac{4}{5}=\square$ ค. $20\times\frac{4}{5}=\square$ ง. $20\div\frac{4}{5}=\square$	“น้ำตาลหนัก 20 กิโลกรัม ตักใส่ถุงเพื่อแบ่งถุงละ $\frac{4}{5}$ กิโลกรัม จะได้กี่ถุง” เขียนเป็นประโยคสัญลักษณ์ได้ดังข้อใดต่อไปนี้ ก. $20+\frac{4}{5}=\square$ ข. $20-\frac{4}{5}=\square$ ค. $20\times\frac{4}{5}=\square$ ง. $20\div\frac{4}{5}=\square$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ข้อที่	ข้อความเดิม	ข้อความหลังปรับปรุง
40	“เชือกเส้นหนึ่งยาว 15 เมตร ตัดออกเป็น 2 ส่วน ส่วนแรกยาว $5\frac{2}{3}$ เมตร อีกส่วนยาวกี่เมตร” โจทย์ข้อนี้ เขียนเป็นประโยคสัญลักษณ์ได้ ดังข้อใดต่อไปนี้ ก. $15-5\frac{2}{3}=\square$ ข. $15\times5\frac{2}{3}=\square$ ค. $15\div5\frac{2}{3}=\square$ ง. $(15\div2)-5\frac{2}{3}=\square$	“เชือกเส้นหนึ่งยาว 15 เมตร ตัดออกเป็น 2 ส่วน ส่วนแรกยาว $5\frac{2}{3}$ เมตร อีกส่วนยาวกี่เมตร” เขียนเป็นประโยคสัญลักษณ์ได้ ดังข้อใดต่อไปนี้ ก. $15-5\frac{2}{3}=\square$ ข. $15\times5\frac{2}{3}=\square$ ค. $15\div5\frac{2}{3}=\square$ ง. $(15\div2)-5\frac{2}{3}=\square$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ค

แบบสอบวัดความสามารถทางคณิตศาสตร์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

แบบวัดผลสัมฤทธิ์ทางการเรียน กลุ่มสาระการเรียนรู้คณิตศาสตร์

<u>คำชี้แจง</u> ทำเครื่องหมายกากบาท (×) ลงบนกระดาษคำตอบที่ถูกต้องหรือเหมาะสมที่สุดเพียงคำตอบเดียวจำนวน 40 ข้อ ใช้เวลา 50 นาที

1. ค่าประมาณใกล้เคียงจำนวนเต็มร้อยของจำนวนในข้อใดต่อไปนี้แตกต่างไปจากข้ออื่น
   - ก. 654
   - ข. 732
   - ค. 645
   - ง. 723

2. 577 + 155 + 423 + 345 ควรทำตามข้อใดต่อไปนี้จึงจะหาคำตอบได้รวดเร็ว
   - ก. (577+345) + (155 + 423)
   - ข. (577 + 423) + (155 + 345)
   - ค. (577 + 345) + (155 + 423)
   - ง. 577 + (155 + 423 + 345)

3. ข้อใดต่อไปนี้เป็นสมบัติการสลับที่
   - ก. $(G \times H) + K = K + (G \times H)$
   - ข. $R \times (T + U) = (R \times U) + (R \times T)$
   - ค. $(X + Y) \times Z = (X \times Z) \times (Y \times Z)$
   - ง. $A + B + C + D = (A + B) + (C + D)$

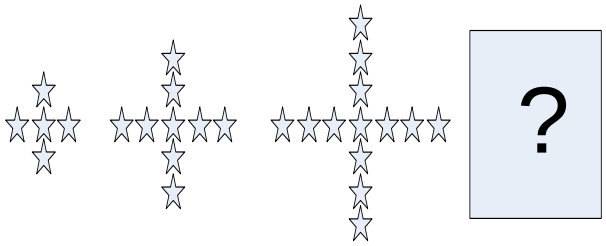
4. "เด็กชายมานะมีเงินเก็บ 287 บาท ได้รับรางวัลจากการประกวดคัดลายมือเป็นเงิน 150 บาท เด็กชายมานะต้องเก็บเงินเพิ่มอีกกี่บาท จึงจะซื้อนาฬิการาคา 579 บาทได้"
   เขียนเป็นประโยคสัญลักษณ์ได้ ดังข้อใดต่อไปนี้
   - ก. 579 - (287 + 150) = ☐
   - ข. (287 + 150) - 579 = ☐
   - ค. 579 - 287 +150 = ☐
   - ง. 579 - (287 - 150) = ☐

5. "มีส้มโอ 986 ผล ครั้งแรกขายไป 189 ผล ครั้งที่สองขายได้มากกว่าครั้งแรก 99 ผล จะเหลือส้มโอกี่ผล"
   เขียนเป็นประโยคสัญลักษณ์ได้ ดังข้อใดต่อไปนี้
   - ก. $986 - 189 - 99 = \square$
   - ข. $986 - 189 + 99 = \square$
   - ค. $986 - 189 - (189 + 99) =$
   - ง. $986 - 189 - (189 - 99) = \square$

6. "ชัยชนะซื้อข้าวสาร 9 ถุง ราคาถุงละ 115 บาท ให้เงินไป 1,500 บาท จะได้รับเงินทอนกี่บาท"
   เขียนเป็นประโยคสัญลักษณ์ได้ ดังข้อใดต่อไปนี้
   - ก. $1{,}500 - (115 \times 9) = \square$
   - ข. $1{,}500 - (115 \div 9) = \square$
   - ค. $(1{,}500 \div 9) - 115 = \square$
   - ง. $(1{,}500 \times 9) - 115 = \square$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

7. จากรูปที่กำหนดให้ จำนวน ☆ ในรูปต่อไปมีทั้งหมดกี่รูป



ก. 17 ข. 18

ค. 19 ง. 20

8. ถ้า K - 5 = 5 แล้วคำตอบของสมการคือ 5 ใช่หรือไม่ เพราะเหตุใด

ก. ไม่ใช่ เพราะ เมื่อนำ 5 ไปแทน K ในสมการ ทำให้สมการเป็นจริง

ข. ไม่ใช่ เพราะ เมื่อนำ 5 ไปแทน K ในสมการ ทำให้สมการเป็นเท็จ

ค. ใช่ เพราะ เมื่อนำ 5 ไปแทน K ในสมการ ทำให้สมการเป็นเท็จ

ง. ใช่ เพราะ เมื่อนำ 5 ไปแทน K ในสมการทำให้สมการเป็นจริง

9. สมการ $3X + 8 = 27$ ถ้าต้องการแก้สมการ จะต้องใช้สมบัติการเท่ากันในข้อใดต่อไปนี้

ก. การคูณ และ การบวก

ข. การลบ และ การคูณ

ค. การลบ และการหาร

ง. การหาร และการบวก

10. ข้อใดต่อไปนี้แสดงวิธีการแก้สมการของ $56 + X = 77$ ได้ถูกต้อง

ก. $56 + X - 56 = 77 + 56$

ข. $56 + X - 56 = 77 - 56$

ค. $56 + X + 56 = 77 - 56$

ง. $56 + X + 56 = 77 + 56$

11. จากสมการ $\frac{M}{5} = 21$ ข้อใดต่อไปนี้เป็นวิธีทำให้สมการเป็นจริง

ก. $\frac{M}{5 \times 5} = 21 \times 5$

ข. $\frac{M}{5} \times 5 = 21 \times 5$

ค. $\frac{M}{5 \times 5} = \frac{21}{5}$

ง. $\frac{M}{5} - 5 = 21 - 5$

12. “เด็กหญิงนางฟ้ามีเงิน 180 บาท ซื้อหนังสือไป A บาท พ่อให้อีก 80 บาท ปรากฏว่าเหลือเงิน 108 บาท เด็กหญิงนางฟ้าซื้อหนังสือไปกี่บาท”
เขียนเป็นสมการได้ตรงกับข้อใดต่อไปนี้

ก. 180 + A + 80 = 108

ข. 180 - A - 80 = 108

ค. 180 - A + 80 = 108

ง. 180 + A - 80 = 108

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

13. “กาละแมร์มีเงิน G บาทแบ่งให้เพื่อน 5 คน ปรากฏว่าได้คนละ 50 บาท เดิมกาละแมร์ มีเงินกี่บาท” ข้อใดต่อไปนี้เป็นวิธีการหาคำตอบของโจทย์ข้อนี้

ก. $\frac{G\times 5}{5}=\frac{50}{5}$

ข. $\frac{G\times 5}{5}=50\times 5$

ค. $(G\times 5)-50=50-5$

ง. $(G\times 5)+50=50+5$

14. “ผงซักฟอกราคาถุงละ 79 บาท ซื้อผงซักฟอก K ถุง เป็นเงินทั้งหมด 711 บาท ซื้อผงซักฟอกกี่ถุง” เขียนเป็นสมการได้ตรงกับข้อใดต่อไปนี้

ก. $\frac{79}{K}=711$

ข. $\frac{K}{79}=711$

ค. $K+79=711$

ง. $79\times K=711$

15. จำนวนในข้อใดต่อไปนี้มี 7 เป็นตัวประกอบทุกจำนวน

ก. 7 , 13 , 34

ข. 14 , 34 , 42

ค. 21 , 37 , 56

ง. 28 , 42 , 56

16. ข้อใดต่อไปนี้เป็นตัวประกอบเฉพาะทั้งหมดของ 60

ก. 1, 3, 5

ข. 2, 3, 5

ค. 2, 5, 6

ง. 2, 5, 10

17. กำหนดให้ A เป็น จำนวนนับ B และ C เป็น จำนวนเฉพาะ ข้อใดต่อไปนี้เป็นการแยกตัวประกอบ

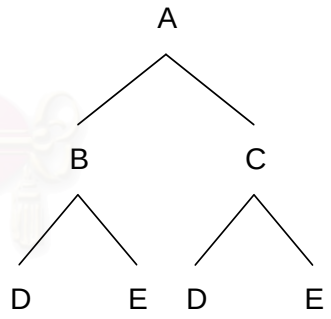
ก. A = B+ C

ข. A = B - C

ค. A = B × C

ง. A = B ÷ C

18.



จากแผนภาพการแยกตัวประกอบที่กำหนดให้ ข้อความใดต่อไปนี้**ถูกต้อง**

ก. A เป็นจำนวนเฉพาะ

ข. B เป็นตัวประกอบเฉพาะ

ค. C เป็นตัวประกอบเฉพาะ

ง. D เป็นตัวประกอบเฉพา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

19. กำหนดให้ M และ N เป็นจำนวนนับ ถ้า M เป็นตัวประกอบของ N แล้ว ข้อใดต่อไปนี้ถูกต้อง

ก. นำ N ไปหาร M ลงตัว
ข. นำ M ไปหาร N ลงตัว
ค. ทั้ง M และ N ต่างเป็นจำนวนเฉพาะ
ง. M × N ได้จำนวนเฉพาะ

20. จำนวนในข้อใดต่อไปนี้ที่แยกตัวประกอบได้เป็นจำนวนนับ 2 จำนวน

ก. 18
ข. 20
ค. 22
ง. 24

21. กำหนดให้ $X = A \times B$ และ $Y = A \times C$ ถ้า X และ Y เป็นจำนวนนับใด ๆ A, B และ C เป็นจำนวนเฉพาะ แล้ว ห.ร.ม.ของ X และ Y ได้เท่ากับข้อใดต่อไปนี้

ก. A
ข. B
ค. C
ง. BC

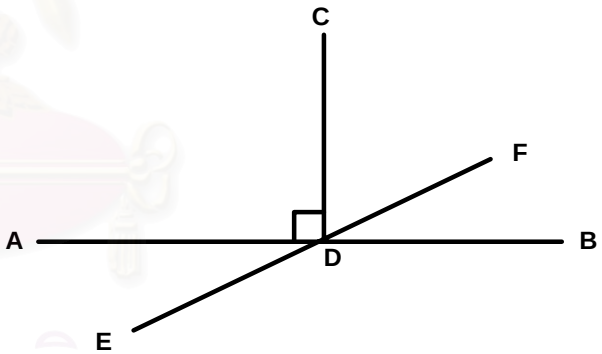
22.

2	12	18	24
3	6	9	12
2	2	3	4
2	1	3	2
3	1	3	1
	1	1	1

จากขั้นตอนการหารที่กำหนดให้ เป็นวิธีการหาสิ่งใด

ก. ห.ร.ม.
ข. ค.ร.น.
ค. แยกตัวประกอบ
ง. ห.ร.ม. และ ค.ร.น.

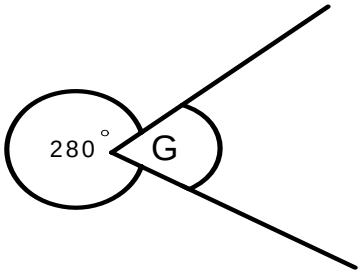
24. จากรูปที่กำหนดให้ มุมในข้อใดต่อไปนี้เป็นมุมป้าน



ก. $A\hat{D}C$
ข. $A\hat{D}E$
ค. $C\hat{D}F$
ง. $C\hat{D}E$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

25.



จากรูปที่กำหนดให้ สามารถหาค่าของมุม G ได้ตามข้อใดต่อไปนี้

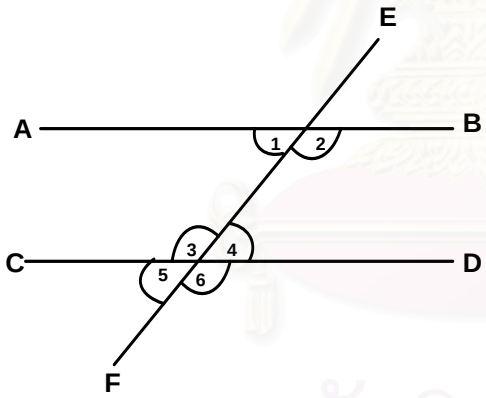
ก. $280° - 180°$

ข. $280° + 180°$

ค. $360° + 280°$

ง. $360° - 280°$

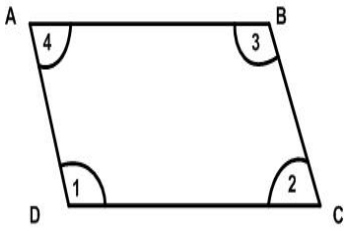
26. จากรูป ถ้า $\hat{2} = \hat{3}$ แล้วข้อใดต่อไปนี้สรุปได้ถูกต้อง



ก. $\hat{1} = \hat{6}$ ข. $\hat{4} + \hat{5} = 180°$

ค. $\overline{AB} = \overline{CD}$ ง. $\overline{AB} // \overline{CD}$

26. จากรูปที่กำหนดให้ ถ้า $\overline{AB} // \overline{DC}$ แล้วข้อใดต่อไปนี้ถูกต้อง



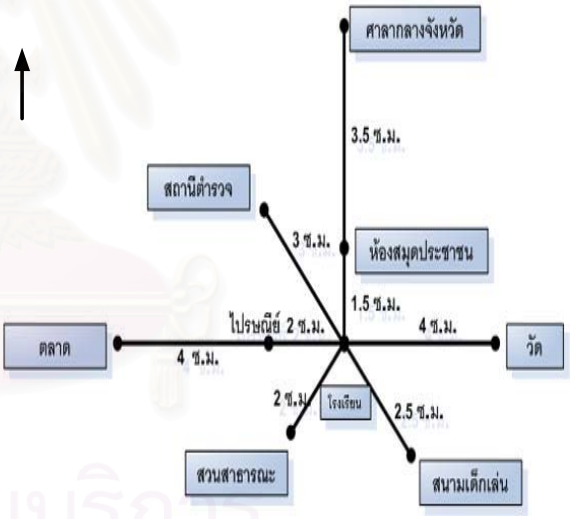
ก. $\hat{1} + \hat{4} = 180°$

ข. $\hat{1} + \hat{2} + \hat{3} = 180°$

ค. $\hat{1} + \hat{3} = 180°$

ง. $\hat{1} + \hat{2} + \hat{3} + \hat{4} = 180°$

27. จากแผนผังที่กำหนดให้ สถานที่ใดต่อไปนี้อยู่ห่างจากโรงเรียนเป็นระยะทาง 300 เมตร



