

การพัฒนาแบบทดสอบมาตรฐานทางคณิตศาสตร์ สำหรับนักเรียน
ระดับมัธยมศึกษาตอนต้น พ.ศ.2548

A Development of Mathematics Standardized Test for Lower Secondary Students

อ.วัฒนา นานาแสงวานิช

โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม
สาขาวิชาการศึกษาคณิตศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร

อ.วิมลมาศ อัมพพวงษ์

โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม
สาขาการวัดและประเมินผล จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร

บทคัดย่อ

รายงานการวิจัยครั้งนี้ มีวัตถุประสงค์เพื่อพัฒนาแบบทดสอบมาตรฐานทางคณิตศาสตร์ สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น เพื่อตรวจสอบคุณภาพของแบบทดสอบมาตรฐานทางคณิตศาสตร์ ในด้านความเที่ยง (Reliability) และความตรง (Validity) และเพื่อจัดทำเกณฑ์มาตรฐาน รายวิชาคณิตศาสตร์ สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น

กลุ่มตัวอย่างประชากรที่ใช้ในการวิจัยครั้งนี้คือ นักเรียนระดับมัธยมศึกษาตอนต้น ปีการศึกษา 2544 ในกรุงเทพมหานคร จำนวน 2,741 คน แบ่งเป็นกลุ่มตัวอย่างสำหรับการทดลองใช้เครื่องมือครั้งที่ 1 จำนวน 589 คน ทดลองใช้เครื่องมือครั้งที่ 2 จำนวน 1,476 คน และสำหรับสร้างเกณฑ์ปกติ จำนวน 676 คน การเลือกกลุ่มตัวอย่างใช้วิธีสุ่มอย่างง่าย (Simple Random Sampling)

เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบสัมฤทธิ์ผลวิชาคณิตศาสตร์ตามหลักสูตร มีรายละเอียดของเนื้อหายึดตามหลักสูตรระดับมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) ของกระทรวงศึกษาธิการ

การวิเคราะห์ข้อมูลใช้การหาค่าความเที่ยง (Reliability) ความตรง (Validity) ความยากง่าย (Difficulty) และอำนาจจำแนก (Discrimination) ของแบบทดสอบรายข้อ โดยใช้ทฤษฎีการทดสอบแบบดั้งเดิม (CTT) ได้แบบทดสอบมาตรฐานที่มีความยากง่ายอยู่ระหว่าง 0.2 - 0.8 และอำนาจจำแนกมากกว่า 0.2 จำนวน 3 ฉบับ ฉบับละ 60 ข้อ แบบทดสอบแต่ละฉบับมีค่าความเที่ยงของแบบทดสอบ (KR-20) เท่ากับ 0.88 , 0.83 และ 0.88 ตามลำดับ และหาเกณฑ์ปกติ (Norms) ที่เป็นคะแนนปกติ (T-score)

คำสำคัญ แบบทดสอบมาตรฐาน / เกณฑ์ปกติ (Norms) / คะแนนปกติ (T-score)

ABSTRACT

The purposes of this study were to develop Mathematics Standardized Tests for lower secondary students and to develop norms for Mathematics subjects for lower secondary students. The test was constructed based on the 1990 Revised Edition of the 1978 Lower Secondary School Curriculum by the Ministry of Education.

Three paralleled tests were developed and tested with 2,741 lower secondary students from schools in Bangkok. 589 students participated in the first pilot study, 1,476 students participated in the second pilot study and 676 students participated in the final stage of developing the norms.

The three tests were analyzed using classical test theory to examine the reliability, validity, difficulty and the power of discrimination of the tests item. The results showed that the three tests had the difficulty level between 0.2 - 0.8 and the power of discrimination of the three tests was over 0.2. The reliability of the three tests obtained from KR - 20 were 0.88, 0.83 and 0.88.

