

การพัฒนาแบบทดสอบมาตรฐานทางคณิตศาสตร์ สำหรับนักเรียน

ระดับมัธยมศึกษาตอนปลาย พ.ศ.2548

A Development of Mathematics Standardized Tests for Upper Secondary Students

ผศ.ชัยศักดิ์ ชั่งใจ

โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร

อ.สุวรรณา ทิมสกลิต

โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร

บทคัดย่อ

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

กลุ่มตัวอย่างประชากรที่ใช้ในการวิจัยครั้งนี้คือ นักเรียนระดับมัธยมศึกษาตอนปลาย ปีการศึกษา 2544 ในกรุงเทพมหานคร จำนวน 2855 คน แบ่งเป็นกลุ่มตัวอย่างสำหรับการทดลองใช้เครื่องมือ ครั้งที่ 1 จำนวน 580 คน ทดลองใช้เครื่องมือครั้งที่ 2 จำนวน 1585 คน และสำหรับสร้างเกณฑ์ปกติ จำนวน 690 คน การเลือกกลุ่มตัวอย่างใช้วิธีสุ่มอย่างง่าย (Simple Random Sampling)

เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบสัมฤทธิ์ผลวิชาคณิตศาสตร์ตามหลักสูตร มีรายละเอียดของเนื้อหายึดตามหลักสูตรระดับมัธยมศึกษาตอนปลาย พุทธศักราช 2524 (ฉบับปรับปรุง พ.ศ.2533) ของกระทรวงศึกษาธิการ

การวิเคราะห์ข้อมูลใช้การหาค่าความเที่ยง (Reliability) ความตรง (Validity) ความยากง่าย (Difficulty) และอำนาจจำแนก (Discrimination) ของแบบทดสอบรายข้อโดยใช้ทฤษฎีการทดสอบแบบดั้งเดิม (CTT) ได้แบบทดสอบมาตรฐานที่มีความยากง่าย (Difficulty) อยู่ระหว่าง 0.2 -0.8 และอำนาจจำแนก (Discrimination) มากกว่า 0.2 จำนวน 3 ฉบับ ฉบับละ 80 ข้อ แบบทดสอบ แต่ละฉบับมีค่าความเที่ยงของแบบทดสอบ (KR-20) เท่ากับ 0.84 0.83 และ 0.82 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 upper secondary students and to develop norms for Mathematics subjects for upper secondary students. The test was constructed based on the 1990 Revised Edition of the 1981 Upper Secondary School Curriculum by the Ministry of Education.

Three paralleled tests were developed and tested with 2,855 upper secondary students from schools in Bangkok. 580 students participated in the first pilot study, 1,585 students participated in the second pilot study and 690 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 0.84 , 0.83 and 0.82.

Keywords : Standardized Test/ Norms/ T-score

บทนำ

การศึกษาเป็นกระบวนการสำคัญในการพัฒนาศักยภาพของมนุษย์ ถ้าประชากรของประเทศใดมีความรู้และความเชี่ยวชาญทางด้านวิทยาศาสตร์และเทคโนโลยี ก็ย่อมมีศักยภาพในการพัฒนาประเทศได้สูงกว่าประเทศอื่นๆ และวิชาที่เป็นพื้นฐานสำคัญของวิทยาศาสตร์และเทคโนโลยีก็คือ วิชาคณิตศาสตร์ ถ้าประชากรของประเทศใดปราศจากความรู้ความสามารถทางคณิตศาสตร์ การที่ประเทศนั้นจะพัฒนาความก้าวหน้า ความเป็นผู้นำที่ทัดเทียมประเทศอื่นๆ คงเป็นไปได้ยากดังที่ ภิญญู สาร (2526:53) กล่าวว่า " ประเทศใดก็ตามที่มี "คน" ที่มีคุณภาพ เป็นพลเมืองประเทศนั้นจะทำการใดก็สำเร็จ สมความปรารถนาทั้งสิ้นในทางตรงกันข้ามประเทศใดที่มี "คน" ที่ขาดคุณภาพเป็นพลเมืองประเทศนั้น จะทำการใดไม่ว่าจะลงทุนด้วยจำนวนมาก สัดส่วนเพียงใดในการนั้นเมื่อคนปฏิบัติขาดคุณภาพผลงานก็ย่อมไม่มีคุณภาพไปด้วย บางทีอาจล้มเหลว ขาดทุนล้มจมไปเลยก็ได้ " การให้การศึกษาแก่บุคคลนั้น จำเป็น