มาตราส่วน 1 เซนติเมตร : 50 เมตร

ก. วัด

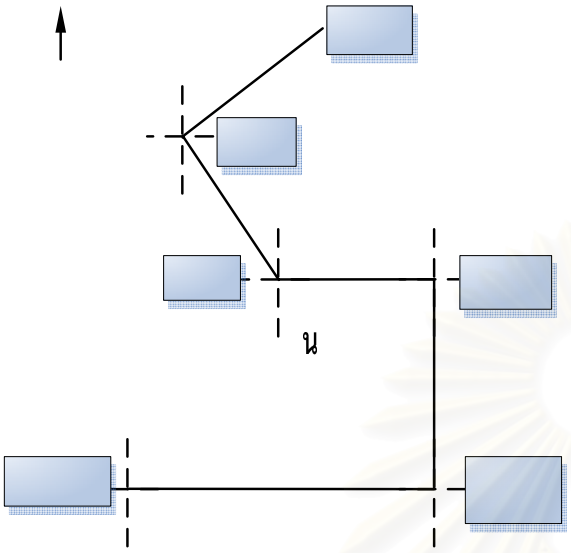
ข. ตลาด

ค. ห้องสมุดประชาชน

ง. สถานีตำรวจ

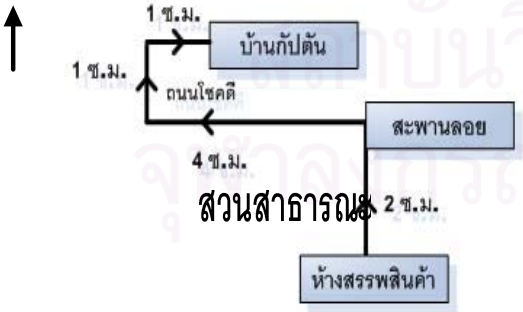
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

28. จากแผนผังที่กำหนดให้ ที่ว่าการอำเภออยู่ทางทิศใดของโรงเรียน



ก. ทิศตะวันออกเฉียงเหนือ
ข. ทิศตะวันออกเฉียงใต้
ค. ทิศตะวันตกเฉียงเหนือ
ง. ทิศตะวันตกเฉียงใต้

29. จากแผนผังที่กำหนดให้ กัปตันต้องเดินทางจากห้างสรรพสินค้ากลับบ้าน เป็นระยะทางเท่าใด เมื่อใช้มาตราส่วน 1 เซนติเมตร : 500 เมตร



ก. 4 กิโลเมตร
ข. 2 กิโลเมตร
ค. 2 กิโลเมตร 500 เมตร
ง. 1 กิโลเมตร 500 เมตร

30. เศษส่วนในข้อใดต่อไปนี้มีค่าเท่ากับ $\frac{6}{18}$

ก. $\frac{1}{3}, \frac{2}{6}, \frac{12}{36}$
ข. $\frac{1}{6}, \frac{3}{9}, \frac{12}{18}$
ค. $\frac{2}{3}, \frac{4}{6}, \frac{16}{24}$
ง. $\frac{6}{36}, \frac{1}{6}, \frac{36}{54}$

31. ข้อใดต่อไปนี้เป็นเศษส่วนอย่างต่ำ

ก. $\frac{3}{16}$
ข. $\frac{5}{10}$
ค. $\frac{8}{14}$
ง. $\frac{7}{28}$

32. ข้อใดต่อไปนี้เรียงลำดับเศษส่วนจากมากไปน้อย ได้ถูกต้อง

ก. $\frac{2}{7}, \frac{3}{7}, \frac{4}{7}$
ข. $\frac{6}{8}, \frac{5}{8}, \frac{7}{8}$
ค. $\frac{3}{7}, \frac{3}{4}, \frac{3}{3}$
ง. $\frac{5}{10}, \frac{5}{12}, \frac{5}{15}$

33. ข้อใดต่อไปนี้เป็นเศษเกิน

ก. $\frac{1}{2}$
ข. $\frac{7}{8}$
ค. $\frac{6}{5}$
ง. $\frac{4}{9}$

34. ข้อใดต่อไปนี้สามารถทำให้เป็นเป็นจำนวนคละได้

ก. $\frac{3}{9}$

ข. $\frac{4}{5}$

ค. $\frac{6}{7}$

ง. $\frac{9}{5}$

35. ข้อใดต่อไปนี้เป็นวิธีการหาคำตอบของ $\frac{A}{B}+\frac{C}{D}$

ก. $\frac{A+C}{B+D}$

ข. $\frac{(A\times D)+(C\times B)}{B\times D}$

ค. $\frac{(A+D)+(C+D)}{B+D}$

ง. $\frac{(A+D)+(C+D)}{B\times D}$

36. $\frac{A}{B}\times\frac{C}{D}\times\frac{B}{C}$ มีค่าเท่ากับข้อใดต่อไปนี้

ก. $\frac{1}{D}$

ข. 1

ค. $\frac{C}{B}$

ง. $\frac{A}{D}$

37. $\frac{\frac{A}{B}\times\frac{B}{C}}{\frac{A}{B}\div\frac{C}{B}}$ มีค่าเท่ากับข้อใดต่อไปนี้

ก. 0 ข. 1

ค. $\frac{1}{BC}$ ง. $\frac{A}{BC}$

38. ข้อใดต่อไปนี้ควรเป็นขั้นตอนแรกในการหาคำตอบของ $5\frac{8}{9}+\frac{5}{7}\div1\frac{5}{8}=[\ ]$

ก. หา ค.ร.น.

ข. กลับตัวเศษเป็นตัวส่วน

ค. เปลี่ยนเครื่องหมายหารเป็นเครื่องหมายคูณ

ง. แปลงจำนวนคละให้เป็นเศษเกิน

39. “น้ำตาลหนัก 20 กิโลกรัม ตักใส่ถุงเพื่อแบ่งขาย ถุงละ $\frac{4}{5}$ กิโลกรัม จะได้กี่ถุง” เขียนเป็นประโยคสัญลักษณ์ได้ดังข้อใดต่อไปนี้

ก. $20+\frac{4}{5}=\square$

ข. $20-\frac{4}{5}=\square$

ค. $20\times\frac{4}{5}=\square$

ง. $20\div\frac{4}{5}=\square$

40. “เชือกเส้นหนึ่งยาว 15 เมตร ตัดออกเป็น 2 ส่วน ส่วนแรกยาว $5\frac{2}{3}$ เมตร อีกส่วนยาวกี่เมตร” เขียนเป็นประโยคสัญลักษณ์ได้ ดังข้อใดต่อไปนี้

ก. $15-5\frac{2}{3}=\square$

ข. $15\times5\frac{2}{3}=\square$

ค. $15\div5\frac{2}{3}=\square$

ง. $(15\div2)-5\frac{2}{3}=\square$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ง

ค่าพารามิเตอร์ของข้อสอบรายข้อใน
แบบสอบวัดความสามารถทางคณิตศาสตร์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ง

ค่าพารามิเตอร์ของข้อสอบในแบบสอบวัดความสามารถทางคณิตศาสตร์ โดยใช้การวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT)

Item	a	b	c	หมายเหตุ
1	0.559	-3.003	0.296	ตัดทิ้ง
2	0.592	3.610	0.200	ตัดทิ้ง
3	0.539	-2.897	0.286	คัดเลือก
4	0.761	1.384	0.200	คัดเลือก
5	0.761	1.947	0.200	คัดเลือก
6	0.778	1.123	0.200	คัดเลือก
7	0.752	2.037	0.200	คัดเลือก
8	0.601	2.829	0.200	คัดเลือก
9	0.509	1.805	0.228	คัดเลือก
10	0.386	3.601	0.204	ตัดทิ้ง
11	0.637	2.967	0.272	คัดเลือก
12	0.624	-2.460	0.147	คัดเลือก
13	0.767	3.875	0.200	ตัดทิ้ง
14	0.319	3.961	0.173	ตัดทิ้ง
15	0.685	3.037	0.303	ตัดทิ้ง
16	0.969	-0.827	0.200	คัดเลือก
17	0.565	3.022	0.200	ตัดทิ้ง
18	0.776	1.660	0.200	คัดเลือก
19	0.874	1.301	0.308	ตัดทิ้ง
20	0.570	3.373	0.301	ตัดทิ้ง
21	0.355	3.300	0.200	ตัดทิ้ง
22	0.623	-3.742	0.200	ตัดทิ้ง
23	0.704	2.359	0.200	คัดเลือก
24	0.418	2.092	-0.005	ตัดทิ้ง
25	0.830	0.725	0.200	คัดเลือก
26	0.625	1.749	0.200	คัดเลือก
27	0.551	3.279	0.200	ตัดทิ้ง

Item	a	b	c	หมายเหตุ
28	0.683	-1.742	0.182	คัดเลือก
29	0.604	2.541	0.200	คัดเลือก
30	0.682	-3.931	0.200	ตัดทิ้ง
31	0.367	3.636	0.129	ตัดทิ้ง
32	0.629	3.083	0.200	ตัดทิ้ง
33	0.592	-2.751	0.268	คัดเลือก
34	0.973	0.559	0.200	คัดเลือก
35	0.757	1.455	0.200	คัดเลือก
36	0.568	2.448	0.200	คัดเลือก
37	0.707	-0.006	0.164	คัดเลือก
38	0.830	1.324	0.200	คัดเลือก
39	0.369	3.572	0.121	ตัดทิ้ง
40	0.332	3.253	0.189	ตัดทิ้ง
41	0.654	-2.334	0.133	คัดเลือก
42	0.629	2.512	0.200	คัดเลือก
43	0.545	-3.300	0.200	ตัดทิ้ง
44	0.731	2.181	0.200	คัดเลือก
45	0.706	0.985	0.192	คัดเลือก
46	0.700	3.312	0.306	ตัดทิ้ง
47	0.506	-3.938	0.200	ตัดทิ้ง
48	0.673	-3.208	0.200	ตัดทิ้ง
49	0.714	-3.487	0.200	ตัดทิ้ง
50	0.813	1.485	0.200	คัดเลือก
51	0.424	2.704	0.195	ตัดทิ้ง
52	0.396	2.983	0.184	ตัดทิ้ง
53	0.578	3.610	0.200	ตัดทิ้ง
54	0.554	2.276	0.200	คัดเลือก
55	0.346	3.723	0.218	ตัดทิ้ง
56	0.402	2.847	0.126	ตัดทิ้ง

Item	a	b	c	หมายเหตุ
57	0.649	2.791	0.297	คัดเลือก
58	0.469	-3.968	0.200	ตัดทิ้ง
59	0.815	2.549	0.243	คัดเลือก
60	0.508	3.234	0.200	ตัดทิ้ง
61	0.481	2.971	0.297	ตัดทิ้ง
62	0.728	-3.879	0.200	ตัดทิ้ง
63	0.378	3.731	0.245	ตัดทิ้ง
64	0.636	2.483	0.200	คัดเลือก
65	0.623	3.523	0.200	ตัดทิ้ง
66	0.843	1.818	0.200	คัดเลือก
67	0.662	1.306	0.236	คัดเลือก
68	0.882	1.346	0.200	คัดเลือก
69	0.596	0.748	0.200	คัดเลือก
70	0.675	-2.393	0.200	คัดเลือก
71	0.542	2.656	0.200	คัดเลือก
72	0.826	1.383	0.200	คัดเลือก
73	0.734	-1.332	0.200	คัดเลือก
74	0.569	3.142	0.200	ตัดทิ้ง
75	0.322	3.300	-0.003	ตัดทิ้ง
76	0.495	2.190	0.148	ตัดทิ้ง
77	0.636	2.432	0.200	คัดเลือก
78	0.397	2.925	0.164	ตัดทิ้ง
79	0.391	3.164	0.175	ตัดทิ้ง
80	0.381	3.645	0.196	ตัดทิ้ง

ภาคผนวก จ

โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก จ

โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

```
unit TestUtil;

interface

uses Classes, Graphics, ItemTime;

const
  ProgramStr = 'The Test data file XP v';
  VersionStr = '2.3';
  UserDataStr = 'The Test - user data file v';
  VersionUserDataStr = '1.0';
  DataSignature: String[Length(ProgramStr + VersionStr) + 1] =
    ProgramStr + VersionStr + #26;
  UserSignature: String[Length(UserDataStr + VersionUserDataStr) + 1]
=
    UserDataStr + VersionUserDataStr + #26;

  UserFileName = 'User.dat';

  Choices: array[0..2] of String[5] = (
    'กขคงจ', 'ABCDE', '12345');

type
  TTestPassword = String[8];

  TTestObj = class(TObject)
    PassWord: TTestPassword;
    TestName: String[40];
    ItemTimeRec: TSetItemTimeRec;
    ChoiceType: Byte;
    NumChoice: Byte;
    MaxItem: Integer;
    DoTime: Integer;
    CurItem: Integer;
  end;

  TItemTextRec = record
    Lines: TStringList;
    Alignment: TAlignment;
    WordWrap: Boolean;
    Font: TFont;
    BkColor: TColor;
  end;

  TItemImageRec = record
    Position: Byte;
    AutoSize: Boolean;
    Center: Boolean;
    Stretch: Boolean;
    Size: Integer;
    BkColor: TColor;
    FileName: String[20];
  end;

  TItemObj = class(TObject)
```

```
    Text: TItemTextRec;
    Image: TItemImageRec;
    ValueA: Single;
    ValueB: Single;
    ValueC: Single;
    CorrAns: Byte;
    DoTime: Integer;
    destructor Destroy; override;
  end;

  TItemList = class(TList)
    destructor Destroy; override;
    procedure InsertItem(Index: Integer; Item: Pointer);
  end;

  TLastItemObj = class(TObject)
    Item: Integer;
  end;

  TLastItemList = class(TList)
    destructor Destroy; override;
  end;

  TUserObj = class(TObject)
    UserName: String[40];
    ThrmID: Integer;
    RateLast: Integer;
    EndSEE: Integer;
    TestCount: Integer;
    LastList: TLastItemList;
    destructor Destroy; override;
  end;

  TUserList = class(TList)
    destructor Destroy; override;
  end;

{ Stream Routines }
function GetTestFromStream(S: TStream): Boolean;
procedure PutTestToStream(S: TStream);
function GetUserFromStream(S: TStream; UserList: TUserList): Boolean;
procedure PutUserToStream(S: TStream; UserList: TUserList);

{ General Routines }
function StrToRealDef(S: String; Def: Real): Real;
function GetDoTestType: String;

{ Statistic Routines }
function Get_D(A, B: Single; Zeta, RowSqr: Double): Double;
function Get_Tp(D: Double): Double;
function Get_Tm(D: Double): Double;
function Get_O_D(D: Double): Double;
function Get_A_D(O_D, T: Double): Double;
function Get_NextZeta(CorrAns: Boolean;
  A, C: Single; Zeta, RowSqr, O_D, A_Dp, A_Dm: Double): Double;
function Get_NextRowSqr(CorrAns: Boolean;
  A, C: Single; D, RowSqr, O_D, A_Dp, A_Dm: Double): Double;
function GetIZetaMax(A, B, C: Single; Zeta: Double): Double;
function GetSEE(RowSqr: Double): Double;
function GetH(N: Integer): Double;
function GetZetaI(I: Single; N: Integer; ZetaH: Double): Double;
```

```
function GetKL(A, B, C: Single; Zeta, ZetaH: Double): Double;
function GetKLZetaH(N: Integer; A, B, C: Single; ZetaH: Double):
Double;

const
  StartRun: Boolean = True;
  DesignMode: Boolean = False;
  CanTestMode: Boolean = False;
  TestingMode: Boolean = False;
  TestFile: String = '';
  TestPassWord: TTestPassword = '';
  TestChanged: Boolean = False;
  ItemChanged: Boolean = False;
  ItemList: TItemList = nil;
  CurrentItem: Integer = 0;
  CurrentTestItem: Integer = 0;
  CurrentScore: Integer = 0;
  CurrentZeta: Double = 0.0;
  LastZeta: Double = 0.0;
  CurrentRowSqr: Double = 1.0;
  SumIZetaMax: Double = 0.0;
  SumKLZetaH: Double = 0.0;
  ZetaAvg: Double = 0.0;
  ZetaSum: Double = 0.0;
  StartTime: TDateTime = 0;

  { TestingCount = 1   --> In first test.
                > 1   --> In next test and doing items in
LastItemList.
                < -1  --> In next test but continue doing items in
ItemList.
                =-1,0 --> No meaning.
  }
  TestingCount: Integer = 1;

  { LastItemList hold items when TestingCount > 1 }
  LastItemList: TLastItemList = nil;
  LastItemG1Count: Integer = 0;  { Number of items of 1st group in
LastItemList }
  LastItemCount: Integer = 0;   { Count items in LastItemList that
doing }
  SaveItemList: TLastItemList = nil;  { Save current items that doing
}

  TheoremID: Integer = 0;       { 1, 2, 3 }
  RateLastItem: Integer = 0;    { 10, 20, 30 % }
  EndingSEE: Double = 0.0;      { 0.3, 0.45 }
  First045: Boolean = True;     { Check if the first time SEE < 0.45
but not < 0.3 }
  RateInG1: Boolean = True;     { Rate in 1st group in LastItemList }

implementation

uses SysUtils, TestProp, Main, DoTest;

{ TItemObj }
destructor TItemObj.Destroy;
begin
  Text.Lines.Free;
  Text.Font.Free;
  inherited Destroy;
```

```
end;

{ TItemList }
destructor TItemList.Destroy;
var
  I: Integer;
begin
  for I := 0 to Count - 1 do
    TItemObj(Items[I]).Free;
  inherited Destroy;
end;

procedure TItemList.InsertItem(Index: Integer; Item: Pointer);
begin
  if (Index >= 0) and (Index < Count) then
  begin
    TItemObj(Items[Index]).Free;
    Remove(Items[Index]);
  end;
  Insert(Index, Item);
end;

{ TLastItemList }
destructor TLastItemList.Destroy;
var
  I: Integer;
begin
  for I := 0 to Count - 1 do
    TLastItemObj(Items[I]).Free;
  inherited Destroy;
end;

{ TUserObj }
destructor TUserObj.Destroy;
begin
  LastList.Free;
  inherited Destroy;
end;

{ TUserList }
destructor TUserList.Destroy;
var
  I: Integer;
begin
  for I := 0 to Count - 1 do
    TUserObj(Items[I]).Free;
  inherited Destroy;
end;

{ Stream Routines }
function GetTestFromStream(S: TStream): Boolean;
var
  DataSig: String;
  TestObj: TTestObj;
  ItemObj: TItemObj;
  I, L, Count: Integer;
  AColor: TColor;
  { have problem about TFontName type when write to stream }
  { because SizeOf(TFontName) = 4, not actual size of data }
  { AName: TFontName; }
  AName: String;
```