Keywords: Standardized Test/ Norms/ T-score

บทนำ

การศึกษาเป็นกระบวนการสำคัญในการพัฒนาศักยภาพของมนุษย์ ถ้าประชากรของประเทศใดมีความรู้ และความเชี่ยวชาญทางด้านวิทยาศาสตร์และเทคโนโลยี ก็ย่อมมีศักยภาพในการพัฒนาประเทศได้สูงกว่าประเทศอื่นๆ และวิชาที่เป็นพื้นฐานสำคัญของวิทยาศาสตร์และเทคโนโลยีก็คือวิชาคณิตศาสตร์ ถ้าประชากรของประเทศใดปราศจากความรู้ ความสามารถทางคณิตศาสตร์ การที่ประเทศนั้นจะพัฒนาความก้าวหน้า ความเป็นผู้นำที่ทัดเทียมประเทศอื่นๆ คงเป็นไปได้ยาก ดังที่ภิญโญ สาร (2526:53) กล่าวว่า "ประเทศใดก็ตามที่มี "คน" ที่มีคุณภาพเป็นพลเมืองประเทศนั้น จะทำการใดก็สำเร็จสมความปรารถนาทั้งสิ้น ในทางตรงกันข้ามประเทศใดที่มี "คน" ที่ขาดคุณภาพเป็นพลเมืองประเทศนั้น จะทำการใดไม่ว่าจะลงทุนด้วยจำนวนมากสักส่วนเพียงใดในการนั้น เมื่อคนปฏิบัติขาดคุณภาพผลงาน ก็ย่อมไม่มีคุณภาพไปด้วย บางที่อาจล้มเหลว ขาดทุนล่มจมไปเลยก็ได้" การให้การศึกษาแก่บุคคลนั้น จำเป็นต้องปลูกฝังตั้งแต่เด็กเพื่อให้เป็นพื้นฐานของการพัฒนากำลังคน

ในอนาคต เด็กและเยาวชนเป็นทรัพยากรที่มีค่ายิ่งของชาติบ้านเมืองและเป็นกำลังสำคัญในการพัฒนาประเทศ ดังนั้นการส่งเสริมและพัฒนาเด็กจึงต้องอาศัยการศึกษาในระดับประถมศึกษาจนถึงมัธยมศึกษา เพราะเป็นการศึกษาขั้นพื้นฐานของพลเมืองที่รัฐจำเป็นต้องจัดให้ทุกคนได้มีโอกาสเรียนโดยเสมอภาคกัน

จากการที่สมาคมนานาชาติว่าด้วยการประเมินผลสัมฤทธิ์ทางการเรียนหรือ IEA (International Association for the Evaluation of Education Achievement) ได้ประเมินผลการเรียนวิชาคณิตศาสตร์และวิทยาศาสตร์ของเด็กนักเรียนประเทศต่างๆ พบว่าเด็กนักเรียนระดับประถมศึกษาปีที่ 3 และ 4 ของประเทศไทย ได้คะแนนวิชาคณิตศาสตร์เป็นอันดับที่ 17 จากจำนวนทั้งหมด 24 ประเทศ และอันดับที่ 22 จากจำนวนทั้งหมด 26 ประเทศตามลำดับ ซึ่งเป็นอันดับที่ค่อนข้างต่ำ และคะแนนที่ได้ต่ำกว่าคะแนนเฉลี่ยระดับนานาชาติ ส่วนนักเรียนระดับมัธยมศึกษาปีที่ 1 และ 2 ของประเทศไทย ได้คะแนนวิชาคณิตศาสตร์เป็นอันดับ

ที่ 18 จากจำนวนทั้งหมด 39 ประเทศ เป็นอันดับที่ 20 จากจำนวนทั้งหมด 41 ประเทศ ซึ่งอยู่อันดับกลางๆ และคะแนนที่ได้สูงกว่าคะแนนเฉลี่ยระดับนานาชาติเพียงเล็กน้อย นอกจากนี้กรมวิชาการกระทรวงศึกษาธิการ ได้ใช้เครื่องมือมาตรฐานกลางประเมินนักเรียนชั้นประถมศึกษาปีที่ 6 ทั่วประเทศพบว่า นักเรียนทำคะแนนวิชาคณิตศาสตร์ได้เพียง 52.76% ซึ่งถือว่าต่ำกว่ามาตรฐานสากลเป็นอย่างมาก