ต้องปลูกฝังตั้งแต่เด็กเพื่อให้เป็นพื้นฐานของการพัฒนากำลังคนในอนาคตเด็กและเยาวชน เป็นทรัพยากรที่มีค่าของชาติบ้านเมือง และเป็นกำลังสำคัญในการพัฒนาประเทศ ดังนั้นการส่งเสริมและพัฒนาเด็ก จึงต้องอาศัยการศึกษาในระดับประถมศึกษาจนถึงมัธยมศึกษา เพราะเป็นการศึกษาขั้นพื้นฐานของพลเมืองที่รัฐจำเป็นต้องจัดให้ทุกคนได้มีโอกาสเรียนโดยเสมอภาคกัน

จากการที่สมาคมนานาชาติว่าด้วยการประเมินผลสัมฤทธิ์ทางการเรียนหรือ IEA (International Association for the Evaluation of Education Achievement) ได้ประเมินผลการเรียนวิชาคณิตศาสตร์และวิทยาศาสตร์ของเด็กนักเรียนประเทศต่างๆ พบว่าเด็กนักเรียนระดับประถมศึกษาปีที่ 3 และ 4 ของประเทศไทยได้คะแนนวิชาคณิตศาสตร์เป็นอันดับที่ 17 จากจำนวนทั้งหมด 24 ประเทศ และอันดับที่ 22 จากจำนวนทั้งหมด 26 ประเทศตามลำดับ ซึ่งเป็นอันดับที่ค่อนข้างต่ำและคะแนนที่ได้ต่ำกว่าคะแนนเฉลี่ยระดับนานาชาติ

ส่วนนักเรียนระดับมัธยมศึกษาปีที่ 1 และ 2 ของประเทศไทยได้คะแนนวิชาคณิตศาสตร์ เป็นอันดับที่ 18 จากจำนวนทั้งหมด 39 ประเทศ เป็นอันดับที่ 20 จากจำนวนทั้งหมด 41 ประเทศ ซึ่งอยู่อันดับกลางๆ และคะแนนที่ได้สูงกว่าคะแนนเฉลี่ยระดับนานาชาติเพียงเล็กน้อย นอกจากนี้กรมวิชาการกระทรวงศึกษาธิการ ได้ใช้เครื่องมือมาตรฐานกลางประเมินนักเรียนชั้นประถมศึกษาปีที่ 6 ทั่วประเทศพบว่า นักเรียนทำคะแนนวิชาคณิตศาสตร์ได้เพียง 52.76% ซึ่งถือว่าต่ำกว่ามาตรฐานสากลเป็นอย่างมาก