จุฬาลงกรณ์มหาวิทยาลัย

```
  ASize: Integer;
  APitch: TFontPitch;
  AStyle: TFontStyles;
  AHeight: Integer;
  APixelsPerInch: Integer;
  SavePos: Longint;
begin
  SetLength(DataSig, Length(DataSignature));
  S.Read(DataSig[1], Length(DataSignature));
  if DataSig <> DataSignature then
  begin
    GetTestFromStream := False;
    Exit;
  end;

  GetTestFromStream := True;
  TestObj := TTestObj.Create;
  with TestObj do
  begin
    S.Read(Password, SizeOf(Password));
    S.Read(TestName, SizeOf(TestName));
    S.Read(ItemTimeRec, SizeOf(ItemTimeRec));
    S.Read(ChoiceType, SizeOf(ChoiceType) + SizeOf(NumChoice)
      + SizeOf(MaxItem) + SizeOf(DoTime) + SizeOf(CurItem));
    TestPassword := Password;
    TestPropRec.TestName := TestName;
    SetItemTimeRec := ItemTimeRec;
    TestPropRec.ChoiceType := ChoiceType;
    TestPropRec.NumChoice := NumChoice;
    TestPropRec.MaxItem := MaxItem;
    TestPropRec.DoTime := DoTime;
    CurrentItem := CurItem;
  end;
  TestObj.Free;

  if (ItemList <> nil) and (ItemList.Count = 0) then
  begin
    S.Read(Count, SizeOf(Integer));
    if Count > 0 then
      for I := 0 to Count - 1 do
      begin
        ItemObj := TItemObj.Create;
        with ItemObj do
        begin
          with Text do
          begin
            S.Read(SavePos, SizeOf(SavePos));
            Lines := TStringList.Create;
            Lines.LoadFromStream(S);
            S.Seek(SavePos, soFromBeginning);
            S.Read(Alignment, SizeOf(Alignment));
            S.Read(WordWrap, SizeOf(WordWrap));
            Font := TFont.Create;
            with Font do
            begin
              S.Read(AColor, SizeOf(AColor));
              Color := AColor;
              S.Read(L, SizeOf(L));
              SetLength(AName, L);
              S.Read(AName[1], L);
              Name := AName;
```

```
              S.Read(ASize, SizeOf(ASize));
              Size := ASize;
              S.Read(APitch, SizeOf(APitch));
              Pitch := APitch;
              S.Read(AStyle, SizeOf(AStyle));
              Style := AStyle;
              S.Read(AHeight, SizeOf(AHeight));
              Height := AHeight;
              S.Read(APixelsPerInch, SizeOf(APixelsPerInch));
              PixelsPerInch := APixelsPerInch;
            end;
            S.Read(BkColor, SizeOf(BkColor));
          end;
          with Image do
          begin
            S.Read(Position, SizeOf(Position));
            S.Read(AutoSize, SizeOf(AutoSize) * 3);
            S.Read(Size, SizeOf(Size));
            S.Read(BkColor, SizeOf(BkColor));
            S.Read(FileName, SizeOf(FileName));
          end;
          S.Read(ValueA, SizeOf(ValueA) * 3);
          S.Read(CorrAns, SizeOf(CorrAns));
          S.Read(DoTime, SizeOf(DoTime));
        end;
        ItemList.Add(ItemObj);
      end;
  end;
end;

procedure PutTestToStream(S: TStream);
var
  TestObj: TTestObj;
  ItemObj: TItemObj;
  I, L: Integer;
  AColor: TColor;
  { have problem about TFontName type when write to stream }
  { because SizeOf(TFontName) = 4, not actual size of data }
  { AName: TFontName; }
  AName: String;
  ASize: Integer;
  APitch: TFontPitch;
  AStyle: TFontStyles;
  AHeight: Integer;
  APixelsPerInch: Integer;
  SavePos1, SavePos2: Longint;
begin
  S.Write(DataSignature[1], Length(DataSignature));

  TestObj := TTestObj.Create;
  with TestObj do
  begin
    Password := TestPassword;
    TestName := TestPropRec.TestName;
    ItemTimeRec := SetItemTimeRec;
    ChoiceType := TestPropRec.ChoiceType;
    NumChoice := TestPropRec.NumChoice;
    MaxItem := TestPropRec.MaxItem;
    DoTime := TestPropRec.DoTime;
    CurItem := CurrentItem;
    S.Write(Password, SizeOf(Password));
```

```
  S.Write(TestName, SizeOf(TestName));
  S.Write(ItemTimeRec, SizeOf(ItemTimeRec));
  S.Write(ChoiceType, SizeOf(ChoiceType) + SizeOf(NumChoice)
    + SizeOf(MaxItem) + SizeOf(DoTime) + SizeOf(CurItem));
end;
TestObj.Free;

if ItemList <> nil then
begin
  S.Write(ItemList.Count, SizeOf(Integer));
  for I := 0 to ItemList.Count - 1 do
    if ItemList.Items[I] <> nil then
    begin
      ItemObj := TItemObj(ItemList.Items[I]);
      with ItemObj do
      begin
        with Text do
        begin
          SavePos1 := 0;  { Make temp value }
          SavePos2 := S.Position;  { Save position of SavePos1 }
          S.Write(SavePos1, SizeOf(SavePos1));
          Lines.SaveToStream(S);
          SavePos1 := S.Position;  { Set to valid value }
          S.Seek(SavePos2, soFromBeginning);  { seek to SavePos }
          S.Write(SavePos1, SizeOf(SavePos1));
          S.Seek(SavePos1, soFromBeginning);
          S.Write(Alignment, SizeOf(Alignment));
          S.Write(WordWrap, SizeOf(WordWrap));
          with Font do
          begin
            AColor := Color;
            S.Write(AColor, SizeOf(AColor));
            AName := Name;
            L := Length(AName);
            S.Write(L, SizeOf(L));
            S.Write(AName[1], L);
            ASize := Size;
            S.Write(ASize, SizeOf(ASize));
            APitch := Pitch;
            S.Write(APitch, SizeOf(APitch));
            AStyle := Style;
            S.Write(AStyle, SizeOf(AStyle));
            AHeight := Height;
            S.Write(AHeight, SizeOf(AHeight));
            APixelsPerInch := PixelsPerInch;
            S.Write(APixelsPerInch, SizeOf(APixelsPerInch));
          end;
          S.Write(BkColor, SizeOf(BkColor));
        end;
        with Image do
        begin
          S.Write(Position, SizeOf(Position));
          S.Write(AutoSize, SizeOf(AutoSize) * 3);
          S.Write(Size, SizeOf(Size));
          S.Write(BkColor, SizeOf(BkColor));
          S.Write(FileName, SizeOf(FileName));
        end;
        S.Write(ValueA, SizeOf(ValueA) * 3);
        S.Write(CorrAns, SizeOf(CorrAns));
        S.Write(DoTime, SizeOf(DoTime));
      end;
```

```
      end;
  end;
end;

function GetUserFromStream(S: TStream; UserList: TUserList): Boolean;
var
  I, J, NU, NL: Integer;
  DataSig: String;
  UserObj: TUserObj;
  LastItemObj: TLastItemObj;
begin
  SetLength(DataSig, Length(UserSignature));
  S.Read(DataSig[1], Length(UserSignature));
  if DataSig <> UserSignature then
  begin
    GetUserFromStream := False;
    Exit;
  end;

  GetUserFromStream := True;
  S.Read(NU, SizeOf(NU));
  if NU <= 0 then Exit;

  for I := 1 to NU do
  begin
    UserObj := TUserObj.Create;
    with UserObj do
    begin
      S.Read(UserName, SizeOf(UserName));
      S.Read(ThrmID, SizeOf(ThrmID));
      S.Read(RateLast, SizeOf(RateLast));
      S.Read(EndSEE, SizeOf(EndSEE));
      S.Read(TestCount, SizeOf(TestCount));
      S.Read(NL, SizeOf(NL));
      if NL <= 0 then
        LastList := nil
      else
      begin
        LastList := TLastItemList.Create;
        for J := 1 to NL do
        begin
          LastItemObj := TLastItemObj.Create;
          with LastItemObj do
            S.Read(Item, SizeOf(Item));
          LastList.Add(LastItemObj);
        end; { for J }
      end; { if NL }
    end; { with UserObj }
    UserList.Add(UserObj);
  end; { for I }
end;

procedure PutUserToStream(S: TStream; UserList: TUserList);
var
  I, J, NL: Integer;
begin
  S.Write(UserSignature[1], Length(UserSignature));

  S.Write(UserList.Count, SizeOf(UserList.Count));
  for I := 0 to UserList.Count - 1 do
  begin
```

```
    with TUserObj(UserList.Items[I]) do
    begin
      S.Write(UserName, SizeOf(UserName));
      S.Write(ThrmID, SizeOf(ThrmID));
      S.Write(RateLast, SizeOf(RateLast));
      S.Write(EndSEE, SizeOf(EndSEE));
      S.Write(TestCount, SizeOf(TestCount));
      if LastList = nil then
        NL := 0
      else
        NL := LastList.Count;
      S.Write(NL, SizeOf(NL));
      if NL > 0 then
        for J := 0 to LastList.Count - 1 do
          with TLastItemObj(LastList.Items[J]) do
            S.Write(Item, SizeOf(Item));
    end; { with TUserObj }
  end; { for I }
end;

{ General Routines }
function StrToRealDef(S: String; Def: Real): Real;
var
  R: Real;
  E: Integer;
begin
  Val(S, R, E);
  if E = 0 then
    Result := R
  else
    Result := Def;
end;

function Power(Base: Double; Exponent: Integer): Double;
var
  I: Integer;
begin
  Result := 1;
  for I := 1 to Exponent do
    Result := Result * Base;
end;

function GetDoTestType: String;
begin
  Result := IntToStr(DoTestRec.TheoremID + 1)
    + IntToStr(DoTestRec.RateLastItem + 1)
    + IntToStr(DoTestRec.EndingSEE + 1)
    + IntToStr(DoTestRec.NumItems + 1)
    + IntToStr(DoTestRec.TimeUsed + 1)
    + '-'
    + IntToStr(Abs(TestingCount));
end;

{ Statistic Routines }
function Get_D(A, B: Single; Zeta, RowSqr: Double): Double;
begin
  Result := (B - Zeta) / Sqrt( (1/Power(A,2)) + RowSqr);
end;

function Get_Tp(D: Double): Double;
begin
```

```
  Result := 1 / (1 + 0.2316419 * D);
end;

function Get_Tm(D: Double): Double;
begin
  Result := 1 / (1 - 0.2316419 * D);
end;

function Get_O_D(D: Double): Double;
begin
  Result := 1 / ( Sqrt(2 * Pi) * Exp(D * D / 2) );
end;

function Get_A_D(O_D, T: Double): Double;
const
  C1 =  0.31938153;
  C2 = -0.356563782;
  C3 =  1.78147937;
  C4 = -1.821255978;
  C5 =  1.330274429;
begin
  Result := 1 - O_D * (C1 * T + C2 * Power(T, 2)
    + C3 * Power(T, 3) + C4 * Power(T, 4) + C5 * Power(T, 5));
end;

function Get_NextZeta(CorrAns: Boolean;
  A, C: Single; Zeta, RowSqr, O_D, A_Dp, A_Dm: Double): Double;
var
  Tmp: Double;
begin
  Tmp := RowSqr / Sqrt(1/(A * A) + RowSqr);
  if CorrAns then
    Result := Zeta + (1 - C) * Tmp * (O_D / (C + (1 - C) * A_Dm))
  else
    Result := Zeta - Tmp * (O_D / A_Dp);
end;

function Get_NextRowSqr(CorrAns: Boolean;
  A, C: Single; D, RowSqr, O_D, A_Dp, A_Dm: Double): Double;
var
  Tmp1, Tmp2: Double;
begin
  Tmp1 := 1 + 1 / (A * A * RowSqr);
  Tmp2 := C + (1 - C) * A_Dm;
  if CorrAns then
  begin
    Result := ((1 - C) / Tmp1) * (O_D / Tmp2) * ( (1 - C) * O_D /
Tmp2 - D);
    Result := RowSqr * (1 - Result);
  end else
  begin
    Result := (O_D / Tmp1) * (O_D / A_Dp + D);
    Result := RowSqr * (1 - Result / A_Dp);
  end;
end;

function GetIZetaMax(A, B, C: Single; Zeta: Double): Double;
var
  Tmp: Double;
begin
  Tmp := 1.7 * A * (Zeta - B);
```

```
  Result := (C + Exp(Tmp)) * Power((1 + Exp(-Tmp)), 2);
  Result := Power(1.7 * A, 2) * (1 - C) / Result;
end;

function GetSEE(RowSqr: Double): Double;
begin
  Result := Sqrt(RowSqr);
end;

(*
function GetSEE(Zeta: Double): Double;
begin
  Result := 1 / Sqrt(Zeta);
end;
*)

{ Kullback - Leibler Formulas }