การทดสอบเป็นการให้ข้อมูลเกี่ยวกับผลการเรียนของนักเรียน ผลการสอนของครู ผลการวางแผนการศึกษาของนักการศึกษาและผลการบริหารจัดการของนักบริหารในสถาบันการศึกษา ในวงการศึกษานั้นการสร้างระบบการตรวจสอบที่ครบวงจร และที่เชื่อถือได้จะทำให้ผู้บริหารการศึกษามีข้อมูลพอที่จะทำการตัดสินใจเกี่ยวกับการบริหารสถาบันการศึกษาได้ในด้านการเรียนการสอน การสร้างระบบการตรวจสอบที่ครบวงจรในห้องเรียนจะทำให้ครูและนักเรียน มีข้อมูลพอที่จะตัดสินใจเกี่ยวกับการเลือกวิชาเรียน การเพิ่ม-ถอนวิชา ตลอดจนการวางแผนการเรียนของนักเรียนและการวางแผนการสอนของครู การตรวจสอบที่ดี จะทำให้เกิดการเพิ่มคุณภาพและประสิทธิภาพของการเรียนการสอนในทำนองเดียวกัน การทดสอบที่ไม่ดีก็ทำให้เกิดผลในทางลบกับผู้เรียน ผู้สอน และผู้เกี่ยวข้องเกิดความเบื่อหน่าย ท้อถอย เสียความเชื่อถือในตัวเองหมดศรัทธา หรือเกิดการเรียนแบบท่องจำแทนการคิดวิเคราะห์ เกิดการเดาแทนการพยายามแสวงหาความรู้ เกิดการทุจริตในการสอบ เป็นต้น ดังที่อุทุมพร จามรมาน (2535) กล่าวว่า การทดสอบมีอิทธิพลต่อการเรียนการสอนและการศึกษา ต่อ นักเรียน ครู ผู้บริหาร ประชาชน ต่อสถาบันการศึกษา ต่อแผนพัฒนาการศึกษาและต่อประเทศในที่สุดด้วย เพราะการศึกษาคือการพัฒนาทรัพยากรมนุษย์ของประเทศ

ในวิชาการวัดผลและประเมินผลการศึกษาตามแนวทางสมัยใหม่ก็ได้ให้ความรู้ที่สำคัญแก่เราในเรื่องการสร้างข้อทดสอบที่ดี ข้อทดสอบที่มีความน่าเชื่อถือได้ มีความแม่นยำเที่ยงตรงสูง มีอำนาจจำแนกความเก่ง ความไม่เก่งของผู้ตอบแบบทดสอบ มีอำนาจจำแนกความถนัด ความไม่ถนัด ความรู้ ความไม่รู้ ด้วยการอาศัยเทคนิคการออกข้อทดสอบ และแบบแผนการวัดผลประเมินผลแบบต่างๆ ได้ และข้อทดสอบเหล่านั้นสามารถนำมาใช้ซ้ำแล้วซ้ำอีกได้ โดยยังคงประสิทธิภาพของแบบทดสอบที่ดีอยู่เสมอ จากแนวความคิดนี้แสดงให้เห็นว่าแบบทดสอบที่เคยสร้างไว้ครั้งหนึ่งนั้น เป็นสิ่งที่อาจนำมาใช้ให้เกิดประโยชน์ในการทดสอบในโอกาสต่อไปได้ ถ้ามีการเก็บรักษาดีและถ้าสามารถพัฒนาแบบทดสอบตามระบบ และเทคนิควิธีการวัดผลประเมินผลจนก้าวหน้าแล้ว แบบทดสอบเหล่านั้นจะเป็นแบบทดสอบที่มีประสิทธิภาพสูง อำนาจความสะดวกและอำนาจประโยชน์ต่อระบบการวัดผลการเรียนการสอน และการประเมินผลการศึกษาของโรงเรียนได้เป็นอย่างมาก (อุทัย บุญประเสริฐ, 2535 : 6)