การทดสอบเป็นการให้ข้อมูลเกี่ยวกับผลการเรียนของนักเรียน ผลการสอนของครู ผลการวางแผนการศึกษาของนักการศึกษา และผลการบริหารจัดการของนักบริหารในสถาบันการศึกษา ในวงการศึกษานั้น การสร้างระบบการตรวจสอบที่ครบวงจรและที่เชื่อถือได้ จะทำให้ผู้บริหารการศึกษาที่มีข้อมูลพอที่จะทำการตัดสินใจเกี่ยวกับการบริหารสถาบันการศึกษาได้ ในด้านการเรียนการสอน การสร้างระบบการตรวจสอบที่ครบวงจรในห้องเรียนจะทำให้ครูและนักเรียนมีข้อมูลพอที่จะตัดสินใจเกี่ยวกับการเลือกวิชาเรียน การเพิ่ม-ถอนวิชา ตลอดจนการวางแผนการเรียนของนักเรียนและการวางแผนการสอนของครู การตรวจสอบที่ดีจะทำให้เกิดการเพิ่มคุณภาพและประสิทธิภาพของการเรียนการสอน ในทำนองเดียวกัน การทดสอบที่ไม่ดีก็ทำให้เกิดผลในทางลบกับผู้เรียน ผู้สอนและผู้เกี่ยวข้อง เกิดความเบื่อหน่าย ท้อถอย เสียความเชื่อถือในตัวเองหมดศรัทธาหรือเกิดการเรียนแบบท่องจำแทนการคิดวิเคราะห์ เกิดการเดาแทนการพยายามแสวงหาความรู้ เกิดการทุจริตในการสอบ เป็นต้น ดังที่ อุทุมพร จามรมาน (2535) กล่าวว่า การทดสอบมีอิทธิพลต่อการเรียนการสอนและการศึกษา ต่อนักเรียน ครู ผู้บริหาร ประชาชน ต่อสถาบันการศึกษา ต่อแผนพัฒนาการศึกษา และต่อประเทศในที่สุดด้วย เพราะการศึกษาคือการ

พัฒนาศักยภาพมนุษย์ของประเทศ

ในวิชาการวัดผลและประเมินผลการ ศึกษาตามแนวทางสมัยใหม่ก็ได้ให้ความรู้ที่สำคัญแก่เราในเรื่องการสร้างข้อทดสอบที่ดี ข้อทดสอบที่มีความน่าเชื่อถือได้ มีความแม่นยำเที่ยงตรงสูง มีอำนาจจำแนกความเก่ง ความไม่เก่งของผู้ตอบแบบทดสอบ มีอำนาจจำแนก ความถนัด ความไม่ถนัด ความรู้ความไม่รู้ ด้วยการอาศัยเทคนิคการออกข้อทดสอบและแบบแผนการวัดผลประเมินผลแบบต่างๆได้ และข้อทดสอบเหล่านั้น สามารถนำมาใช้ซ้ำแล้วซ้ำอีกได้ โดยยังคงประสิทธิภาพของแบบทดสอบที่ดีอยู่เสมอ จากแนวความคิดนี้แสดงให้เห็นว่า แบบทดสอบที่เคยสร้างไว้ครั้งหนึ่งนั้น เป็นสิ่งที่อาจนำมาใช้ให้เกิดประโยชน์ในการทดสอบในโอกาสต่อไปได้ ถ้ามีการเก็บรักษาดี และถ้าสามารถพัฒนาแบบทดสอบตามระบบ และเทคนิควิธีการวัดผลประเมินผลจนก้าวหน้าแล้ว แบบทดสอบเหล่านั้น จะเป็นแบบทดสอบที่มีประสิทธิภาพสูง อำนาจความสะดวกและอำนาจประโยชน์ต่อระบบการวัดผลการเรียนการสอน และการประเมินผลการศึกษาของโรงเรียนได้เป็นอย่างมาก (อุทัย บุญประเสริฐ, 2535: 6)

การพัฒนาการเรียนรู้ให้เกิดขึ้นกับผู้เรียนอย่างบรรลุเป้าหมายที่กำหนดไว้ ต้องอาศัยองค์ประกอบ 3 ประการ ที่มีการประสานสัมพันธ์กัน ได้แก่ การกำหนดจุดมุ่งหมายของการเรียนรู้ การจัดกระบวนการเรียนรู้ และการวัดประเมินผลการเรียนรู้ ในการวัดประเมินผลการเรียนรู้นั้น นักวัดผลให้ความสำคัญอย่างมากต่อการวัดความสามารถ (Ability) ที่แท้จริงของผู้เรียน ซึ่งเป็นคุณลักษณะภายในที่ไม่สามารถสังเกต หรือวัดได้โดยตรงเหมือนการวัดทางกายภาพ จึงต้องวัดคุณลักษณะดังกล่าวทางอ้อม โดยวิธีการตอบสนองต่อสิ่งเร้าที่จัดขึ้น แล้วนำผลที่ได้สรุปอ้างอิงไปยังคุณลักษณะภายในที่ต้องการวัด แต่เนื่องจากการวัดคุณลักษณะ

ภายในโดยทางอ้อมนี้ มีความคลาดเคลื่อนในการวัดเกิดขึ้นเสมอ นักวัดผลจึงได้เสนอทฤษฎีการวัด เพื่อให้การสรุปอ้างอิงเกี่ยวกับคุณลักษณะภายใน หรือความสามารถของผู้เรียนมีความคลาดเคลื่อน น้อยที่สุดเท่าที่จะเป็นไปได้ ทฤษฎีดังกล่าวสามารถ แบ่งเป็น 2 ประเภทใหญ่ๆ ได้แก่ ทฤษฎีแบบ คลาสสิกและทฤษฎีสอบแนวใหม่แบบตอบสนอง ข้อสอบ (Item Response Theory หรือ IRT) ซึ่งเป็น ที่ยอมรับโดยทั่วไปในหมู่นักวิชาการและเป็นที่นิยม กันอย่างแพร่หลายในปัจจุบัน (ศิริชัย กาญจนวาสี, 2535:46)

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาแบบทดสอบมาตรฐาน ทางคณิตศาสตร์ สำหรับนักเรียนระดับมัธยมศึกษา ตอนปลาย
2. เพื่อตรวจสอบคุณภาพของแบบ ทดสอบมาตรฐานทางคณิตศาสตร์ในด้านความ เทียง (Reliability) และความตรง (Validity)
3. เพื่อจัดทำเกณฑ์มาตรฐานรายวิชา คณิตศาสตร์ สำหรับนักเรียนระดับมัธยมศึกษา ตอนปลาย

ขอบเขตของการศึกษา

1. การสร้างแบบทดสอบมาตรฐาน สร้างโดยยึดหลักสูตรคณิตศาสตร์ของกระทรวง ศึกษาธิการ ระดับมัธยมศึกษาตอนปลาย พุทธ ศักราช 2524 (ฉบับปรับปรุง พ.ศ.2533)
2. กลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้ เป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4, 5 และ 6 จำนวน 15 โรงเรียน

เป้าหมายของการวิจัย

ได้แบบทดสอบมาตรฐานทางคณิตศาสตร์ สำหรับนักเรียนระดับมัธยมศึกษาตอนปลาย เป็น แบบทดสอบคู่ขนานจำนวน 3 ฉบับ ฉบับละ 80 ข้อ ใช้เวลาสอบฉบับละ 3 ชั่วโมง

วิธีดำเนินการวิจัย

การพัฒนาแบบทดสอบมาตรฐานดังกล่าว ได้ดำเนินการตั้งแต่ปี 2541 โดยมีวิธีดำเนินการวิจัย ดังนี้ (สุพัฒน์ สุขมลสันต์, 2530: 67)

1. วิเคราะห์หลักสูตรคณิตศาสตร์ของ กระทรวงศึกษาธิการ ระดับมัธยมศึกษาตอนปลาย พุทธศักราช 2524 (ฉบับปรับปรุง พ.ศ.2533) รวมทั้งสมรรถภาพที่นักเรียนควรจะต้องตามมาตรฐาน สากล
2. สร้างผังแบบทดสอบ (Table of Test Specification) โดยมีส่วนประกอบของโครงสร้าง เนื้อหา ระดับพฤติกรรมที่มุ่งวัด ได้แก่ ความรู้ ความจำ ความเข้าใจ การนำไปใช้ และการวิเคราะห์
3. กำหนดลักษณะเฉพาะของแบบ ทดสอบ (Item Specification) สำหรับแบบทดสอบ ระดับมัธยมศึกษาตอนปลาย มีจำนวน 80 ข้อ
4. ร่างแบบทดสอบตามลักษณะเฉพาะ ของแบบทดสอบที่กำหนดเป็นแบบทดสอบคู่ขนาน 3 ฉบับ แต่ละฉบับมีจำนวน 98 ข้อ
5. ผู้เชี่ยวชาญทางคณิตศาสตร์แต่ละ ระดับ ตรวจสอบความตรงตามเนื้อหา (Content Validity) ของแบบทดสอบแต่ละฉบับและตรวจสอบ ความตรงของแบบทดสอบรายข้อกับจุดมุ่งหมาย (Item-Objective Congruence)
6. ปรับปรุงแบบทดสอบและนำแบบ ทดสอบไปทดลองใช้ครั้งที่ 1 กับกลุ่มตัวอย่าง นักเรียนชั้นมัธยมศึกษาปีที่ 6 จำนวน 580 คน โดยใช้จำนวนแบบทดสอบ 3 ฉบับ จำนวนกลุ่ม ตัวอย่างของฉบับที่ 1, 2 และ 3 คือ 190, 193 และ 197 คน ตามลำดับ
7. นำผลที่ได้จากการนำแบบทดสอบ ไปทดลองใช้ครั้งที่ 1 มาวิเคราะห์แบบทดสอบ รายข้อ (Item Analysis) เพื่อคำนวณค่าความ ยากของแบบทดสอบ (Item Difficulty) และอำนาจ จำแนกของแบบทดสอบ (Item Discrimination)
8. ปรับปรุงแบบทดสอบข้อที่มีค่า ความยาก และอำนาจจำแนกไม่อยู่ในเกณฑ์ที่ กำหนด

9. นำแบบทดสอบที่ปรับปรุงแล้วไปทดลองใช้ครั้งที่ 2 กับกลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 6 จำนวน 1,585 คน โดยใช้จำนวนแบบทดสอบ 3 ฉบับ จำนวนกลุ่มตัวอย่างของฉบับที่ 1, 2 และ 3 คือ 545, 517 และ 523 คนตามลำดับ

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

11. คัดเลือกแบบทดสอบที่มีความยากพอเหมาะและอำนาจจำแนกสูง มารวมเป็นแบบทดสอบคู่ขนานจำนวน 3 ฉบับ ฉบับละ 80 ข้อ

12. นำแบบทดสอบคู่ขนานไปใช้กับนักเรียนกลุ่มเป้าหมาย ระดับชั้นมัธยมศึกษาตอนปลาย จำนวน 690 คน เพื่อวิเคราะห์คุณภาพของแบบทดสอบ และสร้างเกณฑ์การแปลผลการสอบ

13. วิเคราะห์ความเที่ยง (Reliability) ของแบบทดสอบโดยใช้สูตร KR-20

14. สร้างเกณฑ์การแปลผล โดยสร้างตารางเปอร์เซ็นต์ไทล์ (Percentile Norms) ของระดับชั้นมัธยมศึกษาตอนปลาย และสร้างตารางเทียบระดับชั้นมัธยมศึกษาตอนปลาย (Grade Norms)

กลุ่มตัวอย่างประชากร

กลุ่มตัวอย่างประชากรที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ปีการศึกษา 2544 ที่เรียนตามหลักสูตรมัธยมศึกษาตอนปลาย พุทธศักราช 2524 (ฉบับปรับปรุง พ.ศ.2533) การเลือกกลุ่มตัวอย่างใช้วิธีสุ่มอย่างง่าย (Simple Random Sampling) ได้โรงเรียนที่เป็นกลุ่มตัวอย่างจำนวน 15 โรงเรียน ได้ตัวอย่างประชากร จำนวน 1,585 คน

เครื่องมือที่ใช้การวิจัย

เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบสัมฤทธิ์ผลวิชาคณิตศาสตร์ระดับชั้นมัธยมศึกษาตอนปลาย ซึ่งผู้วิจัยสร้างขึ้นจากเนื้อหายึดตามหลักสูตรระดับมัธยมศึกษาตอนปลาย พุทธศักราช 2524 (ฉบับปรับปรุง พ.ศ.2533) ของกระทรวงศึกษาธิการ มีลักษณะปรนัยแบบเลือกตอบ 4 ตัวเลือก

การเก็บรวบรวมข้อมูล

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูล โดยนำแบบทดสอบไปทดลองใช้กับกลุ่มตัวอย่างจำนวน 3 ครั้ง โดยแต่ละครั้งนำผลที่ได้มาวิเคราะห์ แล้วปรับปรุงแบบทดสอบ เพื่อนำไปใช้ครั้งต่อไป ซึ่งกลุ่มตัวอย่างที่ใช้แต่ละครั้งเป็นนักเรียนคนละกลุ่ม

การวิเคราะห์ข้อมูล

ผู้วิจัยนำข้อมูลที่รวบรวมได้ในแต่ละครั้ง มาวิเคราะห์แบบทดสอบรายข้อ (Item Analysis) เพื่อคำนวณค่าความยากแบบทดสอบ (Item Difficulty) และอำนาจจำแนกของแบบทดสอบ (Item Discrimination) รวมทั้งวิเคราะห์หาความเที่ยง (Reliability) ของแบบทดสอบโดยใช้สูตร KR-20 แล้วคัดเลือกข้อสอบที่มีความยากพอเหมาะ และค่าอำนาจจำแนกสูงมารวมเป็นแบบทดสอบคู่ขนานจำนวน 3 ฉบับ ฉบับละ 80 ข้อ เพื่อนำแบบทดสอบคู่ขนานไปใช้กับนักเรียนกลุ่มเป้าหมายระดับชั้นมัธยมศึกษาตอนปลายจำนวน 300 คน เพื่อวิเคราะห์คุณภาพของแบบทดสอบ และสร้างเกณฑ์การแปลผลการสอบ โดยสร้างตารางเปอร์เซ็นต์ไทล์ (Percentile Norms) ของระดับชั้นมัธยมศึกษาตอนปลายและสร้างตารางเทียบระดับชั้นมัธยมศึกษาตอนปลาย (Grade Norms)

ผลการวิจัย

ผลการวิเคราะห์ข้อมูลใช้การหาค่าความเที่ยง (Reliability) ความตรง (Validity) ความยากง่าย (Difficulty) และอำนาจจำแนก (Discrimination) ของแบบทดสอบรายข้อโดยใช้ทฤษฎีการทดสอบแบบดั้งเดิม (CTT) ได้แบบทดสอบมาตรฐานที่มีค่าความยากง่าย (Difficulty) อยู่ระหว่าง 0.2 - 0.8 และอำนาจจำแนก (Discrimination) มากกว่า 0.2 จำนวน 3 ฉบับ ฉบับละ 80 ข้อ แบบทดสอบแต่ละฉบับมีค่าความเที่ยงของแบบทดสอบ (KR-20) เท่ากับ 0.84, 0.83 และ 0.82 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 : องค์การค้ำของคุรุสภา ศึกษาภัณฑ์พาณิชย์.
- วิชาการ, กรม. (2535). หลักสูตรมัธยมศึกษาตอนปลาย พุทธศักราช 2524 (ฉบับปรับปรุง พ.ศ. 2533) พิมพ์ครั้งที่ 2. กรุงเทพมหานคร : โรงพิมพ์การศาสนา.
- ศิริชัย กาญจนวาสี และคณิต ไชยมุกต์. (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, มหาวิทยาลัย ศรีนครินทรวิโรฒ ประสานมิตร.
- อนันต์ ศรีโสภณ. (2515). การพัฒนาการทดสอบ : การสร้างข้อสอบมาตรฐาน. กรุงเทพมหานคร : จุฬารัตน์การพิมพ์
- อนันต์ ศรีโสภณ. (2524). การวัดและการประเมินผลการศึกษา. กรุงเทพมหานคร : บริษัทโรงพิมพ์ ไทยวัฒนาพานิช 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.