function GetLower(N: Integer; ZetaH: Double): Double;
begin
  Result := ZetaH - 3 / (Sqrt(N) + 1);
end;

function GetUpper(N: Integer; ZetaH: Double): Double;
begin
  Result := ZetaH + 3 / (Sqrt(N) + 1);
end;

function GetPZetaH(A, B, C: Single; ZetaH: Double): Double;
var
  Tmp: Double;
begin
  Tmp := Exp(1.7 * A * (ZetaH - B));
  Result := C + (1 - C) * Tmp / (1 + Tmp);
end;

function GetH(N: Integer): Double;
begin
  Result := 0.06 / (Sqrt(N) + 1);
end;

function GetZetaI(I: Single; N: Integer; ZetaH: Double): Double;
begin
  Result := (ZetaH - 3 / (Sqrt(N) + 1)) + I * GetH(N);
end;

function GetKL(A, B, C: Single; Zeta, ZetaH: Double): Double;
var
  PZ, PZH: Double;
begin
  PZ := GetPZetaH(A, B, C, Zeta);
  PZH := GetPZetaH(A, B, C, ZetaH);
  Result := (1 - PZH) / (1 - PZ);
  Result := PZH * Ln(PZH / PZ) + (1 - PZH) * Ln(Result);
end;

function GetKLZetaH(N: Integer; A, B, C: Single; ZetaH: Double):
Double;
var
  I: Integer;
  Sum1, Sum2: Double;
```

```
begin
  Sum1 := 0.0;
  I := 1;
  repeat
    Sum1 := Sum1 + GetKL(A, B, C, GetZetaI(I, N, ZetaH), ZetaH);
    I := I + 2;
  until I > N - 1;
  Sum2 := 0.0;
  I := 2;
  repeat
    Sum2 := Sum2 + GetKL(A, B, C, GetZetaI(I, N, ZetaH), ZetaH);
    I := I + 2;
  until I > N - 2;
  Result := (GetKL(A, B, C, GetLower(N, ZetaH), ZetaH)
    + GetKL(A, B, C, GetUpper(N, ZetaH), ZetaH)
    + 4 * Sum1 + 2 * Sum2) * GetH(N) / 3;
end;

end.
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ฉ

ค่าพารามิเตอร์คลังข้อสอบ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ฉ

ค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบ

Item	a	b	c	Item	a	b	c
1	1.0890	0.5350	0.1550	28	0.4060	1.6990	0.1930
2	1.0640	0.6670	0.1520	29	1.3130	0.9870	0.1530
3	0.8560	2.2140	0.2680	30	1.1970	1.3550	0.1870
4	0.9790	2.2110	0.1320	31	1.1730	2.6330	0.2150
5	1.1980	1.3500	0.0920	32	0.6080	-1.5700	0.2130
6	0.7070	1.1370	0.2400	33	0.9180	2.1180	0.2300
7	0.9610	0.2410	0.2490	34	0.9380	1.2480	0.2020
8	0.8100	0.1910	0.2390	35	0.5110	0.0130	0.1880
9	1.2980	1.4610	0.2240	36	1.2030	1.2320	0.2050
10	1.5130	2.0550	0.1920	37	1.7100	1.7060	0.1760
11	0.7530	-0.0750	0.2900	38	0.5980	-1.0800	0.1660
12	0.9440	2.0020	0.1870	39	0.4550	-1.2500	0.2120
13	0.9320	0.6420	0.2180	40	1.2970	0.8020	0.1360
14	1.0820	-0.0300	0.1970	41	1.7440	1.5720	0.1530
15	1.1850	-0.3000	0.2100	42	0.6220	-0.7360	0.2610
16	1.6130	2.7200	0.1800	43	0.9040	0.4550	0.1100
17	0.9460	1.2230	0.1030	44	1.5000	1.6860	0.1540
18	0.9660	0.7270	0.2950	45	1.0420	1.6260	0.1640
19	0.7070	-0.5290	0.1800	46	0.8630	-0.0310	0.1920
20	1.0250	1.6670	0.2260	47	1.1700	1.2810	0.2580
21	1.4390	1.5620	0.2070	48	1.2990	1.8390	0.2780
22	1.1100	1.0440	0.2570	49	0.6980	-0.5210	0.1410
23	1.4730	0.8130	0.2840	50	0.5820	-0.0490	0.1780
24	0.9380	0.9570	0.2010	51	1.2070	0.2540	0.2480
25	1.3640	0.8910	0.1490	52	1.6310	1.2400	0.1980
26	0.4860	-0.9310	0.2440	53	0.7060	-0.0100	0.1140
27	0.9290	0.2540	0.2190	54	1.1210	-0.3990	0.1450

Item	a	b	c	Item	a	b	c
55	1.0620	-0.7470	0.1130	85	1.0460	0.7460	0.2790
56	0.6350	-0.3670	0.1460	86	1.5690	0.8150	0.1610
57	1.3840	0.4120	0.1290	87	0.8250	-0.4740	0.1740
58	1.6840	0.5900	0.1570	88	1.0400	0.3150	0.2240
59	1.1800	1.2530	0.1330	89	1.3750	1.1790	0.2960
60	1.0880	0.2950	0.2040	90	0.8320	2.2920	0.2820
61	0.8040	0.8990	0.2310	91	0.8710	0.6420	0.1870
62	0.8610	-0.0640	0.0880	92	1.1060	1.0990	0.2280
63	1.0050	0.4050	0.2220	93	0.8730	1.1230	0.1290
64	1.1330	2.2340	0.2140	94	1.6750	1.2930	0.1250
65	0.6150	-1.5200	0.2470	95	1.5550	0.7410	0.2800
66	1.1690	1.6790	0.2130	96	0.6110	-1.0600	0.2330
67	1.6510	1.4810	0.2670	97	1.3640	2.5410	0.1480
68	0.7660	-0.3270	0.1600	98	0.8130	0.6660	0.2050
69	1.2910	2.2950	0.2350	99	1.0720	0.1540	0.2450
70	1.4990	0.8750	0.1510	100	0.4620	-0.5140	0.1560
71	0.4950	-2.2300	0.2150	101	0.8230	-0.5070	0.1800
72	0.5870	-0.1960	0.2800	102	0.8670	2.3940	0.2170
73	1.1080	0.4010	0.1470	103	0.6300	-0.5710	0.2330
74	0.7980	0.3520	0.1890	104	0.6950	-0.7570	0.1770
75	1.0900	-0.2050	0.2040	105	1.1340	0.5270	0.2400
76	1.6100	-0.3020	0.2210	106	0.6990	-0.1330	0.1670
77	1.3010	-0.3540	0.2340	107	1.5860	2.1150	0.2010
78	1.1970	-0.1940	0.2480	108	0.7230	-0.8730	0.1350
79	1.5580	0.0400	0.2470	109	1.2060	-0.3060	0.1050
80	0.9840	1.7800	0.2870	110	1.4730	0.9650	0.1220
81	0.8750	-0.9180	0.1390	111	0.9550	1.2360	0.1760
82	0.8820	-0.2660	0.1720	112	1.2660	0.6460	0.2700
83	1.6230	0.5560	0.2350	113	0.8380	0.5780	0.1310
84	0.4390	-0.5950	0.2340	114	0.5110	-1.3100	0.1960

Item	a	b	c	Item	a	b	c
115	1.0380	1.3310	0.1460	145	1.6370	1.9200	0.0610
116	0.5910	-0.9860	0.1850	146	0.7880	-1.4200	0.1880
117	1.5760	2.0880	0.2700	147	0.7230	-1.6700	0.1860
118	1.1960	1.1030	0.2880	148	0.8800	0.5520	0.1760
119	1.5430	2.2530	0.0820	149	1.4990	1.4290	0.1410
120	0.9740	-1.0100	0.1740	150	1.2550	1.0550	0.1380
121	1.0560	1.3650	0.1270	151	0.9380	0.0810	0.2320
122	1.0190	0.0210	0.2280	152	1.5010	1.3190	0.1570
123	0.4700	-0.3100	0.1920	153	1.4900	1.0140	0.2110
124	1.3320	0.2890	0.2150	154	0.9980	-0.5510	0.2180
125	1.2830	0.1190	0.1520	155	1.0240	-0.1150	0.2230
126	1.3500	-0.3830	0.1740	156	1.3650	0.1240	0.2560
127	1.0010	1.0010	0.2220	157	2.3380	1.4060	0.1180
128	1.7010	2.4170	0.1560	158	1.0760	2.6450	0.1920
129	1.0000	1.2220	0.2220	159	0.6620	-0.0600	0.2560
130	0.8340	0.6480	0.2240	160	0.8990	0.4630	0.1080
131	1.8350	1.0400	0.2020	161	1.2480	2.0150	0.1980
132	1.6000	1.1560	0.1420	162	1.2180	1.0520	0.2110
133	1.4490	0.6210	0.1710	163	0.7020	-0.5680	0.1860
134	0.6170	-0.9510	0.1540	164	1.5140	0.0400	0.1850
135	1.2850	0.4710	0.1920	165	1.9290	0.7520	0.1190
136	1.5500	0.6530	0.1720	166	0.7680	-0.1090	0.2060
137	0.6360	-0.9130	0.2420	167	1.3710	0.5070	0.1430
138	0.9430	0.0640	0.1880	168	1.0540	1.3230	0.1800
139	1.3190	0.3820	0.1630	169	1.0780	0.3920	0.2050
140	1.4810	1.7050	0.1980	170	1.1450	-0.1610	0.2940
141	1.0130	1.6520	0.2130	171	1.4030	1.3920	0.1070
142	0.9510	1.6840	0.2410	172	1.7040	1.6070	0.1680
143	1.2570	1.6850	0.1510	173	1.8840	0.9110	0.1580
144	1.3730	1.1650	0.1500	174	1.0260	-0.5380	0.1870

Item	a	b	c	Item	a	b	c
175	1.1980	0.1400	0.2360	205	0.9240	-0.2410	0.2470
176	1.2020	-0.7070	0.1430	206	0.7590	0.2860	0.1860
177	2.1500	2.1890	0.2050	207	0.9640	0.3640	0.1940
178	1.4200	1.4630	0.1110	208	1.0560	-0.1960	0.1630
179	1.1330	0.9000	0.2890	209	1.3300	0.2850	0.1770
180	0.6460	-0.4650	0.2110	210	0.9600	0.2280	0.2000
181	1.5520	0.9760	0.2450	211	1.3860	0.2380	0.1830
182	0.9490	-0.5170	0.1850	212	1.6190	-0.3580	0.1570
183	0.5750	-0.3780	0.2020	213	0.4510	-1.8010	0.2610
184	0.5230	-0.2200	0.2220	214	1.5900	0.6020	0.2690
185	1.2660	1.5390	0.1990	215	1.0210	0.1710	0.2190
186	1.2830	1.1420	0.2290	216	1.2080	0.1550	0.3060
187	2.0700	1.5760	0.1450	217	0.7580	0.1650	0.1820
188	1.0450	0.7640	0.3000	218	1.5960	0.4140	0.2840
189	1.0640	1.3630	0.2670	219	1.4670	1.0050	0.1360
190	0.9850	0.1730	0.2090	220	1.0760	0.6730	0.1780
191	1.0350	0.3440	0.2040	221	0.9270	-0.0490	0.1110
192	1.1690	0.1580	0.2390	222	0.8980	-0.6360	0.2050
193	1.1660	1.0680	0.2120	223	0.7340	0.7740	0.1800
194	1.6680	0.7140	0.1770	224	0.9600	0.8750	0.2320
195	1.0290	-1.4300	0.2060	225	1.1120	0.4860	0.2430
196	0.8190	-0.1510	0.1870	226	0.8110	-0.3250	0.2490
197	1.2080	1.0840	0.1170	227	0.5180	0.0560	0.2960
198	1.4940	2.1440	0.0940	228	0.9640	0.3920	0.3010
199	1.5810	1.6540	0.2560	229	1.0760	-0.2700	0.2400
200	1.4550	1.5200	0.1120	230	1.2430	0.1120	0.2130
201	0.9030	-1.9320	0.2390	231	0.1324	1.1630	0.3000
202	0.6330	0.0350	0.2180	232	0.8160	-0.4970	0.2290
203	1.2120	1.2110	0.1880	233	1.3680	-0.4470	0.2240
204	0.8500	0.8620	0.3000	234	1.4060	0.7500	0.1400

Item	a	b	c	Item	a	b	c
235	1.0510	2.2600	0.2390	240	1.3660	1.3620	0.1350
236	1.3160	1.2890	0.1510	241	1.0260	2.2880	0.2480
237	1.2700	0.4050	0.1760	242	1.3570	1.7060	0.2160
238	1.0244	2.1980	0.2800	243	1.1640	1.3890	0.2570
239	0.9700	1.5150	0.2310	244	0.9710	1.7110	0.2160

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ช

ผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ช

ตัวอย่างผลการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ อัตราการใช้ข้อสอบซ้ำ10 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ในครั้งที่ 1

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: พัฑฒิดา หิรัญภูวิศ
ชั้น/ห้อง: 6/5 โรงเรียน: สุขานารี
วันที่: 16/11/2006 เวลา: 13:58:17

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 11111-1
เวลาที่ใช้ในการทดสอบ: 0:07:48
จำนวนข้อที่ตอบถูก: 4 จาก 7
ตำแหน่งเปอร์เซ็นไทล์: 61
ความสามารถในการสอบ: 0.2914 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 4.0582
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.4452

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	212	ค	ถูก	0.4021	0.6127	0.8568	0.0641	0.0641
2	165	ก	ผิด	-0.0301	0.8595	0.6608	0.0559	0.1199
3	164	ข	ถูก	0.2366	1.9992	0.6123	0.0397	0.1596
4	157	ข	ผิด	0.1742	2.0054	0.5674	0.0000	0.1597
5	76	ก	ถูก	0.3045	2.7572	0.5260	0.0068	0.1665
6	173	ก	ผิด	0.1702	3.0072	0.4672	0.0030	0.1695
7	79	ข	ถูก	0.2914	4.0582	0.4452	0.0077	0.1773

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ในครั้งที่ 2

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: พัฑฒิดา หิรัญภูวิศ
ชั้น/ห้อง: 6/5 โรงเรียน: สุขานารี
วันที่: 27/11/2006 เวลา: 11:08:07

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 11111-2
เวลาที่ใช้ในการทดสอบ: 0:15:37
จำนวนข้อที่ตอบถูก: 8 จาก 16
ตำแหน่งเปอร์เซ็นไทล์: 60
ความสามารถในการสอบ: 0.2588 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 5.8105
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.4328

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	157	ข	ผิด	-0.1763	0.0004	0.8714	0.0000	0.0000
2	86	ค	ผิด	-0.4006	0.0426	0.7511	0.0018	0.0018
3	125	ข	ผิด	-0.6714	0.3224	0.6505	0.0138	0.0156
4	218	ก	ถูก	-0.4930	0.4106	0.7012	0.0010	0.0165
5	209	ค	ถูก	-0.1669	0.9596	0.7095	0.0213	0.0378
6	83	ก	ถูก	0.1343	1.5624	0.7327	0.0179	0.0557
7	133	ค	ผิด	-0.1418	1.8352	0.6240	0.0152	0.0709
8	167	ข	ผิด	-0.3134	2.1006	0.5594	0.0049	0.0758
9	139	ง	ถูก	-0.0758	2.6622	0.5610	0.0128	0.0886
10	214	ค	ผิด	-0.2124	2.8019	0.5050	0.0017	0.0903
11	156	ง	ถูก	-0.0596	3.4897	0.4982	0.0072	0.0975
12	234	ก	ผิด	-0.1492	3.6984	0.4651	0.0010	0.0985
13	124	ค	ถูก	-0.0012	4.3382	0.4609	0.0063	0.1048
14	110	ก	ถูก	0.2003	4.6753	0.4830	0.0051	0.1099
15	237	ข	ถูก	0.3573	5.4766	0.4662	0.0094	0.1193
16	219	ข	ผิด	0.2588	5.8105	0.4328	0.0019	0.1211

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ในครั้งที่ 1

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: พัฑฒิดา หิรัญภูวิศ
ชั้น/ห้อง: 6/5 โรงเรียน: สุขานารี
วันที่: 16/11/2006 เวลา: 13:58:17

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 11111-1