การพัฒนาการเรียนรู้อาจเกิดขึ้นกับผู้เรียนอย่างบรรลุเป้าหมายที่กำหนดไว้ ต้องอาศัยองค์ประกอบ 3 ประการ ที่มีการประสานสัมพันธ์กันได้แก่ การกำหนดจุดมุ่งหมายของการเรียนรู้อาจจัดกระบวนการเรียนรู้ และการวัดประเมินผล การเรียนรู้ ในการวัดประเมินผลการเรียนรู้นั้น นักวัดผลให้ความสำคัญอย่างมากต่อการวัดความสามารถ (Ability) ที่แท้จริงของผู้เรียน ซึ่งเป็นคุณลักษณะภายในที่ไม่สามารถสังเกตหรือวัดได้โดยตรงเหมือนการวัดทางกายภาพ จึงต้องวัดคุณลักษณะดังกล่าวทางอ้อม โดยวิธีการตอบสนองต่อสิ่งเร้าที่จัดขึ้น แล้วนำผลที่ได้สรุปอ้างอิงไปยังคุณลักษณะภายในที่ต้องการวัด แต่เนื่องจากการวัดคุณลักษณะภายในโดยทางอ้อมนี้มีความคลาดเคลื่อนในการวัดเกิดขึ้นเสมอ นักวัดผลจึงได้เสนอทฤษฎี

การวัดเพื่อให้การสรุปอ้างอิงเกี่ยวกับคุณลักษณะภายในหรือความสามารถของผู้เรียนมีความคลาดเคลื่อนน้อยที่สุดเท่าที่จะเป็นไปได้ ทฤษฎีดังกล่าวสามารถแบ่งเป็น 2 ประเภทใหญ่ๆ ได้แก่ ทฤษฎีแบบคลาสสิกและทฤษฎีสอบแนวใหม่แบบตอบสนองข้อสอบ (Item Response Theory หรือ IRT) ซึ่งเป็นที่ยอมรับโดยทั่วไปในหมู่นักวิชาการและเป็นที่นิยมกันอย่างแพร่หลายในปัจจุบัน (ศิริชัย กาญจนวาสี, 2535 : 46)

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาแบบทดสอบมาตรฐานทางคณิตศาสตร์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น
2. เพื่อตรวจสอบคุณภาพของแบบทดสอบมาตรฐานทางคณิตศาสตร์ในด้านความเที่ยง (Reliability) และความตรง (Validity)
3. เพื่อจัดทำเกณฑ์มาตรฐานรายวิชาคณิตศาสตร์ สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น

ขอบเขตของการศึกษา

1. การสร้างแบบทดสอบมาตรฐานสร้างโดยยึดหลักสูตรคณิตศาสตร์ของกระทรวงศึกษาธิการ ระดับมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533)
2. กลุ่มตัวอย่างที่ใช้ในการศึกษา ครั้งนี้เป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 15 โรงเรียน

เป้าหมายของการวิจัย

ได้แบบทดสอบมาตรฐานทางคณิตศาสตร์ สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น (ม.1-ม.3) เป็นแบบทดสอบคู่ขนานจำนวน 3 ฉบับ ฉบับละ 60 ข้อ ใช้เวลาสอบฉบับละ 3 ชั่วโมง

วิธีดำเนินการวิจัย

การพัฒนาแบบทดสอบมาตรฐานดังกล่าว ได้ดำเนินการตั้งแต่ ปี 2541 โดยมีวิธีดำเนินการวิจัยดังนี้ (สุพัฒน์ สุขมลสันต์, 2530 : 67)

1. วิเคราะห์หลักสูตรคณิตศาสตร์ของกระทรวงศึกษาธิการ ระดับมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ.2533) รวมทั้งสมรรถภาพที่นักเรียนควรจะต้องรู้ตามมาตรฐานสากล

2. สร้างผังแบบทดสอบ (Table of Test Specification) โดยมีส่วนประกอบของโครงสร้างเนื้อหา ระดับพฤติกรรมที่มุ่งวัด ได้แก่ ความรู้ความจำ ความเข้าใจ การนำไปใช้และการวิเคราะห์