เวลาที่ใช้ในการทดสอบ: 0:14:20
จำนวนข้อที่ตอบถูก: 6 จาก 15
ตำแหน่งเปอร์เซ็นไทล์: 52
ความสามารถในการสอบ: 0.0630 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 8.7254
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2978

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	212	ค	ถูก	0.4021	0.6127	0.8568	0.0641	0.0641
2	165	ก	ผิด	-0.0301	0.8595	0.6608	0.0559	0.1199
3	164	ข	ถูก	0.2366	1.9992	0.6123	0.0397	0.1596
4	157	ข	ผิด	0.1742	2.0054	0.5674	0.0000	0.1597
5	76	ก	ถูก	0.3045	2.7572	0.5260	0.0068	0.1665
6	173	ก	ผิด	0.1702	3.0072	0.4672	0.0030	0.1695
7	79	ข	ถูก	0.2914	4.0582	0.4452	0.0077	0.1773
8	131	ง	ผิด	0.2169	4.1997	0.4127	0.0005	0.1777
9	211	ข	ถูก	0.3314	5.1765	0.3972	0.0063	0.1840
10	194	ง	ผิด	0.2232	5.7931	0.3653	0.0041	0.1882
11	58	ก	ผิด	0.1283	6.5159	0.3390	0.0037	0.1918
12	136	ง	ถูก	0.2378	7.2297	0.3396	0.0038	0.1956
13	86	ค	ผิด	0.1790	7.6458	0.3231	0.0008	0.1964
14	57	ข	ผิด	0.0972	8.4755	0.3074	0.0029	0.1993
15	70	ก	ผิด	0.0630	8.7254	0.2978	0.0002	0.1995

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ อัตราการใช้ข้อสอบซ้ำ 10 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ในครั้งที่ 2

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: พัฑฒิดา หิรัญภูวิศ
ชั้น/ห้อง: 6/5 โรงเรียน: สุขานารี
วันที่: 27/11/2006 เวลา: 11:08:07

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 11111-2
เวลาที่ใช้ในการทดสอบ: 0:18:50
จำนวนข้อที่ตอบถูก: 13 จาก 27
ตำแหน่งเปอร์เซ็นไทล์: 54
ความสามารถในการสอบ: 0.1070 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 11.3731
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2967

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	157	ข	ผิด	-0.1763	0.0004	0.8714	0.0000	0.0000
2	86	ค	ผิด	-0.4006	0.0426	0.7511	0.0018	0.0018
3	125	ข	ผิด	-0.6714	0.3224	0.6505	0.0138	0.0156
4	218	ก	ถูก	-0.4930	0.4106	0.7012	0.0010	0.0165
5	209	ค	ถูก	-0.1669	0.9596	0.7095	0.0213	0.0378
6	83	ก	ถูก	0.1343	1.5624	0.7327	0.0179	0.0557
7	133	ค	ผิด	-0.1418	1.8352	0.6240	0.0152	0.0709
8	167	ข	ผิด	-0.3134	2.1006	0.5594	0.0049	0.0758
9	139	ง	ถูก	-0.0758	2.6622	0.5610	0.0128	0.0886
10	214	ค	ผิด	-0.2124	2.8019	0.5050	0.0017	0.0903
11	156	ง	ถูก	-0.0596	3.4897	0.4982	0.0072	0.0975
12	234	ก	ผิด	-0.1492	3.6984	0.4651	0.0010	0.0985
13	124	ค	ถูก	-0.0012	4.3382	0.4609	0.0063	0.1048
14	110	ก	ถูก	0.2003	4.6753	0.4830	0.0051	0.1099
15	237	ข	ถูก	0.3573	5.4766	0.4662	0.0094	0.1193
16	219	ข	ผิด	0.2588	5.8105	0.4328	0.0019	0.1211

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
17	135	ก	ถูก	0.3856	6.5865	0.4214	0.0059	0.1271
18	95	ข	ผิด	0.2634	7.0213	0.3868	0.0038	0.1309
19	230	ข	ผิด	0.1111	7.7453	0.3616	0.0087	0.1395
20	25	ค	ผิด	0.0623	7.9978	0.3478	0.0003	0.1399
21	109	ก	ถูก	0.1258	8.7434	0.3367	0.0015	0.1414
22	40	ข	ผิด	0.0765	9.0943	0.3250	0.0004	0.1418
23	175	ง	ผิด	-0.0123	9.6794	0.3114	0.0024	0.1442
24	23	ค	ถูก	0.0408	9.8466	0.3190	0.0002	0.1444
25	126	ค	ถูก	0.0889	10.6229	0.3103	0.0009	0.1454
26	132	ก	ผิด	0.0707	10.7035	0.3035	0.0000	0.1454
27	233	ข	ถูก	0.1070	11.3731	0.2967	0.0004	0.1458

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ในครั้งที่ 1

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: เด็กชายพรรณเชษฐ กิจจะ
ชั้น/ห้อง: ป.6-1 โรงเรียน: วัดสระแก้ว
วันที่: 08/11/2006 เวลา: 13:00:05

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 12111-1
เวลาที่ใช้ในการทดสอบ: 00:10:53
จำนวนข้อที่ตอบถูก: 5 จาก 9
ตำแหน่งเปอร์เซ็นไทล์: 65
ความสามารถในการสอบ: 0.3989 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 6.8112
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.4079

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	212	ค	ถูก	0.4021	0.6127	0.8568	0.0641	0.0641
2	165	ง	ผิด	-0.0301	0.8595	0.6608	0.0559	0.1199
3	164	ข	ถูก	0.2366	1.9992	0.6123	0.0397	0.1596
4	157	ง	ผิด	0.1742	2.0054	0.5674	0.0000	0.1597
5	76	ก	ถูก	0.3045	2.7572	0.5260	0.0068	0.1665
6	173	ค	ถูก	0.5730	3.8702	0.5341	0.0277	0.1942
7	131	ก	ถูก	0.8093	5.0790	0.5325	0.0258	0.2200
8	94	ค	ผิด	0.6474	5.5494	0.4718	0.0078	0.2277
9	58	ก	ผิด	0.3989	6.8112	0.4079	0.0442	0.2719

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ในครั้งที่ 2

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: เด็กชายพรรณเชษฐ กิจจะ
ชั้น/ห้อง: ป.6-1 โรงเรียน: วักสระแก้ว
วันที่: 22/11/2006 เวลา: 12:10:14

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 12111-2
เวลาที่ใช้ในการทดสอบ: 0:04:46
จำนวนข้อที่ตอบถูก: 9 จาก 12
ตำแหน่งเปอร์เซ็นไทล์: 91
ความสามารถในการสอบ: 1.3733 (สูง)
สารสนเทศของแบบทดสอบ: 8.6928
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.4275

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	131	ค	ผิด	-0.2817	0.0090	0.8335	0.0008	0.0008
2	94	ง	ถูก	0.1408	0.0663	0.9777	0.0020	0.0028
3	136	ค	ผิด	-0.3015	0.1936	0.7762	0.0279	0.0307
4	211	ข	ถูก	0.0750	1.0552	0.7548	0.0480	0.0787
5	70	ค	ถูก	0.4959	1.8433	0.7627	0.0451	0.1237
6	83	ก	ถูก	0.7977	3.0082	0.7082	0.0522	0.1759
7	219	ก	ถูก	1.1446	4.1975	0.6470	0.0681	0.2440
8	152	ก	ผิด	0.8371	4.8227	0.5458	0.0391	0.2832
9	133	ง	ถูก	1.0003	5.7772	0.5077	0.0131	0.2963
10	234	ง	ถูก	1.1459	6.7159	0.4735	0.0103	0.3066
11	181	ค	ถูก	1.2663	7.7380	0.4501	0.0075	0.3141
12	153	ข	ถูก	1.3733	8.6928	0.4275	0.0056	0.3197

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ในครั้งที่ 1

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: เด็กชายพรรณเชษฐ กิจจะ
ชั้น/ห้อง: ป.6-1 โรงเรียน: วัดสระแก้ว
วันที่: 08/11/2006 เวลา: 13:00:05

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 12111-1
เวลาที่ใช้ในการทดสอบ: 00:17:17
จำนวนข้อที่ตอบถูก: 13 จาก 19
ตำแหน่งเปอร์เซ็นไทล์: 83
ความสามารถในการสอบ: 0.9909 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 15.2001
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2958

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	212	ค	ถูก	0.4021	0.6127	0.8568	0.0641	0.0641
2	165	ง	ผิด	-0.0301	0.8595	0.6608	0.0559	0.1199
3	164	ข	ถูก	0.2366	1.9992	0.6123	0.0397	0.1596
4	157	ง	ผิด	0.1742	2.0054	0.5674	0.0000	0.1597
5	76	ก	ถูก	0.3045	2.7572	0.5260	0.0068	0.1665
6	173	ค	ถูก	0.5730	3.8702	0.5341	0.0277	0.1942
7	131	ก	ถูก	0.8093	5.0790	0.5325	0.0258	0.2200
8	94	ค	ผิด	0.6474	5.5494	0.4718	0.0078	0.2277
9	58	ก	ผิด	0.3989	6.8112	0.4079	0.0442	0.2719
10	132	ข	ผิด	0.3360	7.0508	0.3849	0.0005	0.2725
11	79	ข	ถูก	0.4104	8.0097	0.3699	0.0027	0.2752
12	194	ข	ถูก	0.5392	9.2159	0.3620	0.0091	0.2843
13	52	ข	ถูก	0.6499	9.6329	0.3726	0.0021	0.2864
14	57	ก	ถูก	0.7354	10.6199	0.3567	0.0037	0.2901
15	86	ง	ถูก	0.8419	11.9160	0.3434	0.0072	0.2972
16	187	ก	ผิด	0.8006	12.1052	0.3256	0.0002	0.2974
17	211	ข	ถูก	0.8427	12.7843	0.3166	0.0006	0.2980
18	136	ง	ถูก	0.9089	13.9662	0.3056	0.0026	0.3006
19	110	ก	ถูก	0.9909	15.2001	0.2958	0.0041	0.3047

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ อัตราการใช้ข้อสอบซ้ำ 20 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ในครั้งที่ 2

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: เด็กชายพรรณเชษฐ กิจจะ
ชั้น/ห้อง: ป.6-1 โรงเรียน: วักสระแก้ว
วันที่: 22/11/2006 เวลา: 12:10:14

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 12111-2
เวลาที่ใช้ในการทดสอบ: 0:11:19
จำนวนข้อที่ตอบถูก: 12 จาก 16
ตำแหน่งเปอร์เซ็นไทล์: 86
ความสามารถในการสอบ: 1.0832 (สูง)
สารสนเทศของแบบทดสอบ: 10.4463
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2828

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	131	ค	ผิด	-0.2817	0.0090	0.8335	0.0008	0.0008
2	94	ง	ถูก	0.1408	0.0663	0.9777	0.0020	0.0028
3	136	ค	ผิด	-0.3015	0.1936	0.7762	0.0279	0.0307
4	211	ข	ถูก	0.0750	1.0552	0.7548	0.0480	0.0787
5	70	ค	ถูก	0.4959	1.8433	0.7627	0.0451	0.1237
6	83	ก	ถูก	0.7977	3.0082	0.7082	0.0522	0.1759
7	219	ก	ถูก	1.1446	4.1975	0.6470	0.0681	0.2440
8	152	ก	ผิด	0.8371	4.8227	0.5458	0.0391	0.2832
9	133	ง	ถูก	1.0003	5.7772	0.5077	0.0131	0.2963
10	234	ง	ถูก	1.1459	6.7159	0.4735	0.0103	0.3066
11	181	ค	ถูก	1.2663	7.7380	0.4501	0.0075	0.3141
12	153	ข	ถูก	1.3733	8.6928	0.4275	0.0056	0.3197
13	25	ก	ถูก	1.4554	9.4294	0.4074	0.0026	0.3222
14	214	ข	ถูก	1.4927	9.8245	0.3939	0.0003	0.3225
15	218	ก	ถูก	1.5145	10.0623	0.3847	0.0001	0.3226
16	125	ข	ผิด	1.0832	10.4463	0.2828	0.0283	0.3509

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ อัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ในครั้งที่ 1

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: มนัสวีร์ รัตนวิศ
ชั้น/ห้อง: 6/9 โรงเรียน: สุขานารี
วันที่: 17/11/2006 เวลา: 8:33:08

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 13111-1
เวลาที่ใช้ในการทดสอบ: 0:04:38
จำนวนข้อที่ตอบถูก: 4 จาก 7
ตำแหน่งเปอร์เซ็นไทล์: 61
ความสามารถในการสอบ: 0.2914 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 4.0582
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.4452

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	212	ค	ถูก	0.4021	0.6127	0.8568	0.0641	0.0641
2	165	ง	ผิด	-0.0301	0.8595	0.6608	0.0559	0.1199
3	164	ข	ถูก	0.2366	1.9992	0.6123	0.0397	0.1596
4	157	ข	ผิด	0.1742	2.0054	0.5674	0.0000	0.1597
5	76	ก	ถูก	0.3045	2.7572	0.5260	0.0068	0.1665
6	173	ง	ผิด	0.1702	3.0072	0.4672	0.0030	0.1695
7	79	ข	ถูก	0.2914	4.0582	0.4452	0.0077	0.1773

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ อัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ในครั้งที่ 2

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: มนัสวีร์ รัตนวิศ
ชั้น/ห้อง: 6-9 โรงเรียน: สุขานารี
วันที่: 28/11/2006 เวลา: 10:06:35

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 13111-2
เวลาที่ใช้ในการทดสอบ: 0:06:12
จำนวนข้อที่ตอบถูก: 6 จาก 10
ตำแหน่งเปอร์เซ็นไทล์: 69
ความสามารถในการสอบ: 0.5054 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 6.5582
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.4014

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	157	ก	ถูก	0.6962	0.2478	1.1588	0.0091	0.0091
2	76	ก	ถูก	0.9844	0.4095	1.0180	0.0087	0.0178
3	83	ก	ถูก	1.3279	0.9474	0.9003	0.0395	0.0573
4	132	ก	ผิด	0.6621	1.6221	0.6666	0.2551	0.3124
5	194	ค	ผิด	0.2847	2.3519	0.5396	0.0754	0.3878
6	70	ค	ถูก	0.5370	3.2001	0.5351	0.0215	0.4092
7	234	ค	ผิด	0.3272	3.8944	0.4756	0.0178	0.4270
8	133	ง	ถูก	0.5051	4.8983	0.4598	0.0145	0.4416
9	219	ก	ถูก	0.7039	5.8070	0.4505	0.0154	0.4570
10	95	ข	ผิด	0.5054	6.5582	0.4014	0.0172	0.4742

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ อัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ในครั้งที่ 1

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: มนัสวีร์ รัตนวิศ
ชั้น/ห้อง: 6/9 โรงเรียน: สุขานารี
วันที่: 17/11/2006 เวลา: 8:33:08

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 13111-1
เวลาที่ใช้ในการทดสอบ: 0:10:33
จำนวนข้อที่ตอบถูก: 8 จาก 17
ตำแหน่งเปอร์เซ็นไทล์: 60
ความสามารถในการสอบ: 0.2548 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 10.6855
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2940

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	212	ค	ถูก	0.4021	0.6127	0.8568	0.0641	0.0641
2	165	ง	ผิด	-0.0301	0.8595	0.6608	0.0559	0.1199
3	164	ข	ถูก	0.2366	1.9992	0.6123	0.0397	0.1596
4	157	ข	ผิด	0.1742	2.0054	0.5674	0.0000	0.1597
5	76	ก	ถูก	0.3045	2.7572	0.5260	0.0068	0.1665
6	173	ง	ผิด	0.1702	3.0072	0.4672	0.0030	0.1695
7	79	ข	ถูก	0.2914	4.0582	0.4452	0.0077	0.1773
8	131	ค	ผิด	0.2169	4.1997	0.4127	0.0005	0.1777
9	211	ข	ถูก	0.3314	5.1765	0.3972	0.0063	0.1840
10	194	ง	ผิด	0.2232	5.7931	0.3653	0.0041	0.1882
11	58	ก	ผิด	0.1283	6.5159	0.3390	0.0037	0.1918
12	136	ง	ถูก	0.2378	7.2297	0.3396	0.0038	0.1956
13	86	ง	ถูก	0.3525	7.9011	0.3418	0.0039	0.1995
14	83	ก	ถูก	0.4440	8.9781	0.3359	0.0042	0.2037
15	132	ก	ผิด	0.3984	9.2789	0.3217	0.0003	0.2041
16	57	ค	ผิด	0.2955	10.2903	0.3044	0.0055	0.2096
17	110	ง	ผิด	0.2548	10.6855	0.2940	0.0003	0.2099

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของฟิชเชอร์ อัตราการใช้ข้อสอบซ้ำ 30 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ในครั้งที่ 2

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: มนัสวีร์ รัตนวิศ
ชั้น/ห้อง: 6-9 โรงเรียน: สุขานารี
วันที่: 28/11/2006 เวลา: 10:06:35
ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 13111-2
เวลาที่ใช้ในการทดสอบ: 0:10:43
จำนวนข้อที่ตอบถูก: 12 จาก 19
ตำแหน่งเปอร์เซ็นไทล์: 77
ความสามารถในการสอบ: 0.7496 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 13.9559
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2965

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	157	ก	ถูก	0.6962	0.2478	1.1588	0.0091	0.0091
2	76	ก	ถูก	0.9844	0.4095	1.0180	0.0087	0.0178
3	83	ก	ถูก	1.3279	0.9474	0.9003	0.0395	0.0573
4	132	ก	ผิด	0.6621	1.6221	0.6666	0.2551	0.3124
5	194	ค	ผิด	0.2847	2.3519	0.5396	0.0754	0.3878
6	70	ค	ถูก	0.5370	3.2001	0.5351	0.0215	0.4092
7	234	ค	ผิด	0.3272	3.8944	0.4756	0.0178	0.4270
8	133	ง	ถูก	0.5051	4.8983	0.4598	0.0145	0.4416
9	219	ก	ถูก	0.7039	5.8070	0.4505	0.0154	0.4570
10	95	ข	ผิด	0.5054	6.5582	0.4014	0.0172	0.4742
11	214	ข	ถูก	0.6100	7.6150	0.3913	0.0055	0.4797
12	94	ข	ผิด	0.5448	7.9431	0.3681	0.0008	0.4805
13	218	ก	ถูก	0.6204	8.9924	0.3567	0.0030	0.4835
14	181	ก	ผิด	0.5309	9.5211	0.3343	0.0024	0.4858
15	167	ก	ถูก	0.6128	10.5522	0.3236	0.0034	0.4893
16	153	ข	ถูก	0.6990	11.2902	0.3222	0.0025	0.4918
17	25	ก	ถูก	0.7856	12.2370	0.3146	0.0034	0.4952
18	40	ข	ผิด	0.6922	13.1190	0.3001	0.0039	0.4992
19	23	ค	ถูก	0.7496	13.9559	0.2965	0.0013	0.5005

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ10 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ในครั้งที่ 1

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: รัฐศาสตร์ สระทองหลาง
ชั้น/ห้อง: 6/8 โรงเรียน: วัดสระแก้ว
วันที่: 10/11/2006 เวลา: 10:58:14

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 31111-1
เวลาที่ใช้ในการทดสอบ: 00:05:02
จำนวนข้อที่ตอบถูก: 6 จาก 7
ตำแหน่งเปอร์เซ็นไทล์: 71
ความสามารถในการสอบ: 0.5750 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 4.0103
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.4498

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	55	ก	ถูก	0.3003	0.3239	0.8893	0.0165	0.0165
2	176	ข	ถูก	0.4915	0.5826	0.8073	0.0053	0.0218
3	212	ค	ถูก	0.6747	0.9120	0.7211	0.0064	0.0282
4	164	ข	ถูก	0.8520	1.4290	0.6537	0.0091	0.0373
5	165	ง	ผิด	0.4035	2.6940	0.5036	0.1834	0.2206
6	76	ก	ถูก	0.4818	3.2275	0.4750	0.0017	0.2224
7	79	ข	ถูก	0.5750	4.0103	0.4498	0.0035	0.2259

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ10 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ในครั้งที่ 2

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: รัฐศาสตร์ สระทองหลาง
ชั้น/ห้อง: 6-8 โรงเรียน: วัดสระแก้ว
วันที่: 22/11/2006 เวลา: 14:47:09

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 31111-2
เวลาที่ใช้ในการทดสอบ: 00:03:21
จำนวนข้อที่ตอบถูก: 7 จาก 9
ตำแหน่งเปอร์เซ็นไทล์: 62
ความสามารถในการสอบ: 0.3204 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 6.6275
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.4376

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	164	ข	ถูก	0.4662	0.9542	0.8920	0.1123	0.1123
2	109	ข	ผิด	-0.3447	1.7974	0.6885	0.2728	0.3850
3	54	ข	ถูก	-0.1013	2.4599	0.6402	0.0198	0.4048
4	233	ข	ถูก	0.0690	3.1594	0.5961	0.0106	0.4154
5	77	ค	ถูก	0.2051	3.7713	0.5614	0.0059	0.4213
6	15	ก	ถูก	0.3197	4.2952	0.5332	0.0035	0.4249
7	156	ง	ถูก	0.4512	5.0721	0.5080	0.0068	0.4317
8	139	ก	ผิด	0.2117	5.8912	0.4537	0.0255	0.4571
9	230	ง	ถูก	0.3204	6.6275	0.4376	0.0043	0.4615

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ10 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ในครั้งที่ 1

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: รัฐศาสตร์ สระทองหลาง
ชั้น/ห้อง: 6/8 โรงเรียน: วัดสระแก้ว
วันที่: 10/11/2006 เวลา: 10:58:14

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 31111-1
เวลาที่ใช้ในการทดสอบ: 00:09:01
จำนวนข้อที่ตอบถูก: 9 จาก 14
ตำแหน่งเปอร์เซ็นไทล์: 58
ความสามารถในการสอบ: 0.2253 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 10.6327
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2942

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	55	ก	ถูก	0.3003	0.3239	0.8893	0.0165	0.0165
2	176	ข	ถูก	0.4915	0.5826	0.8073	0.0053	0.0218
3	212	ค	ถูก	0.6747	0.9120	0.7211	0.0064	0.0282
4	164	ข	ถูก	0.8520	1.4290	0.6537	0.0091	0.0373
5	165	ง	ผิด	0.4035	2.6940	0.5036	0.1834	0.2206
6	76	ก	ถูก	0.4818	3.2275	0.4750	0.0017	0.2224
7	79	ข	ถูก	0.5750	4.0103	0.4498	0.0035	0.2259
8	57	ก	ถูก	0.7036	5.0187	0.4242	0.0085	0.2344
9	58	ก	ผิด	0.4818	6.4079	0.3732	0.0369	0.2713
10	211	ข	ถูก	0.5595	7.3157	0.3595	0.0028	0.2741
11	167	ง	ผิด	0.4261	8.2993	0.3359	0.0091	0.2831
12	125	ข	ผิด	0.2876	9.1817	0.3167	0.0084	0.2915
13	126	ค	ถูก	0.3224	9.7621	0.3090	0.0004	0.2919
14	209	ข	ผิด	0.2253	10.6327	0.2942	0.0042	0.2961

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ10 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ในครั้งที่ 2

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: รัฐศาสตร์ สระทองหลาง
ชั้น/ห้อง: 6-8 โรงเรียน: วัดสระแก้ว
วันที่: 22/11/2006 เวลา: 14:47:09

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 31111-2
เวลาที่ใช้ในการทดสอบ: 00:07:11
จำนวนข้อที่ตอบถูก: 12 จาก 19
ตำแหน่งเปอร์เซ็นไทล์: 48
ความสามารถในการสอบ: -0.0367 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 12.5371
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2984

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	164	ข	ถูก	0.4662	0.9542	0.8920	0.1123	0.1123
2	109	ข	ผิด	-0.3447	1.7974	0.6885	0.2728	0.3850
3	54	ข	ถูก	-0.1013	2.4599	0.6402	0.0198	0.4048
4	233	ข	ถูก	0.0690	3.1594	0.5961	0.0106	0.4154
5	77	ค	ถูก	0.2051	3.7713	0.5614	0.0059	0.4213
6	15	ก	ถูก	0.3197	4.2952	0.5332	0.0035	0.4249
7	156	ง	ถูก	0.4512	5.0721	0.5080	0.0068	0.4317
8	139	ก	ผิด	0.2117	5.8912	0.4537	0.0255	0.4571
9	230	ง	ถูก	0.3204	6.6275	0.4376	0.0043	0.4615
10	124	ก	ผิด	0.1395	7.3789	0.4002	0.0132	0.4747
11	78	ก	ถูก	0.2061	7.9758	0.3893	0.0013	0.4760
12	208	ก	ผิด	0.0487	8.5575	0.3687	0.0072	0.4832
13	75	ข	ถูก	0.1106	9.1215	0.3605	0.0011	0.4843
14	14	ข	ถูก	0.1751	9.6991	0.3529	0.0012	0.4855
15	175	ค	ถูก	0.2415	10.3567	0.3461	0.0014	0.4869
16	237	ก	ผิด	0.1467	11.0368	0.3291	0.0032	0.4901
17	229	ง	ผิด	0.0275	11.5555	0.3156	0.0037	0.4938
18	174	ข	ถูก	0.0630	12.0016	0.3106	0.0003	0.4940
19	170	ก	ผิด	-0.0367	12.5371	0.2984	0.0027	0.4967

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ20 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ในครั้งที่ 1

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: เด็กหญิงณัฏฐณิชา ช่วยค้ำชู
ชั้น/ห้อง: ป.6/4 โรงเรียน: วัดสระแก้ว
วันที่: 09/11/2006 เวลา: 09:46:55

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 32111-1
เวลาที่ใช้ในการทดสอบ: 00:05:44
จำนวนข้อที่ตอบถูก: 6 จาก 7
ตำแหน่งเปอร์เซ็นไทล์: 71
ความสามารถในการสอบ: 0.5750 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 4.0103
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.4498

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	55	ก	ถูก	0.3003	0.3239	0.8893	0.0165	0.0165
2	176	ข	ถูก	0.4915	0.5826	0.8073	0.0053	0.0218
3	212	ค	ถูก	0.6747	0.9120	0.7211	0.0064	0.0282
4	164	ข	ถูก	0.8520	1.4290	0.6537	0.0091	0.0373
5	165	ง	ผิด	0.4035	2.6940	0.5036	0.1834	0.2206
6	76	ก	ถูก	0.4818	3.2275	0.4750	0.0017	0.2224
7	79	ข	ถูก	0.5750	4.0103	0.4498	0.0035	0.2259

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ20 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ในครั้งที่ 2

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: เด็กหญิงณัฏฐณิชา ช่วยค้ำชู
ชั้น/ห้อง: 6-4 โรงเรียน: วัดสระแก้ว
วันที่: 23/11/2006 เวลา: 09:06:03

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 32111-2
เวลาที่ใช้ในการทดสอบ: 00:05:41
จำนวนข้อที่ตอบถูก: 6 จาก 7
ตำแหน่งเปอร์เซ็นไทล์: 96
ความสามารถในการสอบ: 1.7626 (สูง)
สารสนเทศของแบบทดสอบ: 8.6352
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.4175

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	165	ข	ถูก	0.7322	2.0994	0.9476	0.3169	0.3169
2	125	ค	ถูก	1.0281	2.5197	0.8379	0.0214	0.3383
3	57	ก	ถูก	1.2804	3.0014	0.7411	0.0177	0.3560
4	173	ค	ถูก	1.5425	3.8746	0.6455	0.0364	0.3924
5	157	ก	ถูก	1.8428	5.6132	0.5458	0.0985	0.4909
6	187	ง	ถูก	2.0305	7.0514	0.4844	0.0287	0.5195
7	145	ง	ผิด	1.7626	8.6352	0.4175	0.0606	0.5802

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ20 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ในครั้งที่ 1

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: เด็กหญิงณัฏฐณิชา ช่วยค้ำชู
ชั้น/ห้อง: ป.6/4 โรงเรียน: วัดสระแก้ว
วันที่: 09/11/2006 เวลา: 09:46:55

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 32111-1
เวลาที่ใช้ในการทดสอบ: 00:08:45
จำนวนข้อที่ตอบถูก: 10 จาก 14
ตำแหน่งเปอร์เซ็นไทล์: 66
ความสามารถในการสอบ: 0.4162 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 11.0124
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2943

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	55	ก	ถูก	0.3003	0.3239	0.8893	0.0165	0.0165
2	176	ข	ถูก	0.4915	0.5826	0.8073	0.0053	0.0218
3	212	ค	ถูก	0.6747	0.9120	0.7211	0.0064	0.0282
4	164	ข	ถูก	0.8520	1.4290	0.6537	0.0091	0.0373
5	165	ง	ผิด	0.4035	2.6940	0.5036	0.1834	0.2206
6	76	ก	ถูก	0.4818	3.2275	0.4750	0.0017	0.2224
7	79	ข	ถูก	0.5750	4.0103	0.4498	0.0035	0.2259
8	57	ก	ถูก	0.7036	5.0187	0.4242	0.0085	0.2344
9	58	ก	ผิด	0.4818	6.4079	0.3732	0.0369	0.2713
10	211	ข	ถูก	0.5595	7.3157	0.3595	0.0028	0.2741
11	167	ค	ผิด	0.4261	8.2993	0.3359	0.0091	0.2831
12	125	ค	ถูก	0.4866	9.1067	0.3259	0.0015	0.2846
13	209	ค	ถูก	0.5484	9.9848	0.3169	0.0017	0.2863
14	218	ค	ผิด	0.4162	11.0124	0.2943	0.0094	0.2957