3. กำหนดลักษณะเฉพาะของแบบทดสอบ (Item Specification) สำหรับแบบทดสอบระดับมัธยมศึกษาตอนต้น มีจำนวน 60 ข้อ

4. ร่างแบบทดสอบตามลักษณะเฉพาะของแบบทดสอบที่กำหนดเป็นแบบทดสอบคู่ขนาน 3 ฉบับ แต่ละฉบับมีจำนวน 80 ข้อ

5. ผู้เชี่ยวชาญทางคณิตศาสตร์แต่ละระดับ ตรวจสอบความตรงตามเนื้อหา (Content Validity) ของแบบทดสอบแต่ละฉบับและตรวจสอบความตรงของแบบทดสอบรายข้อกับจุดมุ่งหมาย (Item-Objective Congruence)

6. ปรับปรุงแบบทดสอบและนำแบบทดสอบไปทดลองใช้ครั้งที่ 1 กับกลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 589 คน โดยใช้จำนวนแบบทดสอบ 3 ฉบับ จำนวนกลุ่มตัวอย่างของฉบับที่ 1, 2 และ 3 คือ 195, 197 และ 197 คน ตามลำดับ

7. นำผลที่ได้จากการนำแบบทดสอบไปทดลองใช้ครั้งที่ 1 มาวิเคราะห์แบบทดสอบรายข้อ (Item Analysis) เพื่อคำนวณค่าความยากของแบบทดสอบ (Item Difficulty) และอำนาจจำแนกของแบบทดสอบ (Item Discrimination)

8. ปรับปรุงแบบทดสอบข้อที่มีค่าความยาก และอำนาจจำแนกไม่อยู่ในเกณฑ์ที่กำหนด

9. นำแบบทดสอบที่ปรับปรุงแล้วไปทดลองใช้ครั้งที่ 2 กับกลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 1,476 คน โดยใช้จำนวนแบบทดสอบ 3 ฉบับ จำนวนกลุ่มตัวอย่างของฉบับที่ 1, 2 และ 3 คือ 496, 489 และ 491 คน ตามลำดับ

10. นำผลที่ได้จากการนำแบบทดสอบไปทดลองใช้ครั้งที่ 2 มาวิเคราะห์แบบทดสอบ (Item Analysis) เพื่อคำนวณค่าความยากของแบบทดสอบ (Item Difficulty) และอำนาจจำแนกของแบบทดสอบ (Item Discrimination)

11. คัดเลือกแบบทดสอบที่มีความยากพอเหมาะ และอำนาจจำแนกสูง มารวมเป็นแบบทดสอบคู่มือนานจำนวน 3 ฉบับ ฉบับละ 60 ข้อ

12. นำแบบทดสอบคู่มือนานไปใช้กับนักเรียนกลุ่มเป้าหมาย ระดับชั้นมัธยมศึกษาตอนต้น จำนวน 676 คน เพื่อวิเคราะห์คุณภาพของแบบทดสอบ และสร้างเกณฑ์การแปลผลการสอบ

13. วิเคราะห์ความเที่ยง (Reliability) ของแบบทดสอบโดยใช้สูตร KR-20

14. สร้างเกณฑ์การแปลผล โดยสร้างตารางเปอร์เซ็นต์ไทล์ (Percentile Norms) ของระดับชั้นมัธยมศึกษาตอนต้น และสร้างตารางเทียบระดับชั้นมัธยมศึกษาตอนต้น (Grade Norms)

กลุ่มตัวอย่างประชากร

กลุ่มตัวอย่างประชากรที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนระดับชั้นมัธยมศึกษาตอนต้น ปีการศึกษา 2544 ที่เรียนตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) การเลือกกลุ่มตัวอย่างใช้วิธีสุ่มอย่างง่าย (Simple Random Sampling) ได้โรงเรียนที่เป็นกลุ่มตัวอย่างจำนวน 15 โรงเรียน ได้ตัวอย่างประชากรจำนวน 1,476 คน