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ20 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ในครั้งที่ 2

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: เด็กหญิงณัฏฐณิชา ช่วยค้ำชู
ชั้น/ห้อง: 6-4 โรงเรียน: วัดสระแก้ว
วันที่: 23/11/2006 เวลา: 09:06:03

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 32111-2
เวลาที่ใช้ในการทดสอบ: 00:07:27
จำนวนข้อที่ตอบถูก: 11 จาก 13
ตำแหน่งเปอร์เซ็นไทล์: 97
ความสามารถในการสอบ: 1.9736 (สูง)
สารสนเทศของแบบทดสอบ: 15.4514
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2967

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	165	ข	ถูก	0.7322	2.0994	0.9476	0.3169	0.3169
2	125	ค	ถูก	1.0281	2.5197	0.8379	0.0214	0.3383
3	57	ก	ถูก	1.2804	3.0014	0.7411	0.0177	0.3560
4	173	ค	ถูก	1.5425	3.8746	0.6455	0.0364	0.3924
5	157	ก	ถูก	1.8428	5.6132	0.5458	0.0985	0.4909
6	187	ง	ถูก	2.0305	7.0514	0.4844	0.0287	0.5195
7	145	ง	ผิด	1.7626	8.6352	0.4175	0.0606	0.5802
8	94	ง	ถูก	1.8504	9.6104	0.3912	0.0039	0.5841
9	41	ข	ถูก	1.9474	10.9300	0.3675	0.0064	0.5906
10	172	ก	ถูก	2.0248	12.1062	0.3487	0.0036	0.5942
11	37	ข	ถูก	2.0963	13.3197	0.3325	0.0032	0.5974
12	44	ข	ผิด	1.9214	14.4801	0.3070	0.0172	0.6146
13	200	ก	ถูก	1.9736	15.4514	0.2967	0.0014	0.6160

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ30 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ในครั้งที่ 1

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: สิรดนัย แสงไทยทวีพร
ชั้น/ห้อง: 6/4 โรงเรียน: สุขานารี
วันที่: 16/11/2006 เวลา: 13:17:15

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 33111-1
เวลาที่ใช้ในการทดสอบ: 0:03:42
จำนวนข้อที่ตอบถูก: 4 จาก 6
ตำแหน่งเปอร์เซ็นไทล์: 47
ความสามารถในการสอบ: -0.0519 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 3.8641
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.4168

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	55	ก	ถูก	0.3003	0.3239	0.8893	0.0165	0.0165
2	176	ข	ถูก	0.4915	0.5826	0.8073	0.0053	0.0218
3	212	ค	ถูก	0.6747	0.9120	0.7211	0.0064	0.0282
4	164	ข	ถูก	0.8520	1.4290	0.6537	0.0091	0.0373
5	165	ง	ผิด	0.4035	2.6940	0.5036	0.1834	0.2206
6	76	ค	ผิด	-0.0519	3.8641	0.4168	0.1073	0.3280

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ30 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.45 ในครั้งที่ 2

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: สิรดนัย แสงไทยทวีพร
ชั้น/ห้อง: 6-4 โรงเรียน: สุขานารี
วันที่: 27/11/2006 เวลา: 12:21:33

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 33111-2
เวลาที่ใช้ในการทดสอบ: 0:01:58
จำนวนข้อที่ตอบถูก: 6 จาก 8
ตำแหน่งเปอร์เซ็นไทล์: 69
ความสามารถในการสอบ: 0.5018 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 6.9158
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.4405

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	164	ง	ผิด	-0.6481	0.3214	0.7496	0.1463	0.1463
2	55	ก	ถูก	-0.3670	0.9307	0.6870	0.0247	0.1710
3	79	ข	ถูก	-0.0909	1.8806	0.6751	0.0293	0.2003
4	156	ง	ถูก	0.1442	2.6851	0.6509	0.0203	0.2206
5	78	ก	ถูก	0.3030	3.2437	0.6150	0.0072	0.2278
6	57	ก	ถูก	0.5653	4.3144	0.5670	0.0361	0.2639
7	58	ง	ถูก	0.7968	5.7721	0.5202	0.0392	0.3030
8	194	ง	ผิด	0.5018	6.9158	0.4405	0.0586	0.3616

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ30 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ในครั้งที่ 1

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: สิรดนัย แสงไทยทวีพร
ชั้น/ห้อง: 6/4 โรงเรียน: สุขานารี
วันที่: 16/11/2006 เวลา: 13:17:15

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 33111-1
เวลาที่ใช้ในการทดสอบ: 0:07:32
จำนวนข้อที่ตอบถูก: 13 จาก 17
ตำแหน่งเปอร์เซ็นไทล์: 58
ความสามารถในการสอบ: 0.2022 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 12.8566
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2969

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	55	ก	ถูก	0.3003	0.3239	0.8893	0.0165	0.0165
2	176	ข	ถูก	0.4915	0.5826	0.8073	0.0053	0.0218
3	212	ค	ถูก	0.6747	0.9120	0.7211	0.0064	0.0282
4	164	ข	ถูก	0.8520	1.4290	0.6537	0.0091	0.0373
5	165	ง	ผิด	0.4035	2.6940	0.5036	0.1834	0.2206
6	76	ค	ผิด	-0.0519	3.8641	0.4168	0.1073	0.3280
7	126	ก	ผิด	-0.2688	4.8070	0.3794	0.0222	0.3502
8	109	ง	ผิด	-0.3989	5.6340	0.3576	0.0072	0.3574
9	233	ข	ถูก	-0.3218	6.5140	0.3481	0.0026	0.3600
10	54	ข	ถูก	-0.2501	7.1999	0.3396	0.0018	0.3617
11	77	ค	ถูก	-0.1865	7.9793	0.3320	0.0016	0.3633
12	15	ก	ถูก	-0.1276	8.6566	0.3253	0.0012	0.3645
13	79	ข	ถูก	-0.0500	9.6479	0.3201	0.0029	0.3673
14	125	ค	ถูก	0.0293	10.4910	0.3132	0.0026	0.3699
15	78	ก	ถูก	0.0751	11.1232	0.3078	0.0007	0.3706
16	211	ข	ถูก	0.1486	12.0380	0.3020	0.0024	0.3730
17	156	ง	ถูก	0.2022	12.8566	0.2969	0.0012	0.3741

เมื่อใช้เกณฑ์การคัดเลือกข้อสอบขั้นแรกด้วยวิธีสารสนเทศของคูลเบค – ไลเบลอร์ อัตราการใช้ข้อสอบซ้ำ30 เปอร์เซ็นต์ เกณฑ์ยุติการทดสอบที่กำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.30 ในครั้งที่ 2

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
ชื่อผู้ทดสอบ: สิรดนัย แสงไทยทวีพร
ชั้น/ห้อง: 6-4 โรงเรียน: สุขานารี
วันที่: 27/11/2006 เวลา: 12:21:33

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 33111-2
เวลาที่ใช้ในการทดสอบ: 0:04:09
จำนวนข้อที่ตอบถูก: 11 จาก 16
ตำแหน่งเปอร์เซ็นไทล์: 67
ความสามารถในการสอบ: 0.4498 (ปานกลาง)
สารสนเทศของแบบทดสอบ: 14.5607
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2948

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE	KL	SumKL
1	164	ง	ผิด	-0.6481	0.3214	0.7496	0.1463	0.1463
2	55	ก	ถูก	-0.3670	0.9307	0.6870	0.0247	0.1710
3	79	ข	ถูก	-0.0909	1.8806	0.6751	0.0293	0.2003
4	156	ง	ถูก	0.1442	2.6851	0.6509	0.0203	0.2206
5	78	ก	ถูก	0.3030	3.2437	0.6150	0.0072	0.2278
6	57	ก	ถูก	0.5653	4.3144	0.5670	0.0361	0.2639
7	58	ง	ถูก	0.7968	5.7721	0.5202	0.0392	0.3030
8	194	ง	ผิด	0.5018	6.9158	0.4405	0.0586	0.3616
9	209	ก	ผิด	0.2874	7.8102	0.3999	0.0212	0.3828
10	230	ง	ถูก	0.3696	8.5369	0.3877	0.0025	0.3852
11	139	ง	ถูก	0.4718	9.4559	0.3744	0.0047	0.3900
12	167	ก	ถูก	0.5767	10.4868	0.3604	0.0056	0.3955
13	218	ง	ผิด	0.4066	11.5084	0.3284	0.0156	0.4111
14	124	ค	ถูก	0.4691	12.3531	0.3206	0.0016	0.4128
15	83	ก	ถูก	0.5465	13.5258	0.3134	0.0034	0.4162
16	136	ข	ผิด	0.4498	14.5607	0.2948	0.0051	0.4213

ภาคผนวก ซ

แบบประเมินผลการใช้โปรแกรมและ
ความพึงพอใจของผู้ใช้โปรแกรมแบบปรับเหมาะ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

แบบประเมินผลการใช้โปรแกรมและความพึงพอใจของผู้ใช้โปรแกรมแบบปรับเหมาะ

<u>คำชี้แจง</u>แบบสอบถามฉบับนี้เป็นแบบสอบถามความคิดเห็นและความพึงพอใจของนักเรียนที่ใช้โปรแกรมแบบปรับเหมาะที่พัฒนาโดยเกียรติศักดิ์ ส่องแสง (2547) ซึ่งผู้วิจัยได้นำมาปรับให้เข้ากับเงื่อนไขที่ใช้ในการวิจัย

ขอให้นักเรียนอ่านและพิจารณาข้อความต่อไปนี้ ว่านักเรียนมีความคิดเห็นและความพึงพอใจต่อการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มากน้อยเพียงใด โดยทำเครื่องหมาย ✓ ลงในช่องแสดงระดับความคิดเห็นและความพึงพอใจดังนี้

5	หมายถึง	เห็นด้วยอย่างยิ่ง	หรือ	มีความพึงพอใจมาก
4	หมายถึง	เห็นด้วย	หรือ	มีความพึงพอใจ
3	หมายถึง	ไม่แน่ใจ	หรือ	มีความพึงพอใจปานกลาง
2	หมายถึง	ไม่เห็นด้วย	หรือ	ไม่ค่อยพึงพอใจ
1	หมายถึง	ไม่เห็นด้วยอย่างยิ่ง	หรือ	ไม่พึงพอใจอย่างยิ่ง

การใช้โปรแกรมในการทดสอบ					
ข้อคำถาม	ระดับความคิดเห็น				
	5	4	3	2	1
1. โปรแกรมมีความเหมาะสมและให้ความสะดวกในการทดสอบ					
2. มีคำอธิบายความหมายของสัญลักษณ์ที่ปรากฏในโปรแกรม					
3. เมนูคำสั่งการใช้โปรแกรมสามารถทำให้ใช้โปรแกรมได้อย่างถูกต้อง					
4. มีการป้องกันความผิดพลาดที่อาจจะเกิดขึ้นจากการใช้โปรแกรม					
5. มีการแสดงผลการตอบแบบทดสอบในแต่ละข้อได้เหมาะสมและชัดเจน					
6. สามารถให้ผู้ใช้สั่งแสดงผลการทดสอบทั้งทางจอภาพและทางเครื่องพิมพ์					
7. โปรแกรมมีความไวต่อการใช้งาน					
8. โปรแกรมง่ายต่อการใช้งาน					
9. โปรแกรมมีความทันสมัย					
10. การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ให้คะแนนที่เชื่อถือได้					
ความพึงพอใจต่อการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์					
ข้อความ	ระดับความพึงพอใจ				
	5	4	3	2	1
1. ผู้ดำเนินการสอบชี้แจงวิธีการสอบชัดเจน เข้าใจง่าย					
2. บรรยากาศภายในห้องสอบ					
3. ระยะเวลาที่ใช้ในทำข้อสอบ					
4. ลักษณะของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์					
5. การแสดงผลการสอบ					
6. การให้คะแนนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์					

<u>คำชี้แจง</u> ให้นักเรียนเขียนแสดงความคิดเห็นและความพึงพอใจในข้อคำถามต่อไปนี้

1. นักเรียนมีความรู้สึกอย่างไรกับ**การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์**ในครั้งนี้

.....

.....

.....

.....

2. นักเรียนมีความรู้สึกอย่างไรกับ**การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์** เมื่อเปรียบเทียบกับ**การทดสอบโดยใช้กระดาษ** (มีข้อดีและข้อเสียแตกต่างกันอย่างไร)

.....

.....

.....

.....

3. นักเรียนชอบการทดสอบแบบไหนมากกว่ากันระหว่าง**การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์** กับ **การทดสอบโดยใช้กระดาษ** เพราะเหตุใด

.....

.....

.....

.....

.....

4. นักเรียนคิดว่ามีข้อบกพร่องอะไรบ้างใน**การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์**ครั้งนี้

.....

.....

.....

.....

.....

5. ข้อเสนอแนะใน**การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์**

.....

.....

.....

.....

.....

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ซ
ผลการพิจารณาคุณภาพการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ นักเรียนชั้นประถมศึกษาปีที่ 6

จากการเก็บรวบรวมข้อมูลเกี่ยวกับความคิดเห็นด้านคุณภาพและความเหมาะสมของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ตามหลักสูตรกระทรวงศึกษาธิการกลุ่มสาระการเรียนรู้วิชาคณิตศาสตร์ระดับช่วงชั้นที่ 2 จากนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 913 คน ซึ่งผู้วิจัยทำการวิเคราะห์ข้อมูลเบื้องต้น โดยการวิเคราะห์จำนวนและร้อยละของผู้ตอบแบบสอบถามในการวิจัย เพื่อศึกษาสถานภาพโดยทั่วไปของกลุ่มตัวอย่างผู้ให้ข้อมูล ดังรายละเอียดในตารางที่ 45

และนำข้อมูลความคิดเห็นจากการตอบแบบสอบถามมาคำนวณค่าเฉลี่ย (Mean) และค่าส่วนเบี่ยงเบนมาตรฐาน (Standard deviation) โดยใช้เกณฑ์ในการประเมินคุณภาพดังนี้

4.51 – 5.00	หมายถึง	เห็นด้วย หรือ มีความพึงพอใจระดับมากที่สุด
3.51 – 4.50	หมายถึง	เห็นด้วย หรือ มีความพึงพอใจระดับมาก
2.51 – 3.50	หมายถึง	เห็นด้วย หรือ มีความพึงพอใจระดับปานกลาง
1.51 – 2.50	หมายถึง	เห็นด้วย หรือ มีความพึงพอใจระดับน้อย
1.00 – 1.50	หมายถึง	เห็นด้วย หรือ มีความพึงพอใจระดับน้อยที่สุด

ซึ่งได้แสดงไว้ในตารางที่ 46

นอกจากนั้นผู้วิจัยได้นำข้อเสนอแนะของนักเรียนมาระบุเพิ่มเติม เพื่อเป็นแนวทางสำหรับการพิจารณาความเหมาะสม ประกอบด้วยจุดแข็งและจุดด้อย อันจะนำไปสู่หนทางแก้ไขปรับปรุงทั้งทางด้านการพัฒนาโปรแกรมในการทดสอบ รูปแบบการนำเสนอ และความเหมาะสมกับการนำไปใช้จริง ดังแสดงไว้ในแผนภาพที่ 28 ถึง 31

ตารางที่ 45 สถานภาพของกลุ่มตัวอย่างผู้ให้ข้อมูลที่ใช้ในการวิจัย จำแนกตามเพศ

ตัวแปร	จำนวน	ร้อยละ
เพศ (N=913)		
1. ชาย	436	47.8
2. หญิง	458	50.2
3. ไม่ระบุ	19	2.1
รวม	913	100