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบสัมฤทธิ์ผลวิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาตอนต้น ซึ่งผู้วิจัยสร้างขึ้นจากเนื้อหายึดตามหลักสูตรระดับมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ.2533) ของกระทรวงศึกษาธิการ มีลักษณะปรนัยแบบเลือกตอบ 5 ตัวเลือก

การเก็บรวบรวมข้อมูล

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลโดยนำแบบทดสอบไปทดลองใช้กับกลุ่มตัวอย่างจำนวน 3 ครั้ง โดยแต่ละครั้งนำผลที่ได้มาวิเคราะห์แล้วปรับปรุงแบบทดสอบ เพื่อนำไปใช้ครั้งต่อไป ซึ่งกลุ่มตัวอย่างที่ใช้แต่ละครั้งเป็นนักเรียนคนละกลุ่ม

การวิเคราะห์ข้อมูล

ผู้วิจัยนำข้อมูลที่รวบรวมได้ในแต่ละครั้ง มาวิเคราะห์แบบทดสอบรายข้อ (Item Analysis) เพื่อคำนวณค่าความยากแบบทดสอบ (Item Difficulty) และอำนาจจำแนกของแบบทดสอบ (Item Discrimination) รวมทั้งวิเคราะห์หาความเที่ยง (Reliability) ของแบบทดสอบโดยใช้สูตร KR-20 แล้วคัดเลือกข้อสอบที่มีความยากพอเหมาะ และอำนาจจำแนกสูงมารวมเป็นแบบทดสอบคู่มือนานจำนวน 3 ฉบับ ฉบับละ 60 ข้อ เพื่อนำไปใช้กับนักเรียนกลุ่มเป้าหมาย ระดับชั้นมัธยมศึกษาตอนต้นจำนวน 676 คน เพื่อวิเคราะห์คุณภาพของแบบทดสอบ และสร้างเกณฑ์การแปลผลการสอบโดยสร้างตารางเปอร์เซ็นต์ไทล์ (Percentile Norms) ของระดับชั้นมัธยมศึกษาตอนต้น และสร้างตารางเทียบระดับชั้นมัธยมศึกษาตอนต้น (Grade Norms)

ผลการวิจัย

ผลการวิเคราะห์ข้อมูลใช้การหาค่าความเที่ยง (Reliability) ความตรง (Validity) ความยากง่าย (Difficulty) และอำนาจจำแนก (Discrimination) ของแบบทดสอบรายข้อ โดยใช้ทฤษฎีการทดสอบแบบดั้งเดิม (CTT) ได้แบบทดสอบมาตรฐานที่มีค่าความยากง่าย (Difficulty) อยู่ระหว่าง 0.2-0.8 และอำนาจจำแนก (Discrimination) มากกว่า 0.2 จำนวน 3 ฉบับ ฉบับละ 60 ข้อ แบบทดสอบแต่ละฉบับมีค่าความเที่ยงของแบบทดสอบ(KR-20) เท่ากับ 0.88, 0.83 และ 0.88 ตามลำดับ และหาเกณฑ์ปกติ (Norms) ที่เป็นคะแนนปกติ (T - score)

ข้อเสนอแนะ

ข้อเสนอแนะสำหรับผู้คุมสอบ

1. ผู้คุมสอบควรดำเนินการสอบ โดยปฏิบัติตามคู่มือการดำเนินการสอบอย่างเคร่งครัด เพื่อความเป็นมาตรฐาน
2. ก่อนการดำเนินการสอบทุกครั้ง ผู้คุมสอบควรชี้แจงให้นักเรียนทำแบบทดสอบอย่างตั้งใจ เพื่อให้เกิดผลตรงตามความสามารถแท้จริง

ข้อเสนอแนะในการวิจัย

1. ควรมีการพัฒนาแบบทดสอบอย่างต่อเนื่องในทุกรายวิชา เพื่อใช้เป็นเครื่องมือตรวจสอบคุณภาพการศึกษาของนักเรียน
2. ควรนำผลที่ได้จากการสอบด้วยแบบทดสอบมาตรฐานชุดดังกล่าวของนักเรียนไปใช้ประโยชน์ในการปรับปรุงรูปแบบการเรียนการสอน



บรรณานุกรม

- ภิญโญ สาร. (2526). **หลักการบริหารการศึกษา**. กรุงเทพมหานคร : องค์การค้าของคุรุสภา ศึกษาภัณฑ์พาณิชย์.
- วิชาการ , กรม. (2535). **หลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2524** (ฉบับปรับปรุง พ.ศ. 2533) พิมพ์ครั้งที่ 2 . กรุงเทพมหานคร : โรงพิมพ์การศาสนา.
- ศิริชัย กาญจนวาสี และคณิต ไช้มุกต์. (2535). **การพัฒนาโปรแกรมคอมพิวเตอร์สำหรับวิเคราะห์ข้อสอบ และประเมินค่าความสามารถของผู้สอบตามทฤษฎีการตอบสนองด้วยวิธีของเบส์**. กรุงเทพมหานคร : จุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2538). **การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive Testing)**. กรุงเทพมหานคร : ศูนย์ตำราและเอกสารทางวิชาการ, คณะครุศาสตร์, จุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2544). **ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory)**. พิมพ์ครั้งที่ 4 คณะครุศาสตร์, จุฬาลงกรณ์มหาวิทยาลัย. กรุงเทพมหานคร : โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย
- ศิริชัย กาญจนวาสี. (2545). **ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theories)** พิมพ์ครั้งที่ 2 กรุงเทพมหานคร : จุฬาลงกรณ์มหาวิทยาลัย.
- สุพัฒน์ สุกมลสันต์. (2530). **การวิเคราะห์ข้อสอบแนวใหม่ด้วยเครื่องคอมพิวเตอร์**. กรุงเทพมหานคร : สถาบันภาษา, จุฬาลงกรณ์มหาวิทยาลัย.
- สุพัฒน์ สุกมลสันต์. (2535). **คู่มือการใช้โปรแกรม CTIA/Grading (Classical Test Item Analysis and Grading)**. กรุงเทพมหานคร : สถาบันภาษา, จุฬาลงกรณ์มหาวิทยาลัย.
- สุพัฒน์ สุกมลสันต์. (2542). **วิธีวิทยาการวิจัย**. สำนักทดสอบทางการศึกษาและจิตวิทยา, มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- อนันต์ ศรีโสภณ. (2515). **การพัฒนาการทดสอบ : การสร้างข้อสอบมาตรฐาน**. กรุงเทพมหานคร : จุฬารัตนการพิมพ์.
- อนันต์ ศรีโสภณ. (2524). **การวัดและการประเมินผลการศึกษา**. กรุงเทพมหานคร : บริษัทโรงพิมพ์ไทยวัฒนาพานิช จำกัด
- อุทัย บุญประเสริฐ. (2535). **การจัดทำคลังข้อสอบ**. กรุงเทพมหานคร : ศรีมงคลการพิมพ์.
- Cronbach, L.J. and Others. (1972). **The Dependability of Behavioral Measurements**. New York : John Wiley & Sons, Inc.
- Ebel , R.L.(1965). **Measuring Educational Achievement**. Englewood Cliffs , N.J. : Prentice - Hall, Inc.
- Ebel , R.L. (1972). **Essentials of Educational Measurement**. New Jersey : Prentice-Hall, Inc.
- Gronlund,N.E. (1976). **Measurement and Evaluation in Teaching**. New York : MacMillan.
- Gulliksen, H. (1950). **Theory of Mental Tests**. New York : John Wiley.
- Hambleton, R.K.et.al. (1978). Development in latent trait theory : models technical issues and applications. **Review of Educational Research** pp.467-510 .
- Mehrens, William A. and Lehmann , Irvin J. **Measurement and Evaluation in Education Psychology**. New York : Holt, Rinehart and Winston pp. 398-459.
- Stanley, J.C. (1971). "Reliability." **In Educational Measurement**, pp.356 - 441. Edited by R.L. Thorndike. Washington. D.C. : American Council on Education.