จากตารางที่ 45 ผลการวิเคราะห์สถานภาพของผู้ตอบแบบสอบถามการประเมินคุณภาพการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ตามหลักสูตรกระทรวงศึกษาธิการ กลุ่มสาระการเรียนรู้วิชาคณิตศาสตร์ระดับช่วงชั้นที่ 2 ของนักเรียนชั้นประถมศึกษาปีที่ 6 พบว่า กลุ่มตัวอย่างผู้ให้ข้อมูลมีจำนวนทั้งสิ้น 913 โดยมีสัดส่วนเพศชายและหญิงใกล้เท่าเทียมกัน ประกอบด้วยเพศชาย จำนวน 436 คน คิดเป็นร้อยละ 47.8 เป็นเพศหญิง จำนวน 458 คน คิดเป็นร้อยละ 50.2

ตารางที่ 46 ค่าเฉลี่ยเลขคณิต และส่วนเบี่ยงเบนมาตรฐาน ของคุณภาพการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

องค์ประกอบ	ข้อคำถาม	M	SD
1. การใช้โปรแกรมในการทดสอบ			
1.) ลักษณะของโปรแกรมแบบปรับเหมาะ	1. โปรแกรมมีความเหมาะสมและให้ความสะดวกในการทดสอบ	4.55	.637
	2. มีคำอธิบายความหมายของสัญลักษณ์ที่ปรากฏในโปรแกรม	4.10	.718
	3. เมนูคำสั่งการใช้โปรแกรมสามารถทำให้ใช้โปรแกรมได้อย่างถูกต้อง	4.38	.747
2.) คุณภาพของโปรแกรมแบบปรับเหมาะ	4. มีการป้องกันความผิดพลาดที่อาจจะเกิดขึ้นจากการใช้โปรแกรม	3.96	.959
	5. มีการแสดงผลการตอบแบบทดสอบในแต่ละข้อได้เหมาะสมและชัดเจน	4.49	.753
	6. สามารถให้ผู้ใช้สั่งแสดงผลการทดสอบทั้งทางจอภาพและทางเครื่องพิมพ์	4.35	.765
	7. โปรแกรมมีความไวต่อการใช้งาน	4.52	.704
	8. โปรแกรมง่ายต่อการใช้งาน	4.58	1.197
	9. โปรแกรมมีความทันสมัย	4.64	.632
	10. การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ให้คะแนนที่เชื่อถือได้	4.41	.740

ตารางที่ 46 (ต่อ)

องค์ประกอบ	ข้อคำถาม	M	SD
2. ความพึงพอใจต่อการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์			
	1. ผู้ดำเนินการสอบชี้แจงวิธีการสอบชัดเจน เข้าใจง่าย	4.58	.638
	2. บรรยากาศภายในห้องสอบ	4.45	.739
	3. ระยะเวลาที่ใช้ในการทำข้อสอบ	4.24	.804
	4. ลักษณะของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์	4.42	.739
	5. การแสดงผลการสอบ	4.52	.737
	6. การให้คะแนนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์	4.47	.691

จากตารางที่ 46 ผลการประเมินการใช้โปรแกรมแบบปรับเหมาะและความพึงพอใจในการทดสอบ ซึ่งพิจารณาจากค่าสถิติพื้นฐานขององค์ประกอบการใช้โปรแกรมในการทดสอบ ด้านคุณลักษณะของโปรแกรมแบบปรับเหมาะ พบว่า โปรแกรมมีความเหมาะสมและให้ความสะดวกในการทดสอบ มีค่าเฉลี่ยสูงสุด (M = 4.55, SD = .637) รองลงมาได้แก่ เมนูคำสั่งการใช้โปรแกรมสามารถทำให้ใช้โปรแกรมได้อย่างถูกต้อง (M = 4.38, SD = .747) และมีคำอธิบายความหมายของสัญลักษณ์ที่ปรากฏในโปรแกรม (M = 4.10, SD = .718)

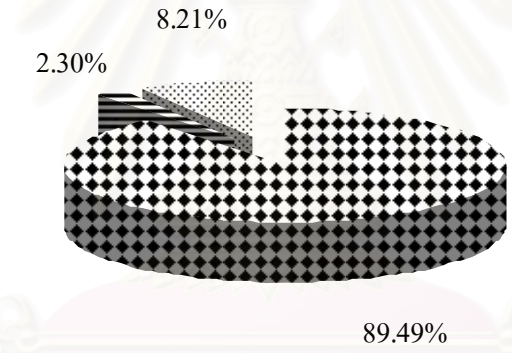
ด้านคุณภาพของโปรแกรมแบบปรับเหมาะ พบว่า โปรแกรมมีความทันสมัย มีค่าเฉลี่ยสูงสุด (M = 4.64, SD = .632) รองลงมาได้แก่ โปรแกรมง่ายต่อการใช้งาน (M = 4.58, SD = 1.197) โปรแกรมมีความไวต่อการใช้งาน (M = 4.52, SD = .704) มีการแสดงผลการตอบแบบทดสอบในแต่ละข้อได้เหมาะสมและชัดเจน (M = 4.49, SD = .753) การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ให้คะแนนที่เชื่อถือได้ (M = 4.41, SD = .740) สามารถให้ผู้ใช้สั่งแสดงผลการทดสอบทั้งทางจอภาพและทางเครื่องพิมพ์ (M = 4.35, SD = .765) และมีการป้องกันความผิดพลาดที่อาจจะเกิดขึ้นจากการใช้โปรแกรม (M = 3.96, SD = .959) ตามลำดับ

องค์ประกอบความพึงพอใจต่อการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ พบว่า ผู้ดำเนินการสอบชี้แจงวิธีการสอบชัดเจน เข้าใจง่าย (M = 4.58, SD = .638) รองลงมาได้แก่ การแสดงผลการสอบ (M = 4.52, SD = .737) การให้คะแนนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (M = 4.47, SD = .691) บรรยากาศภายในห้องสอบ (M = 4.45, SD = .739) ลักษณะ

ของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (M = 4.42, SD = .739) และระยะเวลาที่ใช้ในการทำข้อสอบ (M =4.24, SD = .804)

ตารางที่ 47 แสดงทัศนคติของนักเรียนกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เมื่อเปรียบเทียบกับการทดสอบโดยใช้กระดาษ

ทัศนคติ	จำนวน	ร้อยละ
พึงพอใจ	817	89.49
พึงพอใจทั้งสองรูปแบบ	21	2.30
ไม่พึงพอใจ	75	8.21
รวม	913	100



แผนภาพที่ 31 ทัศนคติของนักเรียนกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เมื่อเปรียบเทียบกับการทดสอบโดยใช้กระดาษ

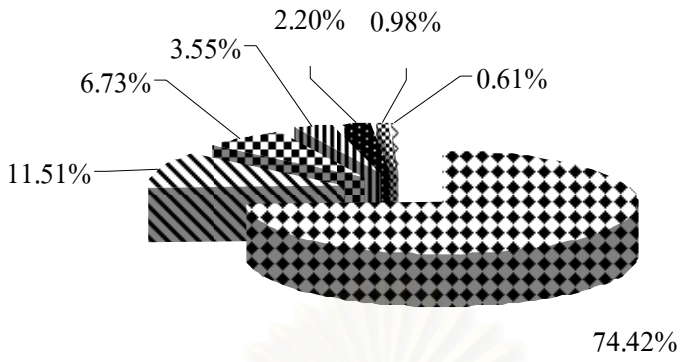
จากตารางที่ 47 และ แผนภาพที่ 28 ผลการประเมินความพึงพอใจในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เมื่อเปรียบเทียบกับการทดสอบโดยใช้กระดาษ พบว่า นักเรียนมีความพึงพอใจกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 817 คน (89.49%) มีความพึงพอใจเท่าเทียมกันระหว่างการทดสอบแบบปรับเหมาะโดยใช้กับคอมพิวเตอร์และกระดาษ จำนวน 21 คน (2.30%) และไม่พึงพอใจ จำนวน 75 คน (8.21%)

ทัศนคติของนักเรียนที่มีความพึงพอใจกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ พบว่า สะดวกและง่ายต่อการใช้งาน เป็นทัศนคติสำคัญต่อความพึงพอใจมากที่สุด จำนวน 608 คน (74.42%) รองลงมาได้แก่ ทราบผลคะแนนทันทีหลังจากทำการทดสอบ จำนวน 94 คน (11.51%) เป็นนวัตกรรมทันสมัยและรู้สึกตื่นเต้นที่ได้ทำการทดสอบ จำนวน 55 คน (6.73%) ยากต่อการทุจริตในการทดสอบ จำนวน 29 คน (3.55%) มีความแม่นยำของการตรวจให้คะแนน จำนวน 18 คน (2.20%) เป็นการประหยัดทรัพยากรกระดาษ จำนวน 8 คน (0.98%) และรูปแบบการนำเสนอสำหรับการทดสอบมีความน่าสนใจ จำนวน 5 คน (0.61%) แสดงได้ดังตารางที่ 48 และ แผนภาพที่ 29

ตารางที่ 48 แสดงความพึงพอใจของนักเรียนกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำแนกตามทัศนคติของนักเรียน

ความพึงพอใจ	จำนวน	ร้อยละ
สะดวกและง่ายต่อการใช้งาน	608	74.42
ทราบผลคะแนนทันทีหลังจากทำการทดสอบ	94	11.51
เป็นนวัตกรรมทันสมัย	55	6.73
ยากต่อการทุจริตในการทดสอบ	29	3.55
มีความแม่นยำของการตรวจให้คะแนน	18	2.20
เป็นการประหยัดทรัพยากรกระดาษ	8	0.98
รูปแบบการนำเสนอสำหรับการทดสอบมีความน่าสนใจ	5	0.61
รวม	817	100

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



แผนภาพที่ 32 ความพึงพอใจของนักเรียนกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์จำแนกตามทัศนคติของนักเรียน

ทัศนคติของนักเรียนที่มีความไม่พึงพอใจกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ พบว่า ไม่สามารถแก้ไขคำตอบที่ได้เลือกไปแล้วได้ เป็นทัศนคติสำคัญต่อความไม่พึงพอใจมากที่สุด จำนวน 65 คน (86.67%) รองลงมาได้แก่ กังวลเรื่องสายตาจากการสังเกตหน้าจอคอมพิวเตอร์เป็นระยะเวลานาน จำนวน 3 คน (4.00%) รูปแบบการนำเสนอไม่น่าสนใจ จำนวน 1 คน (1.33%) มองเป็นเรื่องการสร้างอุปนิสัยความเคยชิน รักความสะดวก จำนวน 1 คน (1.33%) และไม่ระบุ จำนวน 5 คน (6.67%) แสดงได้ดังตารางที่ 49 และ แผนภาพที่ 30

ตารางที่ 49 แสดงความไม่พึงพอใจของนักเรียนกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำแนกตามทัศนคติของนักเรียน

ความไม่พึงพอใจ	จำนวน	ร้อยละ
ไม่สามารถแก้ไขคำตอบที่ได้เลือกไปแล้วได้	65	86.67
กังวลเรื่องสายตาจากการสังเกตหน้าจอคอมพิวเตอร์เป็นระยะเวลานาน	3	4.00
รูปแบบการนำเสนอไม่น่าสนใจ	1	1.33
มองเป็นเรื่องการสร้างอุปนิสัยความเคยชิน รักความสะดวก	1	1.33
ไม่ระบุ	5	6.67
รวม	75	100

แผนภาพที่ 33 ความไม่พึงพอใจของนักเรียนกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำแนกตามทัศนคติของนักเรียน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ผลการพิจารณาคุณภาพการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ พบว่า นักเรียนส่วนใหญ่มีข้อเสนอแนะสำหรับการปรับปรุงเพิ่มเติม เรื่องการขาดความเสมอภาคในการทดสอบจากการที่จำนวนข้อที่ได้รับการทดสอบของนักเรียนแต่ละคนไม่เท่าเทียมกัน อันเป็นผลจากการทดสอบนักเรียนจำแนกตามระดับความสามารถ ส่งผลให้นักเรียนรู้สึกถึงความแตกต่างมากที่สุด จำนวน 29 คน (20.86%) รองลงมาได้แก่ ปรับความยากง่ายของข้อสอบให้ง่ายขึ้น จำนวน 22 คน (15.83%) ควรให้เวลาสร้างความคุ้นเคยในรายละเอียดและทดลองปฏิบัติก่อนใช้จริงเพื่อสร้างความมั่นใจ จำนวน 19 คน (13.67%) ควรมีการเฉลยคำตอบและวิธีการคำนวณ หลังทราบผลคะแนน จำนวน 10 คน (7.19%) รูปแบบการนำเสนอของโปรแกรมควรมีสีสันดึงดูดความสนใจ จำนวน 10 คน (7.19%) ประสิทธิภาพเครื่องคอมพิวเตอร์ของทางโรงเรียนไม่เอื้ออำนวย จำนวน 8 คน (5.76%) เพิ่มระยะเวลาในการทดสอบให้มากขึ้น จำนวน 7 คน (5.04%) ไม่สามารถข้ามข้อสอบที่ทำไม่ได้ไปทำข้อสอบถัดไปได้ ทำให้เสียเวลา จำนวน 6 คน (4.32%) ควรกำหนดระยะเวลาในการเลือกตอบในแต่ละข้อสอบ จำนวน 5 คน (3.60%) ผลการประเมินควรเพิ่มการแสดงผลอาทิ ดีมาก, ดี, ปานกลาง, พอใช้, ควรปรับปรุง เป็นต้น จำนวน 5 คน (3.60%) ควรมีข้อคำถามยืนยันการเลือกคำตอบก่อนเปลี่ยนไปทำข้อถัดไป จำนวน 5 คน (3.60%) มีโปรแกรมเครื่องคิดเลขเพิ่มเติมในโปรแกรม จำนวน 4 คน (2.88%) ควรมีดนตรีประกอบระหว่างการทดสอบ จำนวน 3 คน (2.16%) ข้อสอบมีค่าความยากง่ายไม่สมดุล จำนวน 3 คน (2.16%) ควรมีฉากกั้นระหว่างเครื่องคอมพิวเตอร์เพื่อป้องกันการทุจริต จำนวน 1 คน (0.72%) เป็นกังวลที่จะต้องทำข้อสอบใหม่หากเกิดไฟฟ้าขัดข้อง จำนวน 1 คน (0.72%) และหากนักเรียนลงทะเบียน ชื่อ-นามสกุล ผิด จะเปิดโอกาสให้ทำการทดสอบซ้ำได้ ซึ่งจะส่งผลต่อความเหลื่อมล้ำในการทดสอบ จำนวน 1 คน (0.72%) แสดงได้ดังตารางที่ 50 และแผนภาพที่ 31

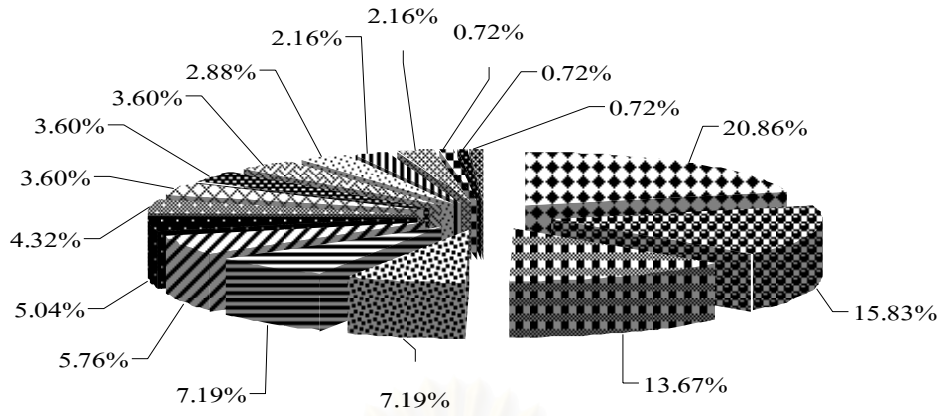
ตารางที่ 50 แสดงผลการพิจารณาคุณภาพการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์จำแนกตามข้อเสนอแนะสำหรับการปรับปรุงเพิ่มเติมของนักเรียน

ข้อเสนอแนะ	จำนวน	ร้อยละ
ขาดความเสมอภาคในการทดสอบ จากการที่จำนวนข้อที่ได้รับการทดสอบของนักเรียนแต่ละคนไม่เท่าเทียมกัน	29	20.86
ปรับความยากง่ายของข้อสอบให้ง่ายขึ้น	22	15.83
ควรให้เวลาสร้างความคุ้นเคยในรายละเอียดและทดลองปฏิบัติก่อนใช้จริงเพื่อสร้างความมั่นใจ	19	13.67

ตารางที่ 50 (ต่อ)

ข้อเสนอแนะ	จำนวน	ร้อยละ
ควรมีการเฉลยคำตอบและวิธีการคำนวณ หลังทราบผลคะแนน	10	7.19
รูปแบบการนำเสนอของโปรแกรมควรมีสีสันดึงดูดความสนใจ	10	7.19
ประสิทธิภาพเครื่องคอมพิวเตอร์ของทางโรงเรียนไม่เอื้ออำนวย	8	5.76
เพิ่มระยะเวลาในการทดสอบให้มากขึ้น	7	5.04
ไม่สามารถข้ามข้อสอบที่ทำไม่ได้ไปทำข้อสอบถัดไปได้ ทำให้เสียเวลา	6	4.32
ควรกำหนดระยะเวลาในการเลือกตอบในแต่ละข้อสอบ	5	3.60
ผลการประเมินควรเพิ่มการแสดงผลอาทิ ดีมาก, ดี, ปานกลาง, พอใช้, ควรปรับปรุง	5	3.60
ควรมีข้อคำถามยืนยันการเลือกคำตอบก่อนเปลี่ยนไปทำข้อถัดไป	5	3.60
มีโปรแกรมเครื่องคิดเลขเพิ่มเติมในโปรแกรม	4	2.88
ควรมีดนตรีประกอบระหว่างการทดสอบ	3	2.16
ข้อสอบมีค่าความยากง่ายไม่สมดุล	3	2.16
ควรมีฉากกั้นระหว่างเครื่องคอมพิวเตอร์เพื่อป้องกันการทุจริต	1	0.72
เป็นกังวลที่จะต้องทำข้อสอบใหม่หากเกิดไฟฟ้าขัดข้อง	1	0.72
หากนักเรียนลงทะเบียน ชื่อ-นามสกุล ผิด จะเปิดโอกาสให้ทำการทดสอบซ้ำได้ ซึ่งจะส่งผลต่อความเหลื่อมล้ำในการทดสอบ	1	0.72
รวม	139	100

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



ขาดความเสมอภาคในการทดสอบ

ปรับความยากง่ายของข้อสอบให้ง่ายขึ้น

ให้เวลาสร้างความคุ้นเคยในรายละเอียดและทดลองปฏิบัติก่อนใช้จริงเพื่อสร้างความมั่นใจ

ควรมีการเฉลยคำตอบและวิธีการคำนวณ หลังทราบผลคะแนน

รูปแบบการนำเสนอของโปรแกรมควรมีสีสันดึงดูดความสนใจ

ประสิทธิภาพเครื่องคอมพิวเตอร์ของทางโรงเรียนไม่เอื้ออำนวย

เพิ่มระยะเวลาในการทดสอบให้มากขึ้น

ไม่สามารถข้ามข้อสอบที่ทำไม่ได้ไปทำข้อสอบถัดไปได้ ทำให้เสียเวลา

ควรกำหนดระยะเวลาในการเลือกตอบในแต่ละข้อสอบ

ผลการประเมินควรเพิ่มการแสดงผลอาทิ ดีมาก, ดี, ปานกลาง, พอใช้, ควรปรับปรุง เป็นต้น

ควรมีข้อคำถามยืนยันการเลือกคำตอบก่อนเปลี่ยนไปทำข้อถัดไป

น่าจะมีโปรแกรมเครื่องคิดเลขเพิ่มเติมในโปรแกรม

ควรมีดนตรีประกอบระหว่างการทดสอบ

ข้อสอบมีค่าความยากง่ายไม่สมดุล

ควรมีฉากกั้นระหว่างเครื่องคอมพิวเตอร์เพื่อป้องกันการทุจริต

เป็นกังวลที่จะต้องทำข้อสอบใหม่หากเกิดไฟฟ้าขัดข้อง

หากลงทะเบียน ชื่อ-นามสกุล ผิดจะมีโอกาสได้ทำการทดสอบซ้ำ

แผนภาพที่ 34 ผลการพิจารณาคุณภาพการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์จำแนกตามข้อเสนอแนะสำหรับการปรับปรุงเพิ่มเติมของนักเรียน

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวสิริลักษณ์ เกษรปทุมานันท์ สำเร็จการศึกษาวิทยาศาสตรบัณฑิต สาขาวิชาคณิตศาสตร์ จากคณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ เมื่อปีการศึกษา 2545 เข้าศึกษาต่อหลักสูตรครุศาสตรมหาบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อปีการศึกษา 2547

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย