

ฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

ปริญญาโท
ของ
เกียรติศักดิ์ ส่องแสง

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษาคุณวุฒิปบัณฑิต สาขาวิชาการทดสอบและวัดผลการศึกษา
พฤษภาคม 2547

371.261
ก 855พ
ร.2

ฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

บทคัดย่อ
ของ
เกียรติศักดิ์ ส่องแสง



เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษาดุขุภัณฑ์บัณฑิต สาขาวิชาการทดสอบและวัดผลการศึกษา
พฤษภาคม 2547

h 251292

20 ก.ย. 2547

เกียรติศักดิ์ ส่องแสง. (2547), *ฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์*. ปรินซ์นิพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา).
กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. คณะกรรมการควบคุม :
รศ.ดร.ส.วาสนา ประवालพฤษ์, รศ.ดร.วิรัช วรรณรัตน์, ดร.รองอาจ นัยพัฒน์

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ และศึกษาประสิทธิภาพของโปรแกรมการทดสอบ และเพื่อเปรียบเทียบเวลาที่ใช้ในการทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ เมื่อมีการจัดชุดข้อสอบให้ตอบในแต่ละชั้นต่างกัน การกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อต่างกัน และการจัดรูปแบบของข้อสอบที่ไม่มีสื่อประสม และรูปแบบของข้อสอบที่มีสื่อประสมประกอบการทดสอบ

การดำเนินการวิจัยจัดทำเป็น 2 ระยะ ระยะแรกทำการสร้างข้อสอบเพิ่มเติมขึ้น สำหรับข้อสอบที่ผู้วิจัยสร้างเพิ่มเติมขึ้นเป็นข้อสอบวิชาคณิตศาสตร์ วัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 จำนวน 54 ข้อ พัฒนาโปรแกรมการทดสอบ และศึกษาประสิทธิภาพของโปรแกรม ระยะที่สอง เป็นการศึกษาผลของตัวแปรอิสระ 3 ตัว ต่อตัวแปรตาม ได้แก่ เวลาที่ใช้ในการทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ ซึ่งใช้ข้อสอบของรังสรรค์ มณีเล็ก จำนวนข้อสอบ 200 ข้อ และข้อสอบของผู้วิจัยจำนวน 44 ข้อ กลุ่มตัวอย่างได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 1 ปีการศึกษา 2546 จังหวัดปทุมธานี จำนวน 597 คน ใช้สำหรับเก็บข้อมูลเพื่อหาคุณภาพของข้อสอบ และนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2546 จังหวัดปทุมธานี จำนวน 968 คน ใช้สำหรับเก็บรวบรวมข้อมูลเพื่อพัฒนาและศึกษาคุณภาพของโปรแกรมการทดสอบ และใช้สำหรับทำการวิจัย

ผลการวิจัยปรากฏว่า ได้ข้อสอบวิชาคณิตศาสตร์ วัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 ที่มีคุณภาพตามเกณฑ์เพิ่มเติมขึ้นจำนวน 44 ข้อ เมื่อนำไปรวมกับข้อสอบในคลังของรังสรรค์ มณีเล็ก จึงมีข้อสอบในคลังทั้งสิ้น 244 ข้อ และได้โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ที่สามารถทำงานบนระบบปฏิบัติการวินโดวส์ 98 ขึ้นไป โปรแกรมสามารถจัดชุดข้อสอบให้ตอบในแต่ละชั้นได้ 3 แบบ กำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อได้ 3 แบบ และจัดรูปแบบของข้อสอบได้ 2 รูปแบบ โปรแกรมเริ่มต้นการทดสอบด้วยข้อสอบที่มีความยากง่ายปานกลาง ประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์ คัดเลือกข้อสอบข้อต่อไปจากค่าฟังก์ชันสารสนเทศของข้อสอบที่สูงที่สุด ยุติการทดสอบเมื่อความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบเท่ากับหรือต่ำกว่า 0.3 โปรแกรมสามารถทำงานได้อย่างมีประสิทธิภาพ และผู้สอบมีความพึงพอใจในการทดสอบยิ่งขึ้น สำหรับการศึกษาผลของตัวแปรอิสระต่อตัวแปรตาม พบว่าการจัดชุดข้อสอบให้ตอบชั้นละ 3 ข้อ ให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงที่สุด รองลงมาได้แก่ การจัดชุดข้อสอบให้ตอบชั้นละ 2 ข้อ ส่วนการกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อนั้น การกำหนดเวลาให้ตอบข้อสอบเฉลี่ยเท่ากันทุกข้อหรือการกำหนดเวลาให้ตอบข้อสอบในแต่ละข้อตามค่าความยากของข้อสอบ จะทำให้ผู้สอบใช้เวลาในการ

ทดสอบน้อยลง และการจัดชุดของข้อสอบที่มีสื่อประสมประกอบการทดสอบ จะทำให้ผู้สอบใช้เวลาในการทดสอบน้อยลง และผู้สอบมีความพึงพอใจในการทดสอบมากยิ่งขึ้น

TEST INFORMATION FUNCTIONS IN COMPUTERIZED ADAPTIVE TESTING

AN ABSTRACT

BY

KIATTISAK SONGSAENG

Presented in partial fulfillment of the requirements
for the Doctor of Education Degree in Testing and Educational Measurement
at Srinakharinwirot University

May 2004

Kiattisak SongSaeng. (2004). *Test Information Functions In Computerized Adaptive Testing*. Dissertation, Ed.D. (Testing and Educational Measurement). Bangkok : Graduate School, Srinakharinwirot University. Advisor Committee: Assoc.Prof.Dr. Sor Wasna Pravalpruk, Assoc.Prof.Dr.Wirach Wannaratn, Dr. Ong-Art Naiyapatana

The objectives of this research are to develop program for test information in computerized adaptive testing and to study efficiency of testing program. Also, the comparison of time spend, numbers of items, and test information are investigated for 3 different condition of testing: number of items in one step of testing , fixed and unfixed time of each step ; and multimedia and nonmultimedia of test presentation.

This process of research have two steps. First construct 54 test items in fraction mathematics class for student Pratomsuksa 6. Develop testing program and efficiency of the program. Secondly, study different condition of testing ; 1) time for testing 2) number of items and 3) test information function. The samples of study consist of 1) 597 students in Mathayom 1 in the first semester of 2003 academic year province Prathumthani, for item try-out for quality of the items and 2) 968 students Pratomsuksa 6 in second semester of 2003 academic year province Prathumthani for research purpose of developing and efficiency of testing program.

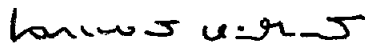
Test construction of 55 items, only 44 items are qualified to include in the item pool of 244 items. The computerized adaptive testing program can be used on window 98/2000/ME. The program provides three different number of item in one step: 1, 2, and 3 items; 1 equal time limit and different time limit for testing ; and multimedia and nonmultimedia of test presentation. The initial item starts at moderate item, then estimates level of testee's ability by bayesian technique, termination at $SE \leq 0.30$. The program works effectively. The major hypotheses, confirm that three item per step gets the highest test information function, and two items per step follower. Equal time used for each item and different time according to item difficulty effect to reduce testing time. Finally presenting multimedia effects on reducing testing time and testee 's satisfaction.

ปริญญาโท
เรื่อง

ฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

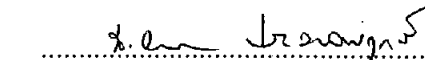
ของ
นายเกียรติศักดิ์ ส่องแสง

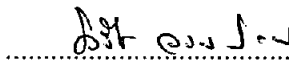
ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษาดุสิตบัณฑิต สาขาวิชาการทดสอบและวัดผลการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

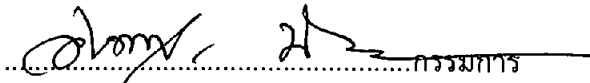



.....คณบดีบัณฑิตวิทยาลัย
(รองศาสตราจารย์ ดร.นภาพร หะวานนท์)
วันที่ ๒๗ เดือน พฤษภาคม พ.ศ 2547

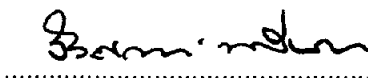
คณะกรรมการสอบปริญญาโท


.....ประธาน
(รองศาสตราจารย์ ดร.ส.วาสนา ประवालพฤษ)


.....กรรมการ
(รองศาสตราจารย์ ดร.วิรัช วรรณรัตน์)


.....กรรมการ
(อาจารย์ ดร.องอาจ นัยพัฒน์)


.....กรรมการที่แต่งตั้งเพิ่มเติม
(รองศาสตราจารย์ ดร.บุญเชิด ภิญโญนนท์พงษ์)


.....กรรมการที่แต่งตั้งเพิ่มเติม
(อาจารย์ ดร.รังสรรค์ มณีเล็ก)

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ สำเร็จสมบูรณ์ได้เพราะได้รับความกรุณาจากประธานและกรรมการควบคุมปริญญานิพนธ์ รองศาสตราจารย์ ดร.ส.วาสนา ประवालพฤษ์ รองศาสตราจารย์ ดร.วิรัช วรณรัตน์ ดร.องอาจ นัยพัฒน์ กรรมการแต่งตั้งเพิ่มเติม รองศาสตราจารย์ ดร.บุญเชิด ภิญโญนนตพงษ์ และ ดร.รังสรรค์ มณีเล็ก ที่ได้กรุณาให้คำปรึกษาชี้แนะในสิ่งอันเป็นประโยชน์ต่อการทำวิจัย ตลอดจนแนะนำแก้ไขข้อบกพร่องต่าง ๆ ด้วยความเอาใจใส่อย่างดียิ่ง ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณผู้บริหารโรงเรียน อาจารย์ผู้สอนทุกท่าน อาจารย์ลัดดา พรหมเมศรี อาจารย์วรรณมา บุญฉิม อาจารย์วิไล แจ่มเจริญ ที่ให้ความช่วยเหลือประสานงาน และอำนวยความสะดวกด้วยห้องคอมพิวเตอร์สำหรับการเก็บรวบรวมข้อมูล

ขอขอบคุณนักเรียนที่เป็นกลุ่มตัวอย่าง ที่ให้ความร่วมมือในการเก็บรวบรวมข้อมูล โดยทำการทดสอบด้วยความเอาใจใส่เป็นอย่างดี

ขอขอบพระคุณคุณผู้เชี่ยวชาญทุกท่านที่ให้คำแนะนำ และตรวจสอบขั้นตอนการสร้างแบบทดสอบ และขั้นตอนการพัฒนาโปรแกรมการทดสอบ จนลุล่วงด้วยดี

ขอขอบคุณสถาบันเทคโนโลยีราชมงคล ที่สนับสนุนเวลาและทุนการศึกษาแก่ผู้วิจัย

ขอขอบคุณ ดร.ฉัตรศิริ ปิยะพิมลสิทธิ์ ที่กรุณาช่วยเหลือด้านการประมวลผลและจัดรูปแบบปริญญานิพนธ์

ขอขอบคุณ คุณวรรณแสงชัย แซ่เฮง ผู้เชี่ยวชาญและให้คำปรึกษาในการเขียนโปรแกรม และขอขอบคุณคุณคุณวิชาญ ชำนาญยา ที่กรุณาช่วยเหลือด้านการสืบค้นข้อมูล

ขอขอบพระคุณวัดผลรุ่นพี่ รุ่นน้อง เพื่อนๆ และผู้เกี่ยวข้องทุกคน ที่มีส่วนช่วยเหลือในการทำปริญญานิพนธ์ฉบับนี้

ท้ายที่สุดนี้ ผู้วิจัยขอน้อมรำลึกถึงพระคุณของบิดา-มารดา ครู-อาจารย์ และญาติพี่น้องทุกท่าน ที่ให้กำลังใจ และสนับสนุนผู้วิจัยด้วยดีตลอดมา

เกียรติศักดิ์ ส่องแสง

ปริญญานิพนธ์นี้ได้รับทุนอุดหนุนจากทบวงมหาวิทยาลัย
ประเภททุน "ปริญญานิพนธ์ที่สร้างองค์ความรู้ใหม่ที่เป็นประโยชน์ต่อวงวิชาการ"

สารบัญ

บทที่	หน้า
1 บทนำ.....	1
ภูมิหลัง.....	1
วัตถุประสงค์ของการวิจัย.....	5
ความสำคัญของการวิจัย.....	6
ขอบเขตของการวิจัย.....	6
สมมุติฐานในการวิจัย.....	8
กรอบความคิดในการวิจัย.....	11
นิยามศัพท์เฉพาะ.....	12
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	15
การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์.....	15
ทฤษฎีการตอบสนองข้อสอบ.....	28
งานวิจัยที่เกี่ยวข้อง.....	45
3 วิธีการดำเนินการวิจัย.....	64
ประชากรและกลุ่มตัวอย่าง.....	64
การดำเนินการวิจัยระยะที่ 1.....	68
การพัฒนาเครื่องมือที่ใช้ในการวิจัย.....	68
การเก็บรวบรวมข้อมูล.....	87
การวิเคราะห์ข้อมูลและสถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	89
การดำเนินการวิจัยระยะที่ 2.....	90
เครื่องมือที่ใช้ในการวิจัย.....	90
การเก็บรวบรวมข้อมูล.....	91
การวิเคราะห์ข้อมูลและสถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	92
4 ผลการวิเคราะห์ข้อมูล.....	94
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล.....	94
การนำเสนอผลการวิเคราะห์ข้อมูล.....	94
ผลการวิเคราะห์ข้อมูลในขั้นตอนเตรียมการ.....	95
ค่าพารามิเตอร์ของข้อสอบ.....	95
การวิเคราะห์ข้อมูลเพื่อพัฒนาและประเมินโปรแกรมการทดสอบ.....	104

สารบัญ(ต่อ)

บทที่	หน้า
4 (ต่อ) ผลการวิเคราะห์ข้อมูลในขั้นตอนหลัก.....	112
การวิเคราะห์ค่าสถิติพื้นฐาน.....	112
การวิเคราะห์ข้อมูลเพื่อศึกษาผลของตัวแปรอิสระที่มีต่อตัวแปรตาม.....	116
5 สรุป อภิปราย และข้อเสนอแนะ.....	123
วัตถุประสงค์ของการวิจัย.....	123
สมมุติฐานการวิจัย.....	123
วิธีการดำเนินการวิจัย.....	123
สรุปผลการวิจัย.....	125
อภิปรายผลการวิจัย.....	130
ข้อเสนอแนะ.....	134
บรรณานุกรม.....	137
ภาคผนวก ก.....	143
ภาคผนวก ข.....	156
ภาคผนวก ค.....	193
ประวัติย่อของผู้วิจัย.....	220

บัญชีตาราง

ตาราง	หน้า
1 ความแตกต่างระหว่างแบบทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ กับแบบทดสอบที่ใช้กระดาษเขียนตอบ.....	19
2 โครงสร้างของการจัดแบบทดสอบเฟลิกซิเลเวล.....	26
3 ลักษณะของข้อมูล โมเดลที่ใช้วิเคราะห์และผู้นำเสนอโมเดล.....	31
4 จำนวนนักเรียนกลุ่มตัวอย่าง จำแนกเป็นรายโรงเรียน ขนาดของโรงเรียน และห้องเรียน.....	65
5 จำนวนนักเรียนกลุ่มตัวอย่าง จำแนกเป็นรายโรงเรียน ขนาดของโรงเรียน และห้องเรียน.....	67
6 การแบ่งชุดข้อสอบตามช่วงระดับความยาก.....	71
7 จำนวนชุดข้อสอบที่นำไปทดสอบเพื่อกำหนดเวลาที่ใช้ในการสอบ.....	72
8 แสดงการกำหนดคุณลักษณะของข้อสอบตามตัวแปรที่ศึกษา.....	75
9 ค่าอำนาจจำแนก ค่าความยากง่าย ค่าการเดา และความเบี่ยงเบนมาตรฐานของข้อสอบ วัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 จำนวน 54 ข้อ ที่ผู้วิจัยสร้างเพิ่มเติมขึ้น.....	95
10 ค่าไอเกน (eigen value) ที่ได้จากการวิเคราะห์องค์ประกอบของแบบทดสอบ วัดความเข้าใจเรื่องเศษส่วน จำนวน 14 องค์ประกอบ ที่มีค่าไอเกน มากกว่า 1.....	98
11 ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนรายข้อ จำนวน 44 ข้อ กับคะแนนของแบบทดสอบ วัดความเข้าใจเรื่องเศษส่วนทั้งฉบับ.....	100
12 จำนวนข้อสอบ จำนวนผู้เข้าสอบ ค่าเฉลี่ย และความเบี่ยงเบนมาตรฐาน ของเวลาที่ใช้ในการทดสอบ จากการทดสอบจากกลุ่มตัวอย่าง เพื่อกำหนดเวลาที่ใช้ ในการตอบแต่ละข้อของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์.....	104
13 ชุดของข้อสอบ ความยากง่ายของข้อสอบ และเวลาที่ใช้ในการตอบข้อสอบแต่ละข้อ โดยการกำหนดเวลาให้กับข้อสอบในคลังแบ่งเป็น 12 ชุด.....	105
14 ผลการวิเคราะห์ ค่าอำนาจจำแนกรายข้อของแบบสอบวัดความพอใจในการสอบ ด้วยการทดสอบแบบปรับเหมาะกับความสามารถ ของผู้สอบด้วยคอมพิวเตอร์.....	107
15 ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนรายข้อกับคะแนนรวมทั้งฉบับของแบบสอบ วัดความพึงพอใจในการสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ด้วยคอมพิวเตอร์.....	109

บัญชีตาราง (ต่อ)

ตาราง	หน้า
16 จำนวนผู้สอบ ค่าเฉลี่ย และค่าความเบี่ยงเบนมาตรฐานของคะแนนความพึงพอใจในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ รูปแบบที่ไม่มีสื่อประสมและรูปแบบที่มีสื่อประสม.....	111
17 ผลการวิเคราะห์ จำนวนผู้เข้าสอบ ค่าเฉลี่ย และค่าความเบี่ยงเบนมาตรฐานของเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ จำแนกตามระดับของอิทธิพลหลักของตัวแปรต้น 3 ตัว.....	112
18 ผลการวิเคราะห์ข้อมูล จำนวนผู้เข้าสอบ ค่าเฉลี่ย และค่าความเบี่ยงเบนมาตรฐานของเวลาที่ใช้ทำแบบทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จากอิทธิพลร่วมของระดับต่างๆ ของตัวแปรต้นร่วมกัน 2 ระดับ.....	113
19 ผลการวิเคราะห์ข้อมูล จำนวนผู้เข้าสอบ ค่าเฉลี่ย และค่าความเบี่ยงเบนมาตรฐานของเวลาที่ใช้ทำแบบทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จากระดับต่างๆ ของตัวแปรต้นร่วมกัน 3 ระดับ.....	115
20 ผลการวิเคราะห์ความแปรปรวนของเวลาที่ใช้ในการทดสอบ จำนวนข้อสอบของการทดสอบ และค่าฟังก์ชันสารสนเทศของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามแหล่งความแปรปรวน.....	117
21 ความแตกต่างของค่าเฉลี่ยของเวลาที่ใช้ในการทดสอบ จำนวนข้อที่ทำการทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามตัวแปรจำนวนข้อสอบในแต่ละชั้น.....	119
22 ความแตกต่างของค่าเฉลี่ยของเวลาที่ใช้ในการทดสอบ จำนวนข้อที่ทำการทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามตัวแปรเวลาที่ใช้ในการทำข้อสอบแต่ละข้อ.....	120
23 ความแตกต่างของค่าเฉลี่ยของเวลาที่ใช้ในการทดสอบ จำนวนข้อที่ทำการทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามตัวแปรเวลาที่ใช้ในการทำข้อสอบแต่ละข้อ.....	121
24 เวลาที่ใช้ในการทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามอิทธิพลของตัวแปรต้น ได้แก่การจัดชุดข้อสอบให้ตอบในแต่ละชั้น การกำหนดเวลาให้ตอบในแต่ละข้อ และการจัดรูปแบบของข้อสอบที่ไม่มีสื่อประสมกับรูปแบบที่มีสื่อประสม.....	122

บัญชีภาพประกอบ

ภาพประกอบ	หน้า
1 กรอบความคิดในการวิจัยเพื่อศึกษาฟังก์ชันสารสนเทศของแบบทดสอบ แบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์.....	11
2 โครงสร้างของการทดสอบเฉพาะบุคคลแบบสองชั้นตอน.....	22
3 โครงสร้างของการทดสอบเฉพาะบุคคลรูปแบบพีระมิดขนาดชั้นคงที่.....	23
4 โครงสร้างของแบบทดสอบรูปพีระมิดขนาดชั้นแปรผัน.....	23
5 โครงสร้างของแบบทดสอบรูปพีระมิดแบบโรบินส์มอนโร.....	24
6 โครงสร้างของแบบทดสอบรูปพีระมิดแบบสกัดการสะท้อนกลับ.....	24
7 โครงสร้างของแบบทดสอบรูปพีระมิดแบบรักษาการสะท้อนกลับ.....	25
8 โครงสร้างของแบบทดสอบรูปพีระมิดแบบมีหลายข้อในแต่ละชั้น.....	25
9 โครงสร้างของแบบทดสอบรูปพีระมิดแบบให้น้ำหนักแก่ตัวเลือก ของข้อสอบเพื่อแยกทาง.....	26
10 โครงสร้างของแบบทดสอบแบบปรับเหมาะกับความสามารถ ของผู้สอบแบบแบ่งชั้น.....	27
11 โค้งแสดงลักษณะของข้อสอบตามโมเดลโลจิสติกแบบ 3 พารามิเตอร์.....	33
12 โค้งแสดงลักษณะของข้อสอบตามโมเดลโลจิสติกแบบ 1 พารามิเตอร์.....	34
13 ช่วงความเชื่อมั่น (confidence interval ; $\underline{\theta}$, $\bar{\theta}$) ของความสามารถที่ได้จากการ ประมาณค่า (θ)	36
14 ช่วงแห่งความเชื่อมั่นเอซิมโทติก 95% ($\underline{\theta}$, $\bar{\theta}$) สำหรับความสามารถ θ	37
15 แสดงผังการทำงานของโปรแกรมคอมพิวเตอร์ในการดำเนินการสอบ แบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์.....	76
16 โลโก้หน้าแรกแสดงโปรแกรมจัดพิมพ์แบบทดสอบ.....	77
17 หน้าจอคำสั่งเพิ่ม-สร้าง.....	78
18 หน้าจอโปรแกรมสร้างแบบทดสอบ.....	78
19 หน้าจอโปรแกรมสำหรับทำแบบทดสอบ.....	79
20 หน้าจอโปรแกรมสำหรับทำแบบทดสอบ.....	80
21 หน้าจอโปรแกรมเมื่อบันทึกผลการสอบเรียบร้อยแล้ว.....	81
22 หน้าจอโปรแกรมแสดงผลการทดสอบ.....	81
23 ลำดับขั้นของการสร้างเครื่องมือวัดความพึงพอใจในการสอบ	85
24 ค่าไอเกน (eigen value) ที่ได้จากการวิเคราะห์องค์ประกอบของแบบทดสอบ วัดความเข้าใจ เรื่องเศษส่วน จำนวน 14 องค์ประกอบที่มีค่าไอเกน มากกว่า 1.....	99

บัญชีภาพประกอบ(ต่อ)

ภาพประกอบ	หน้า
25 จำนวนข้อสอบวัดความเข้าใจ เรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 ในคลังข้อสอบเดิม 200 ข้อ และข้อสอบใหม่ 44 ข้อ จำแนกตามช่วงของค่าอำนาจจำแนก.....	101
26 จำนวนข้อสอบวัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 ในคลังข้อสอบ จำนวน 244 ข้อ จำแนกตามช่วงของค่าความยากง่าย.....	102
27 จำนวนข้อสอบวัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 ในคลังข้อสอบ 244 ข้อ จำแนกตามช่วงของค่าการเดา.....	103

บทที่ 1

บทนำ

ภูมิหลัง

พระราชบัญญัติการศึกษาแห่งชาติ พุทธศักราช 2542 หมวด 4 มาตรา 26 กล่าวถึงการประเมินผู้เรียนไว้ว่า ให้สถานศึกษาจัดการประเมินผู้เรียนโดยพิจารณาจากพัฒนาการของผู้เรียน ความประพฤติ การสังเกตพฤติกรรมการเรียน การร่วมกิจกรรมและการทดสอบควบคู่ไปในกระบวนการเรียนการสอน (สำนักงานคณะกรรมการการศึกษาแห่งชาติ. 2542) การประเมินผู้เรียนโดยใช้แบบทดสอบยังมีความจำเป็นและใช้กันอย่างแพร่หลายในปัจจุบัน ถึงแม้ว่านักการศึกษา ครูอาจารย์ นักวิชาการ ตลอดจนนักวัดผล ได้มีความพยายามในการพัฒนารูปแบบและเครื่องมือการประเมินผู้เรียน ให้มีความหลากหลายยิ่งขึ้น เพื่อให้สามารถวัดคุณลักษณะของผู้เรียนได้เที่ยงตรง ครอบคลุมมีความน่าเชื่อถือและได้ข้อมูลที่มีความหมายต่อตัวผู้เรียน เครื่องมือวัดผลต่างๆ ที่ได้พัฒนาขึ้นนั้น ยังไม่มีเครื่องมือชนิดใดใช้แทนแบบทดสอบได้ทั้งหมด เพราะการวัดและการประเมินผลการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญต้องวัดและประเมินให้ครอบคลุมทุกด้าน ทั้งในส่วนของกระบวนการและผลงาน ทั้งด้านความรู้ ด้านความรู้สึก และทักษะการแสดงออกทุกด้าน และการประเมินตามสภาพจริง (ส.วาสนา ประवालพฤกษ์. 2545) ทั้งนี้เนื่องจากคุณลักษณะที่พึงประสงค์ในตัวผู้เรียนนั้นไม่สามารถใช้เครื่องมือชนิดเดียวทำการประเมินแล้วได้ข้อมูลสะท้อนตัวผู้เรียนได้ทั้งหมด คุณลักษณะต่าง ๆ ในตัวผู้เรียนยังต้องพึงเครื่องมือหลายชนิดทำการประเมิน จึงจะสามารถสะท้อนข้อมูลในตัวผู้เรียนได้อย่างสมบูรณ์ (บุญเชิด ภิญญอนันตพงษ์. 2545) แบบทดสอบยังคงเป็นเครื่องมือที่มีความสำคัญและจำเป็นต้องใช้ประเมินคุณลักษณะในตัวผู้เรียน โดยเฉพาะคุณลักษณะทางด้านความรู้ความสามารถหรือผลสัมฤทธิ์ทางการเรียน เพราะแบบทดสอบเป็นเครื่องมือที่สามารถใช้วัดได้ครอบคลุมเนื้อหา คุณลักษณะทางด้านสติปัญญา และยังสามารถวัดผล ผู้เรียนได้ครั้งละหลาย ๆ คนพร้อมกัน ประหยัดทั้งเวลา ค่าใช้จ่าย สะดวกในการแปลความหมาย และให้ผลเป็นที่น่าเชื่อถือ แบบทดสอบที่มีคุณภาพจะสามารถจำแนกผู้เรียน และให้ความยุติธรรมแก่ผู้เรียนได้เป็นอย่างดี เพราะแบบทดสอบมีความชัดเจนหรือมีความเป็นปรนัยในการให้คะแนนสูง ในขณะที่เครื่องมือวัดผลชนิดอื่น ๆ มีความชัดเจนในการให้คะแนนน้อยกว่า จึงอาจเกิดความคลาดเคลื่อนในการให้คะแนนได้มากกว่าโดยเฉพาะถ้าผู้ให้คะแนนมีความลำเอียง หรือไม่มีความรอบคอบในการให้คะแนนยังมีโอกาสทำให้ขาดความยุติธรรมต่อผู้สอบมากยิ่งขึ้น

แบบทดสอบที่ใช้อยู่โดยทั่วไปมีหลายชนิดและหลายรูปแบบ ที่ใช้กันอย่างแพร่หลายคือแบบทดสอบที่ใช้สอบกับผู้สอบครั้งละหลาย ๆ คนโดยใช้แบบทดสอบฉบับเดียวกันหรือบางครั้งอาจใช้แบบทดสอบที่มีความเป็นคู่ขนานกัน จำนวนข้อสอบหลาย ๆ ข้อ ความยากง่ายของข้อสอบก็ต้องกระจายออกไปให้ครอบคลุมความสามารถของผู้เข้าสอบเหล่านั้น โดยเฉพาะอย่างยิ่งถ้าผู้สอบมีจำนวนมาก ๆ และมีความสามารถแตกต่างกันสูง ข้อสอบจะต้องมีความยากง่ายกระจายออกไปเพื่อให้แบบทดสอบสามารถวัดได้ครอบคลุมช่วงความสามารถของผู้สอบกลุ่มนั้น โดยเฉพาะถ้าข้อสอบมีความ

ยาก-ง่ายไม่ครอบคลุมช่วงความสามารถของผู้เข้าสอบทั้งหมด อาจจะทำให้แบบทดสอบฉบับนั้นยากเกินไป ถ้าช่วงความยากง่ายของข้อสอบไม่ครอบคลุมผู้ที่มีความสามารถต่ำ และแบบทดสอบฉบับนั้นอาจจะง่ายเกินไปถ้าแบบทดสอบฉบับนั้นมีความยากง่ายพอเหมาะกับผู้ที่มีความสามารถต่ำ ด้วยเหตุผลดังกล่าวแบบทดสอบที่ดีตามทฤษฎีมาตรฐานเดิม (Classical Test Theory) จึงเป็นแบบทดสอบที่มีข้อสอบจำนวนมาก และค่าความยากต้องกระจายครอบคลุมจากผู้ที่มีความสามารถต่ำจนถึงผู้ที่มีความสามารถสูง แต่เมื่อพิจารณาให้ถี่จะพบปัญหาที่ว่ามีข้อสอบบางส่วนไม่เหมาะสมกับผู้เข้าสอบ กล่าวคือจะมีข้อสอบบางส่วนเป็นข้อสอบที่ยากมากสำหรับผู้สอบที่มีความสามารถต่ำ แต่พอเหมาะกับผู้ที่มีความสามารถสูง ดังนั้นข้อสอบเหล่านี้จะทำให้ผู้สอบเบื่อหน่าย หรือท้อแท้หมดกำลังใจในการสอบ หรืออาจจะทำข้อสอบโดยวิธีการเดา ในทางตรงกันข้ามจะมีข้อสอบบางส่วนอาจง่ายเกินไปสำหรับผู้สอบที่มีความสามารถสูง จึงอาจทำให้ผู้สอบทำข้อสอบด้วยความสะเพร่าจนเสียคะแนนไปทั้ง ๆ ที่มีความสามารถเหนือข้อสอบนั้นๆ อยู่มาก สาเหตุดังกล่าวจึงทำให้เกิดความคลาดเคลื่อนในการวัดสูงขึ้น ดังนั้นจึงไม่ควรใช้แบบทดสอบฉบับเดียวกันสอบวัดนักเรียนที่มีความสามารถต่างกัน (ต่าย เชียงฉิน. 2534 ; ศิริชัย กาญจนวาสี. 2538 ; รังสรรค์ มณีเล็ก. 2540) แบบทดสอบที่มีประสิทธิภาพสูงนั้นต้องเป็นแบบทดสอบที่เหมาะสมกับความสามารถของผู้สอบ คือแบบทดสอบที่ไม่ยากเกินไปหรือง่ายเกินไปสำหรับผู้เข้าสอบแต่ละคน ผู้สอบที่มีความสามารถสูงควรได้รับข้อสอบยาก ส่วนผู้ที่มีความสามารถต่ำก็ควรได้รับ ข้อสอบง่าย หลังจากการตอบข้อสอบจึงทำการประมาณค่าความสามารถของผู้สอบลงบนคุณลักษณะที่ถูกวัด แปลผลระดับความสามารถตามความจริงที่ว่าผู้ที่มีความสามารถสูงทำข้อสอบข้อที่ยากกว่าถูกต้องย่อมได้คะแนนมากกว่าคนที่มีความสามารถต่ำที่ทำข้อสอบข้อที่ง่ายถูก การทดสอบด้วยแบบทดสอบดังกล่าวนี้ เรียกว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive Testing) และเมื่อนำคอมพิวเตอร์มาประยุกต์ใช้ในกระบวนการสอบจึงเรียกรูปแบบการสอบในลักษณะนี้ว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized Adaptive Testing) (Weiss.1973 ; ศิริชัย กาญจนวาสี. 2538)

จุดมุ่งหมายของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เป็นการสร้างสรรค์แบบทดสอบให้เหมาะสม (optimal test) กับความสามารถของผู้เข้าสอบแต่ละคน การที่จะได้มาซึ่งแบบทดสอบนี้ ผู้เข้าสอบจะถูกวัดระดับคุณลักษณะหรือความสามารถ (Trait level : θ) ในระหว่างการสอบ และข้อสอบที่เหมาะสมกับความสามารถของผู้เข้าสอบจะถูกเลือกจากธนาคารข้อสอบ (Item bank) ให้กับผู้เข้าสอบ โดยอาศัยรูปแบบทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT model) ซึ่งเชื่อว่าเป็นทฤษฎีที่สามารถอธิบายระดับความสามารถของผู้สอบจากพฤติกรรมกรรมการตอบข้อสอบ มาเป็นพื้นฐานในการดำเนินการสอบ ซึ่งผู้เข้าสอบแต่ละคนจะได้รับแบบทดสอบที่แตกต่างกันออกไป ซึ่งแตกต่างจากการทดสอบแบบเขียนตอบลงบนกระดาษ (Paper-and-pencil test) ที่ผู้เข้าสอบทุกคนจะได้รับข้อสอบเหมือนกันหมดทุกข้อและจำนวนข้อสอบเท่ากันหมด (ศิริชัย กาญจนวาสี. 2538 ; Meijer and Michael. 1999)

จากรายงานการวิจัยเกี่ยวกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์นั้นได้ถูกพัฒนาขึ้นเป็นครั้งแรกในทศวรรษ 1970 และ 1980 โดย Lord พัฒนาขึ้นในปี 1971 และปี 1977 ส่วน Weiss พัฒนาขึ้นในปี 1982 และปี 1983 หลังจากนั้นได้มีรายงานวิจัยนำเสนอเรื่องนี้

ออกมาอย่างต่อเนื่อง เช่น Van Der Linden และ Glas, in press; Wainer นำเสนอในปี 1990 ส่วน Sands Waters และ McBride นำเสนอในปี 1997; Drasgow และ Olson-Buchan นำเสนอในปี 1999 และในงานประชุมสัมมนาประจำปีของ National Council on Measurement in Education ในปี 1999 มีงานวิจัยกว่า 25% นำเสนอรายงานวิจัยเกี่ยวกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ หัวข้อของงานวิจัยนั้นนำเสนอเรื่อง การพัฒนารายการข้อสอบ กระบวนการการคัดเลือกข้อสอบ กระบวนการการประมาณค่าความสามารถ และหัวข้อที่เกี่ยวข้อง เช่นความปลอดภัยของแบบทดสอบ และความเชื่อมั่นของคะแนนสอบ (Meijer and Michael. 1999) เนื่องจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ทำการประมาณค่าความสามารถของผู้สอบโดยให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Test information function) ได้เที่ยงตรง (Validity) และเชื่อถือได้ (Reliability) มากยิ่งขึ้นโดยใช้จำนวนข้อสอบและเวลาในการทดสอบเพียงหนึ่งในสามของการทดสอบแบบดั้งเดิม (Conventional test) (Weiss. 1983 ; ศิริชัย กาญจนวาสี. 2538) ซึ่งสอดคล้องกับผลการวิจัยของ โมรีโนและคนอื่น ๆ (ต่าย เชียงฉิน. 2534 ; อ้างอิงจาก Moreno and others. 1984) ที่พบว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ให้ค่าความเที่ยงตรงและค่าความเชื่อมั่นเท่ากับหรือมากกว่าการทดสอบแบบดั้งเดิมทั้ง ๆ ที่ใช้จำนวนข้อสอบน้อยกว่าถึงครึ่งหนึ่ง ดังนั้นจึงทำให้มีผู้สนใจศึกษาและนำการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มาใช้ในกระบวนการทดสอบวัดผลมากยิ่งขึ้น

สหรัฐอเมริกา นำการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มาทำการทดสอบเชิงปฏิบัติการกับการทดสอบหลายประเภท เช่น Placement Test โดย College Board ในปี 1993 และการทดสอบ Graduate-Record Examination โดย Educational Testing Service ในปี 1996 และยังมีคณะกรรมการออกใบอนุญาตหลายแขนงได้นำเอารูปแบบการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ไปใช้ในการทดสอบเพื่อออกใบอนุญาตของแต่ละหน่วยงาน รวมถึง The National Council of State Boards of Nursing and National Board of Medical Examiners นอกเหนือจากนี้ The U.S. Department of Defense ยังได้นำเอารูปแบบการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มาใช้ใน The Armed Services Vocational Aptitude Battery ในปี 1997 นอกจากนี้จะเป็นที่นิยมอย่างมากในสหรัฐอเมริกาแล้วการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ยังใช้อย่างกว้างขวางในเนเธอร์แลนด์ โดย The National Institute for Educational Measurement ได้เสนอการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์สอง Version ออกมาเพื่อใช้สำหรับกำหนดระดับความสามารถของผู้สอบในวิชาคณิตศาสตร์ และใช้ในการประเมินผลสัมฤทธิ์ในวิชาเฉพาะ คณิตศาสตร์ (Meijer and Michael. 1999)

ประเทศไทยมีผู้ศึกษาวิจัยและพัฒนาโปรแกรมสำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์และศึกษาผลของตัวแปรต่าง ๆ ดังนี้ นันทิยา พึ่งคำ (2531) ทำการเปรียบเทียบคุณภาพการทดสอบแบบซี เอ ที และแบบประเพณีนิยม ในการวัดความสามารถด้านคำศัพท์ภาษาอังกฤษ ผลปรากฏว่าการทดสอบทั้งสองแบบมีความตรงเชิงเกณฑ์สัมพัทธ์ค่อนข้างสูงและมีความตรงไม่แตกต่างกัน ส่วนค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ (SEE) ของการ

ทดสอบแบบ ซี เอ ที ต่ำกว่าการทดสอบแบบประเพณีนิยม และจำนวนข้อสอบของการทดสอบแบบ ซี เอ ที น้อยกว่าครึ่งหนึ่งของการทดสอบแบบประเพณีนิยม ดริงใจ พูนผลอำนาจ (2534) ได้พัฒนา โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ โดยใช้รูปแบบปิรามิด ขนาดชั้น คงที่ 10 ชั้น โดยใช้คอมพิวเตอร์ ผลปรากฏว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ดังกล่าวให้ความสะดวกในการทดสอบมากยิ่งขึ้น ผู้สอบให้ความสนใจ และสร้างความดึงดูดใจทำให้ผู้สอบใส่ใจในการทำข้อสอบเพิ่มมากขึ้น รังสรรค์ มณีเล็ก (2540) ทำการสร้างโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบและศึกษาผลของตัวแปรค่าความยากง่ายของข้อสอบข้อแรก เกณฑ์การคัดเลือกข้อสอบจากคลังข้อสอบ วิธีการประมาณค่าความสามารถของผู้สอบ และตัวแปรเกณฑ์ยุติการทดสอบ ต่อความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการทดสอบ การสร้างโปรแกรมการทดสอบโดยใช้ภาษาซี (C Language) ในการเขียนโปรแกรม เนื่องจากภาษาซีมีความ ยืดหยุ่น ง่ายต่อการใช้คำสั่งและเกิดการผิดพลาดน้อย สำหรับผลของตัวแปรที่ศึกษาพบว่า การเริ่มต้นสอบด้วยข้อสอบที่ง่าย ปานกลาง หรือยาก ไม่ทำให้ความเที่ยงตรงเชิงสภาพแตกต่างกัน แต่ควรเริ่มต้นสอบด้วยข้อสอบที่มีความยากง่ายปานกลาง เพราะจะใช้จำนวนข้อสอบน้อยกว่า

การคัดเลือกข้อสอบข้อต่อไปในคลังข้อสอบ 3 วิธี คือ ใช้ค่าความสามารถของผู้สอบจากการทำข้อสอบที่ผ่านมาแล้วเทียบกับค่าความยากง่ายของข้อสอบในคลัง วิธีใช้ค่าความสามารถของผู้สอบจากการทำข้อสอบที่ผ่านมาแล้วเทียบกับค่าความสามารถที่ทำให้ข้อสอบมีสารสนเทศสูงสุด และใช้ค่าสารสนเทศของข้อสอบที่สูงที่สุด พบว่า การคัดเลือกข้อสอบจาก 3 วิธี ไม่ทำให้ค่าความเที่ยงตรงเชิงสภาพแตกต่างกัน แต่วิธีที่ 2 และ 3 จะใช้ข้อสอบน้อยกว่าวิธีแรก และพบว่าวิธีที่ 2 เหมาะสำหรับคลังข้อสอบขนาดเล็ก วิธีที่ 3 เหมาะกับคลังข้อสอบขนาดใหญ่

การประมาณค่าความสามารถของผู้สอบจากการศึกษาเปรียบเทียบวิธีการประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์และวิธีความเป็นไปได้สูงสุดพบว่าวิธีการประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์มีความเที่ยงตรงเชิงสภาพสูงกว่าการประมาณค่าความสามารถด้วยวิธีความเป็นไปได้สูงสุด

สำหรับเกณฑ์ยุติการทดสอบนั้น พบว่าเกณฑ์ยุติการสอบต่างกันทำให้ความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการสอบต่างกัน ถ้าใช้เกณฑ์ค่าความคลาดเคลื่อนมาตรฐานน้อยกว่าหรือเท่ากับ .30 จะใช้จำนวนข้อสอบมากกว่าการใช้เกณฑ์ค่าความคลาดเคลื่อนมาตรฐานเท่ากับ หรือน้อยกว่า .40 อยู่ถึง 2 เท่า ดังนั้น ถ้าเป็นการทดสอบโดยทั่ว ๆ ไปจึงควรใช้เกณฑ์ค่าความคลาดเคลื่อนมาตรฐาน .40 จะได้ค่าความเชื่อมั่นในการวัดเท่ากับ .9165 แต่ใช้ข้อสอบจำนวนน้อยกว่าครึ่ง

ส่วนตัวแปรความสามารถของผู้สอบ พบว่า ผู้สอบที่มีความสามารถต่างกันจะใช้ข้อสอบมากน้อยต่างกัน และมีความเที่ยงตรงเชิงสภาพต่างกันด้วย ผู้สอบที่มีความสามารถสูงกว่าจะใช้ข้อสอบน้อยกว่า และความเที่ยงตรงเชิงสภาพสูงกว่าด้วยเช่นกัน สายชล ออบทม (2539) ทำการพัฒนาโปรแกรม สำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ และทำการประเมินผลการใช้โปรแกรม ผลปรากฏว่าได้โปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์รูปแบบทางแยกแปรผัน ประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์ ใช้ภาษาปาสคาลในการเขียนโปรแกรม โปรแกรม

ปฏิบัติงานบนวินโดวส์ 95 และยุติการทดสอบเมื่อความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบต่ำกว่า 0.3 ส่วนผลการใช้โปรแกรมการทดสอบนั้น พบว่านักเรียนมีความกระตือรือร้นที่จะทำการทดสอบ มีเจตคติที่ดีต่อการทดสอบ และอยากให้มีการทดสอบแบบนี้ในวิชาอื่น ๆ อีกด้วย

สำหรับการค้นคว้าวิจัยเพื่อหาประสิทธิภาพของการทดสอบในเรื่องของโมเดลการตอบสนองข้อสอบนั้น พบว่าโมเดลการตอบสนองข้อสอบตามทฤษฎีการตอบสนองข้อสอบแบบโลจิสติกแบบ 3 พารามิเตอร์มีความเหมาะสมกว่าโมเดลอื่น ๆ (ริงสรณ์ มณีเล็ก. 2540 ; อ้างอิงจาก Urry. 1970; 1977 ; Mckinley and Reckase. 1980 ; Koch and Reckase. 1978) ส่วนคลังข้อสอบนั้น พบว่าขนาดของคลังข้อสอบส่งผลต่อการประมาณค่าของผู้สอบน้อยมาก แต่อย่างไรก็ตามคลังข้อสอบควรมีข้อสอบประมาณ 100 – 200 ข้อ และคลังข้อสอบที่มีข้อสอบ 116 – 150 ข้อ จะให้ผลดีที่สุด (ริงสรณ์ มณีเล็ก. 2540 ; อ้างอิงจาก Ho. 1989; weiss. 1988)

จากการศึกษาค้นคว้าและผลการวิจัยที่กล่าวมา พบว่า ได้มีความพยายามพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์และนำโปรแกรมการทดสอบนี้มาใช้ในการทดสอบพร้อมทั้งศึกษาผลของตัวแปรต่าง ๆ ที่ส่งผลต่อประสิทธิภาพของการทดสอบ ซึ่งพบว่า มีตัวแปรที่มีความชัดเจน จากผลการวิจัยที่ผ่านมา ซึ่งจะทำให้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีประสิทธิภาพสูงสุด คือตัวแปรความยากของข้อสอบที่ใช้เริ่มต้นในการสอบ วิธีการคัดเลือกข้อสอบข้อถัดไปจากคลังข้อสอบ วิธีการประมาณค่าความสามารถของผู้สอบ เกณฑ์ยุติการสอบ ความสามารถของผู้สอบ โมเดลการตอบสนองข้อสอบ ขนาดของคลังข้อสอบ และเจตคติต่อการสอบ ส่วนตัวแปรที่ยังไม่มีความชัดเจน และผู้วิจัยก่อนหน้านี้นี้มักให้ข้อเสนอแนะว่าควรศึกษาตัวแปรที่น่าจะส่งผลต่อคุณภาพของการทดสอบเพิ่มเติม คือ ตัวแปรการจัดชุดข้อสอบให้ตอบในแต่ละชั้น การกำหนดเวลาที่ใช้ในการตอบแต่ละข้อ และการสร้างรูปแบบของข้อสอบที่มีสื่อประสมต่างกัน ดังนั้น จึงทำให้ผู้วิจัยสนใจศึกษาว่าตัวแปรที่ยังไม่ชัดเจนเหล่านี้จะส่งผลต่อเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์หรือไม่ มากน้อยเพียงไร ซึ่งจะเป็นแนวทางในการพัฒนาและนำการทดสอบแบบปรับเหมาะนี้ไปใช้ได้อย่างมีประสิทธิภาพต่อไป

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ รูปแบบหลายขั้นตอนแบบแยกทางแปรผัน โดยการจัดชุดข้อสอบให้ตอบในแต่ละชั้นต่างกัน การกำหนดเวลาที่ใช้ในการตอบแต่ละข้อต่างกัน และการจัดรูปแบบของข้อสอบที่มีสื่อประสมและไม่มี สื่อประสม และศึกษาประสิทธิภาพของโปรแกรมการทดสอบ

2. เพื่อเปรียบเทียบเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ รูปแบบหลายขั้นตอนแบบทางแยกแปรผัน เมื่อมีการจัดชุดข้อสอบให้ตอบในแต่ละชั้นต่างกัน การกำหนดเวลาที่ใช้ในการตอบแต่ละข้อต่างกัน และการจัดรูปแบบของข้อสอบที่ไม่มีสื่อประสมและมีสื่อประสม

ความสำคัญของการวิจัย

จากการศึกษาวิจัยในครั้งนี้ ทำให้ทราบข้อมูลและได้รับผลจากการวิจัยดังต่อไปนี้

1. ได้โปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบที่มีคุณภาพ และมีประสิทธิภาพมากยิ่งขึ้น เพื่อใช้สำหรับการทดสอบวัดสมรรถภาพและเนื้อหาวิชาต่าง ๆ ตามที่ต้องการ และใช้โปรแกรมการทดสอบวัดผลเพื่อปรับปรุงการเรียนการสอนหรือวัดผลสัมฤทธิ์ทางการเรียน โดยที่นักเรียนไม่เบื่อหน่ายการทดสอบ แต่มีความสนใจ กระตือรือร้น เพราะมีสิ่งเร้าสร้างแรงจูงใจจากโปรแกรมการทดสอบด้วยคอมพิวเตอร์ที่พัฒนาขึ้นนี้
2. ได้ข้อมูลเพื่อเป็นแนวทางในการเลือกใช้รูปแบบการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์อย่างมีประสิทธิภาพ มีความน่าเชื่อถือ ประหยัดเวลาในการทดสอบ เพราะใช้ข้อสอบจำนวนน้อยลง และใช้เวลาสอบที่เหมาะสมขึ้น
3. ได้เผยแพร่ความรู้เกี่ยวกับการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ให้เป็นที่ยอมรับหลายออกไป ทั้งนี้เพราะปัจจุบันเทคโนโลยีด้านคอมพิวเตอร์ที่รองรับโปรแกรมการทดสอบมีความพร้อมมากขึ้น โดยเฉพาะอย่างยิ่งในอนาคตอันใกล้นี้คอมพิวเตอร์ยิ่งจะเป็นเครื่องมือสำคัญที่ใช้ในการศึกษาค้นคว้า หรือเป็นแหล่งความรู้ที่สำคัญอีกแหล่งหนึ่งอีกด้วย

ขอบเขตของการวิจัย

1. ประชากรที่ใช้ในการวิจัย
 - ประชากรที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2546 ของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาปทุมธานี จำนวนนักเรียน 7,933 คน จากโรงเรียน 176 โรงเรียน
 2. กลุ่มตัวอย่างที่ใช้ในการวิจัย
 - กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2546 ของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาปทุมธานี จำนวนนักเรียน 968 คน มีห้องเรียน 35 ห้องเรียน จากโรงเรียน 21 โรงเรียน ซึ่งได้มาโดยวิธีการสุ่มแบบแบ่งชั้น (Stratified random sampling) (การสุ่มกลุ่มตัวอย่างมีละเอียดในบทที่ 3)
3. เครื่องมือที่ใช้ในการวิจัย ประกอบด้วย
 - 3.1 โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เป็นโปรแกรมสำเร็จรูปเขียนด้วยโปรแกรมบอร์แลนเดลฟาย (BORLAND DELPHI) บนระบบปฏิบัติการวินโดว์ (Window 98) ขึ้นไปที่ผู้วิจัยพัฒนาขึ้น ซึ่งโปรแกรมสามารถทำงานได้อย่างมีประสิทธิภาพ ภายใต้เงื่อนไขกฎเกณฑ์ของตัวแปรที่ผู้วิจัยทำการศึกษาค้นคว้า ประกอบด้วยโปรแกรมย่อย คือ
 - 3.1.1 โปรแกรมย่อยการพิมพ์ข้อสอบเข้าเก็บไว้ในคลังข้อสอบ เป็นโปรแกรมที่สามารถบันทึกข้อคำถาม คำตอบ คำพารามิเตอร์ของข้อสอบ การกำหนดเงื่อนไขตามตัวแปรที่ศึกษา คือกำหนดการจัดชุดข้อสอบให้ตอบในแต่ละชั้น กำหนดเวลาที่ใช้ในการตอบแต่ละข้อ และกำหนดรูปแบบของข้อสอบที่มีสื่อประสมต่างกัน

3.1.2 โปรแกรมย่อยการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบเป็นโปรแกรมการทดสอบที่กำหนดเงื่อนไขตามตัวแปรที่ศึกษาคือ กำหนดการจัดชุดข้อสอบให้ตอบในแต่ละชั้น การกำหนดเวลาที่ใช้ในการตอบแต่ละข้อ และการกำหนดรูปแบบของข้อสอบที่มีสื่อประสมต่างกัน รวมทั้งหมด 18 โปรแกรม เริ่มต้นการสอบจากการคัดเลือกข้อสอบหรือชุดของข้อสอบที่มีค่าความยากระดับปานกลาง ประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์ โปรแกรมคัดเลือกข้อสอบหรือชุดของข้อสอบถัดไปโดยดูค่าสารสนเทศข้อสอบหรือชุดของข้อสอบที่สูงที่สุด และโปรแกรมจะยุติการสอบเมื่อค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบต่ำกว่าหรือเท่ากับ 0.3

3.2 ข้อสอบที่จะนำมาใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จะต้องเป็นข้อสอบที่ผ่านการวิเคราะห์ค่าพารามิเตอร์ 3 พารามิเตอร์ คือค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าการเดา (c) ตามแนวทฤษฎีการตอบสนองข้อสอบ โดยมีการบันทึกข้อสอบ ค่าพารามิเตอร์ประจำข้อสอบแต่ละข้อ และการกำหนดเงื่อนไขตามตัวแปรที่ศึกษาคือ กำหนดการจัดชุดของข้อสอบให้ตอบในแต่ละชั้น การกำหนดเวลาที่ใช้ในการตอบแต่ละข้อ และการกำหนดรูปแบบของข้อสอบที่มีสื่อประสมต่างกัน ในโปรแกรมการทดสอบย่อย 18 โปรแกรม ซึ่งในการทดสอบเพื่อทำการวิจัยในครั้งนี้ผู้วิจัยขอใช้ข้อสอบวัดความเข้าใจเรื่องเศษส่วน ในระดับชั้นประถมศึกษาปีที่ 6 ของ รังสรรค์ มณีเล็ก จำนวน 200 ข้อ และข้อสอบที่ผู้วิจัยสร้างขึ้นจำนวน 44 ข้อ รวมเป็นข้อสอบ 244 ข้อ ซึ่งเป็นข้อสอบที่มีคุณภาพสร้างขึ้นตามหลักวิชาการสร้างแบบทดสอบและตรวจสอบคุณภาพวิเคราะห์ค่าพารามิเตอร์ตามแนวทฤษฎีการตอบสนองข้อสอบ

3.3 แบบสอบวัดความพึงพอใจในการสอบเพื่อศึกษาประสิทธิภาพของโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เป็นแบบสอบวัดที่ผู้วิจัยสร้างขึ้นเป็นข้อความวัดความพึงพอใจที่เกิดจากการได้รับการตอบสนองความต้องการในปัจจุบันต่าง ๆ ที่ก่อให้เกิดแรงจูงใจหรือกระตุ้นให้ผู้สอบเกิดความผ่อนคลายไม่วิตกกังวลและเต็มใจที่จะทำการสอบปัจจุบันที่ก่อให้เกิดความพึงพอใจในการสอบนี้ ประกอบด้วย รูปแบบการนำเสนอข้อสอบ เวลาที่ใช้ในการตอบข้อสอบแต่ละข้อ กระบวนการและบรรยากาศในการดำเนินการสอบ และการใช้สื่อประกอบการทดสอบ มีลักษณะเป็นข้อความ กำหนดน้ำหนักความพึงพอใจในการสอบไว้ 5 ระดับ คือ พึงพอใจมาก พึงพอใจ เฉยๆ ไม่พึงพอใจ และไม่พึงพอใจอย่างมาก

4. ตัวแปรที่ศึกษาวิจัยในครั้งนี้ประกอบด้วย

4.1 ตัวแปรอิสระ

4.1.1 การจัดชุดข้อสอบให้ตอบในแต่ละชั้น

- 1) ชุดข้อสอบชั้นละ 1 ข้อ
- 2) ชุดข้อสอบชั้นละ 2 ข้อ
- 3) ชุดข้อสอบชั้นละ 3 ข้อ

4.1.2 การกำหนดเวลาที่ใช้ในการตอบข้อสอบแต่ละข้อ

- 1) ไม่กำหนดเวลาในการตอบแต่ละข้อ
- 2) กำหนดเวลาในการตอบแต่ละข้อเฉลี่ยเท่ากันทุกข้อ

3) กำหนดเวลาในการตอบแต่ละข้อตามน้ำหนักค่าความยากของข้อสอบ ซึ่งกำหนดเวลาตามช่วงค่าความยากของข้อสอบได้ข้อสอบ 12 ชุด กำหนดเวลา 12 ค่าตามระดับความยากง่ายของข้อสอบทั้ง 12 ชุด ดังนี้

ชุดที่ 1	ช่วงค่าความยากของข้อสอบตั้งแต่	-3.00	ถึง	-2.51
ชุดที่ 2	ช่วงค่าความยากของข้อสอบตั้งแต่	-2.50	ถึง	-2.01
ชุดที่ 3	ช่วงค่าความยากของข้อสอบตั้งแต่	-2.00	ถึง	-1.51
ชุดที่ 4	ช่วงค่าความยากของข้อสอบตั้งแต่	-1.50	ถึง	-1.01
ชุดที่ 5	ช่วงค่าความยากของข้อสอบตั้งแต่	-1.00	ถึง	-0.50
ชุดที่ 6	ช่วงค่าความยากของข้อสอบตั้งแต่	-0.50	ถึง	-0.01
ชุดที่ 7	ช่วงค่าความยากของข้อสอบตั้งแต่	0.00	ถึง	0.50
ชุดที่ 8	ช่วงค่าความยากของข้อสอบตั้งแต่	0.51	ถึง	1.00
ชุดที่ 9	ช่วงค่าความยากของข้อสอบตั้งแต่	1.01	ถึง	1.50
ชุดที่ 10	ช่วงค่าความยากของข้อสอบตั้งแต่	1.51	ถึง	2.00
ชุดที่ 11	ช่วงค่าความยากของข้อสอบตั้งแต่	2.51	ถึง	2.50
ชุดที่ 12	ช่วงค่าความยากของข้อสอบตั้งแต่	2.51	ถึง	3.00

4.1.3 รูปแบบของข้อสอบ

- 1) รูปแบบของข้อสอบที่ไม่มีสื่อประสม
- 2) รูปแบบของข้อสอบที่มีสื่อประสมภาพกราฟฟิค ภาพวิดิโอ เสียงประกอบการทดสอบ

4.2 ตัวแปรตาม

- 1) เวลาที่ใช้ในการทำแบบทดสอบ
- 2) จำนวนข้อของแบบทดสอบ
- 3) ค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

ผู้สอบด้วยคอมพิวเตอร์

สมมุติฐานในการวิจัย

1. เมื่อจัดชุดข้อสอบให้ตอบในแต่ละชั้นต่างกันค่าฟังก์ชันสารสนเทศของแบบทดสอบน่าจะต่างกัน

ถ้าจัดชุดข้อสอบให้ตอบในแต่ละชั้นเพิ่มมากขึ้น และมีจำนวนไม่มากเกินไปค่าฟังก์ชันสารสนเทศของข้อสอบหรือชุดของข้อสอบในชั้นนั้น ๆ น่าจะมีความน่าเชื่อถือมากยิ่งขึ้น และจะส่งผลต่อค่าฟังก์ชันสารสนเทศของแบบทดสอบให้สูงขึ้นตามไปด้วย ทั้งนี้เนื่องจาก กุลลิกเซน ลอร์ดและโนวิก (ถ่าย เชียงฉวี, 2534 ; อ้างอิงจาก Gulliksen, 1967 ; Lord and Novick, 1968) กล่าวว่า โดยทฤษฎีแล้วจำนวนข้อสอบยิ่งมากขึ้น ค่าความเชื่อมั่นก็ยิ่งสูงขึ้น แต่ถ้าข้อสอบมากเกินไปจะมีผลเสียเช่นกัน กล่าวคือ จะเสียเวลาในการสอบมากขึ้น ทำให้ผู้สอบเกิดความเบื่อหน่ายในการสอบ และมีการทำข้อสอบโดยการเดามากขึ้น ซึ่งจะมีผลให้ค่าความเชื่อมั่นลดลง ดังนั้น ถ้าจัดชุดข้อสอบให้ตอบในแต่ละ

ชั้นมีจำนวนเหมาะสมจึงน่าจะเพิ่มค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงขึ้น แต่จำนวนข้อสอบของแบบทดสอบไม่น่าจะเพิ่มขึ้นเพราะ ถ้าค่าฟังก์ชันสารสนเทศในแต่ละชั้นมีความน่าเชื่อถือมากยิ่งขึ้น ค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบในแต่ละชั้นจะต่ำลง นั่นคือการประมาณค่าความสามารถได้ใกล้เคียงความสามารถจริงของผู้สอบ การทดสอบก็จะยุติเร็วขึ้น (นันทิยา พึ่งคำ. 2531)

2. เมื่อกำหนดเวลาที่ใช้ในการตอบข้อสอบแต่ละข้อตามน้ำหนักค่าความยากของข้อสอบ ฟังก์ชันสารสนเทศของแบบทดสอบจะมีค่ามากที่สุด

ไวส์ สตีเวน (Wise Steven L. <http://ercir.syr.edu/plweb/cgi/fastweb>) ได้กล่าวถึง ประเด็นสำคัญของผู้เข้าทดสอบด้วยการสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (CAT) ว่าในขณะที่ดำเนินการสอบด้วยการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์นั้น มีองค์ประกอบสำคัญที่ควรคำนึงถึงอีกประเด็นหนึ่งคือการกำหนดระยะเวลาที่ใช้สอบ เพราะเวลาที่ใช้ในการสอบนั้นจะเป็นประโยชน์ต่อผู้เข้าสอบ และจะประหยัดเวลาของผู้ดำเนินการสอบด้วย ดังนั้นผู้พัฒนา CAT ควรจะพิจารณาการกำหนดระยะเวลาที่เหมาะสมสำหรับการทดสอบของแต่ละคน และ รังสรรค์ มณีเล็ก. (2540) เสนอแนะไว้ว่า ควรมีการศึกษาเวลาที่ใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เป็นรายข้อว่าจะกำหนดเวลาในลักษณะใด และศึกษาผลของการกำหนดเวลาในลักษณะดังกล่าวกับการไม่กำหนดเวลา จะให้ผลต่างกันหรือไม่

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มีคุณสมบัติแตกต่างจากการทดสอบแบบเขียนตอบลงบนกระดาษคำตอบอีกประการหนึ่งก็คือการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สามารถกำหนดเวลาในการตอบข้อสอบแต่ละข้อได้ตามต้องการซึ่งต่างจากการทดสอบลักษณะอื่น ๆ ที่จะกำหนดเวลาในการตอบข้อสอบแต่ละข้อตามต้องการได้ยาก และเนื่องจากการทดสอบแบบปรับเหมาะนั้น ผู้สอบจะได้รับข้อสอบที่เหมาะสมกับความสามารถของตนทุกข้อ ดังนั้นผู้สอบก็ควรใช้เวลาที่เหมาะสมกับความยากง่ายในการตอบข้อสอบแต่ละข้อนั้นด้วย

การไม่กำหนดเวลาในการตอบข้อสอบแต่ละข้อหรือการกำหนดเวลาในการตอบข้อสอบแต่ละข้อให้เท่ากันทุกข้อ ไม่น่าจะเกิดผลดีต่อการทดสอบ กล่าวคือ ถ้าไม่กำหนดเวลาในการตอบข้อสอบแต่ละข้อ อาจทำให้ผู้สอบขาดความกระตือรือร้นหรือไม่ใส่ใจในการสอบ หรืออาจใช้เวลาในการตอบข้อสอบบางข้อโดยเฉพาะในข้อที่ง่าย ๆ นานเกินไปจะทำให้ผู้สอบเสียเวลาในการทำแบบทดสอบไปโดยไม่เพิ่มความน่าเชื่อถือในการสอบจากการใช้เวลายาวนานขึ้นแต่อย่างใด แต่อาจจะทำให้ผู้สอบมีความประมาทหรือมีความสะเพร่าในการตอบข้อสอบจะส่งผลให้เกิดความคลาดเคลื่อนในการประมาณค่าความสามารถมากยิ่งขึ้นและต้องใช้ข้อสอบมากขึ้นด้วย และถ้ากำหนดเวลาในการตอบข้อสอบแต่ละข้อให้เท่ากันหมดทุกข้อก็ไม่น่าจะสอดคล้องกับความเป็นจริง เพราะว่า ทั้งข้อสอบข้อง่ายและข้อสอบข้อยาก ใช้เวลาในการตอบเท่ากัน ดังนั้นการกำหนดเวลาที่ใช้ในการตอบข้อสอบแต่ละข้อให้เหมาะสมตามระดับความยากของข้อสอบ กล่าวคือ ข้อสอบง่ายใช้เวลาตอบน้อยกว่าข้อสอบยากน่าจะสมเหตุสมผลกว่าผู้เข้าสอบน่าจะใช้เวลาในการสอบน้อยลง แต่ยังคงได้ค่าฟังก์ชันสารสนเทศของแบบทดสอบและใช้จำนวนข้อสอบเท่าเดิม หรืออาจจะเพิ่มค่าฟังก์ชันสารสนเทศของแบบทดสอบและลดจำนวนข้อสอบของแบบทดสอบลง ดังผลการวิจัยของ สิริชัย กาญจนवासี. (2538) และผลการวิจัยของไวส์ โมรีโนและคน

อื่น ๆ (ถ่าย เชียงฉี. 2534 ; อ้างอิงจาก Weiss. 1982 ; Morino and others.1984) ที่พบว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ทำการประมาณค่าความสามารถของผู้สอบโดยให้ค่าสารสนเทศของแบบทดสอบได้เที่ยงตรงและน่าเชื่อถือยิ่งขึ้น โดยใช้จำนวนข้อสอบและเวลาในการสอบลดลงกว่าครึ่ง ถึงหนึ่งในสามของการสอบแบบดั้งเดิม และถ้าผู้สอบได้ทราบว่าคุณเองได้ทำข้อสอบที่มีการกำหนดเวลาให้ตอบแต่ละข้อที่เหมาะสมกับความยากง่ายของข้อสอบที่ตนเองทำจะทำให้ ผู้สอบรู้สึกว่าคุณเองได้รับความยุติธรรมจากการสอบมากยิ่งขึ้น

3. ฟังก์ชันสารสนเทศของแบบทดสอบที่รูปแบบของข้อสอบมีสื่อประสม น่าจะมีค่ามากกว่ารูปแบบของข้อสอบที่ไม่มีสื่อประสม

ทั้งนี้ เนื่องจากเมเยอร์ (Meijer and Michael. 1999); ริงสเร็ค มณีเล็ก. (2540) เสนอแนะไว้ว่า น่าจะศึกษารูปแบบของข้อสอบที่มีภาพกราฟฟิก เสียง และภาพเคลื่อนไหว หรือรูปแบบของข้อสอบที่มีสื่อประสม ซึ่งสอดคล้องกับแนวความคิดของสตรีทแมนและเอจเจน (Straetmens and Eggen. 1998) ที่กล่าวว่า การนำคอมพิวเตอร์มาใช้ในการทดสอบจะได้ข้อสอบที่ตรงกับความต้องการ และสามารถนำเสนอวิธีการทดสอบที่รวมเอาภาพกราฟฟิก เสียง วีดิโอเคลื่อนไหว และเนื้อหาสาระเข้าไว้ด้วยกันจะคล้ายคลึงกับวิถีชีวิตจริง และจากผลการวิจัยของ ตรังใจ พูนผลอำนวย (2534) ที่ได้พัฒนาแบบทดสอบแบบปรับเหมาะกับระดับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ ผลปรากฏว่า การทดสอบสร้างความดึงดูดใจให้ผู้สอบสนใจในการทำข้อสอบเพิ่มขึ้นซึ่งสอดคล้องกับผลการวิจัยของสายชล ออบทม (2539) กล่าวคือ การทดสอบด้วยโปรแกรมสำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ทำให้นักเรียนมีความกระตือรือร้นที่จะทำการทดสอบ และมีทัศนคติที่ดีในการทดสอบ ดังนั้นเมื่อมีการเพิ่มรูปแบบของข้อสอบที่มีสื่อประสมเข้าไปในโปรแกรมการทดสอบ น่าจะสร้างความสนใจแก่ผู้สอบมากยิ่งขึ้น เพราะในโปรแกรมการทดสอบนั้นจะมีทั้งภาพกราฟฟิก ภาพเคลื่อนไหวและเสียง ซึ่งจะช่วยผ่อนคลายในขณะที่ทำการทดสอบ ไม่ทำให้ผู้สอบเกิดความเบื่อหน่าย ผู้สอบน่าจะมีความตั้งใจในการทำข้อสอบมากขึ้น และอาจทำข้อสอบด้วยวิธีการเดาน้อยลง ค่าฟังก์ชันสารสนเทศของแบบทดสอบจะมีความน่าเชื่อถือยิ่งขึ้น

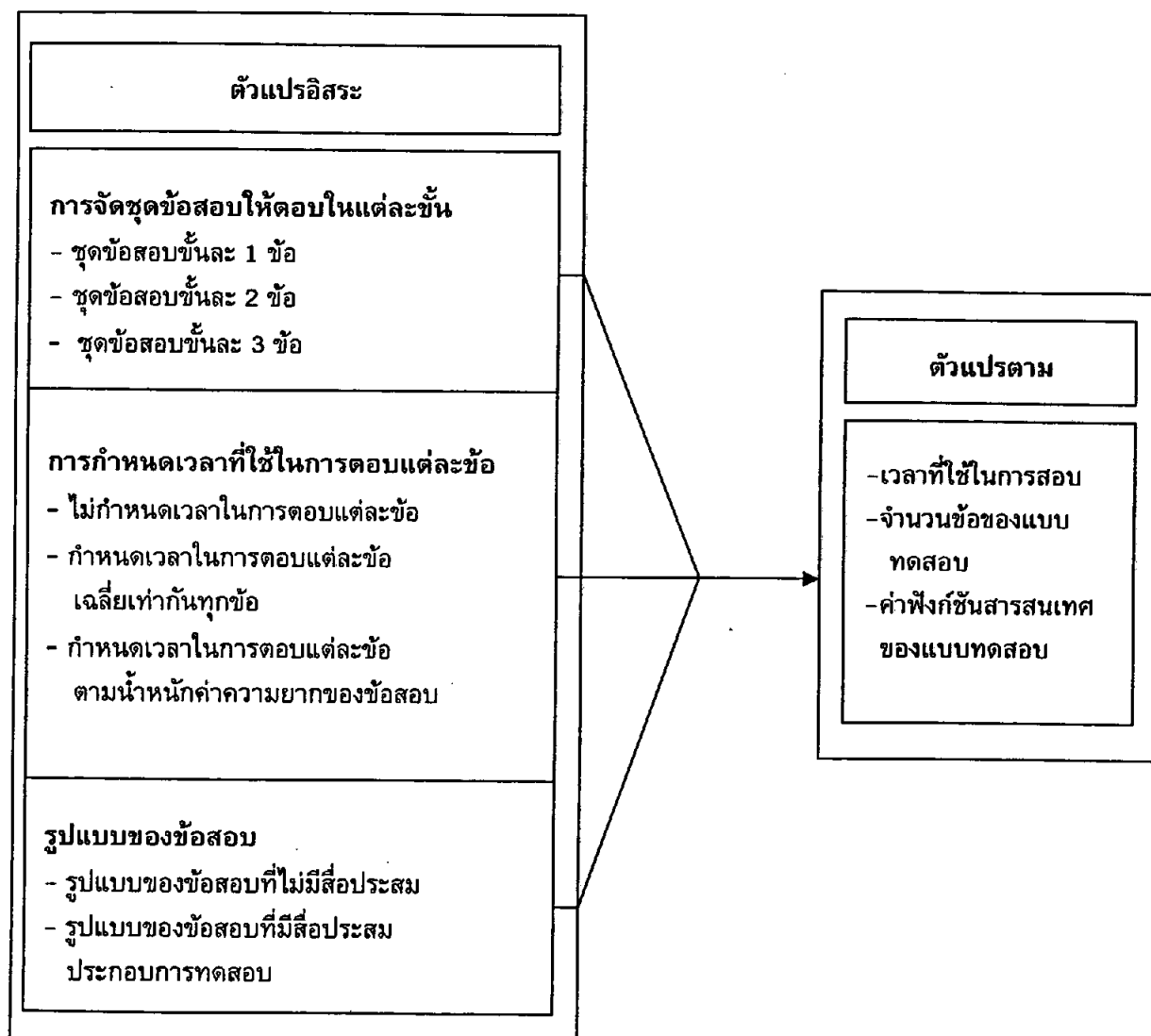
นอกจากอิทธิพลของผลหลัก (Main effect) แล้ว ผู้วิจัยยังไม่พบมีงานวิจัยของนักวิจัยท่านใดได้กล่าวถึงอิทธิพลร่วมกันระหว่างตัวแปรอิสระที่จะส่งผลกระทบต่อเวลาที่ใช้ในการทำแบบทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ดังนั้น ผู้วิจัยจึงตั้งสมมุติฐานไว้เพื่อศึกษาเรื่องนี้เพิ่มอีกเพียง 2 ข้อ ดังนี้

4. เมื่อจัดชุดข้อสอบให้ตอบในแต่ละชั้นและเมื่อกำหนดเวลาให้ตอบข้อสอบในแต่ละข้อต่างกันจะร่วมกันส่งผลทำให้ฟังก์ชันสารสนเทศของแบบทดสอบต่างกัน

5. เมื่อกำหนดเวลาให้ตอบข้อสอบในแต่ละข้อและการจัดรูปแบบของข้อสอบที่มีสื่อประสมต่างกันจะร่วมกันส่งผลทำให้ฟังก์ชันสารสนเทศของแบบทดสอบต่างกัน

กรอบความคิดในการวิจัย

ผู้วิจัยได้กำหนดกรอบความคิดในการวิจัยตามพื้นฐานความเชื่อจากที่กล่าวไว้ในสมมุติฐาน
ดังแสดงในภาพประกอบ 1



ภาพประกอบ 1 กรอบความคิดในการวิจัยเพื่อศึกษาเวลาที่ใช้ในการทดสอบ จำนวนข้อของแบบทดสอบ ฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

นียมศัพท์เฉพาะ

1. การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized Adaptive Testing) หมายถึง การทดสอบวัดความเข้าใจเรื่องเศษส่วน ในระดับชั้นประถมศึกษาปีที่ 6 ซึ่งดำเนินการสอบโดยใช้โปรแกรมคอมพิวเตอร์เพื่อคัดเลือกข้อสอบที่มีความยากง่ายเหมาะกับระดับความสามารถของผู้สอบ นำเสนอแก่ผู้สอบตามเงื่อนไขของตัวแปรที่กำหนด และข้อสอบที่นำเสนอแต่ละข้อจะขึ้นอยู่กับผลการตอบข้อสอบในข้อหรือในชุดของข้อสอบที่ผ่านมา

2. คลังข้อสอบ หมายถึง ระบบการจัดเก็บข้อสอบที่มีคุณภาพ ซึ่งได้จากการสร้างตามหลักเกณฑ์ที่ถูกต้อง ข้อสอบที่อยู่ในคลังเพื่อเก็บข้อมูลทำวิจัยในครั้งนี้เป็นข้อสอบวัดความเข้าใจเรื่องเศษส่วน ในระดับชั้นประถมศึกษาปีที่ 6 ของรังสรรค์ มณีเล็ก จำนวน 200 ข้อ และข้อสอบที่ผู้วิจัยสร้างขึ้นจำนวน 44 ข้อ รวมข้อสอบ 244 ข้อ เป็นข้อสอบแบบเลือกตอบ 4 ตัวเลือก วิเคราะห์ค่าพารามิเตอร์ของข้อสอบแบบ 3 พารามิเตอร์ โดยใช้ทฤษฎีการตอบสนองข้อสอบ

3. การจัดชุดข้อสอบให้ตอบในแต่ละชั้น หมายถึง การจัดชุดของจำนวนข้อสอบที่คอมพิวเตอร์จะคัดเลือกนำเสนอแก่ผู้สอบในแต่ละครั้งหลังจากทราบผลการตอบข้อสอบในข้อหรือชุดของข้อสอบก่อนหน้านั้น จำนวนข้อสอบที่นำเสนอแต่ละครั้งนั้นจะแตกต่างกันสามแบบคือ นำเสนอครั้งละ 1 ข้อ นำเสนอครั้งละ 2 ข้อ และนำเสนอครั้งละ 3 ข้อ แล้วจึงนำค่าความสามารถที่ได้จากการประมาณค่าความสามารถของผู้สอบครั้งหลังสุด ไปเป็นข้อมูลเพื่อคัดเลือกข้อสอบหรือชุดของข้อสอบที่เหมาะสมกับความสามารถของผู้สอบต่อไป โดยข้อสอบแต่ละชุดนั้นจะต้องมีค่าฟังก์ชันสารสนเทศของข้อสอบเท่ากัน หรือมีค่าใกล้เคียงกันมากที่สุด

4. เวลาที่ใช้ในการตอบแต่ละข้อ หมายถึง เวลาที่ผู้เข้าสอบใช้ในการตอบข้อสอบอย่างมากที่สุดในแต่ละข้อได้ไม่เกินเวลาตามที่ผู้วิจัยกำหนดไว้ในแต่ละข้อนั้น ซึ่งผู้วิจัยได้กำหนดเวลาไว้ 3 รูปแบบ คือไม่กำหนดเวลาในการตอบข้อสอบแต่ละข้อ กำหนดเวลาในการตอบข้อสอบแต่ละข้อเฉลี่ยเท่ากันทุกข้อ และกำหนดเวลาในการตอบข้อสอบแต่ละข้อตามน้ำหนักค่าความยากของข้อสอบ การกำหนดเวลาในการตอบข้อสอบตามน้ำหนักค่าความยากของข้อสอบแต่ละข้อนั้น ผู้วิจัยได้กำหนดเวลาโดยแบ่งค่าความยากของข้อสอบออกเป็นช่วง ๆ ตามระดับของค่าความยาก แล้วกำหนดเวลาให้ตอบข้อสอบแต่ละข้อ มาก น้อย ตามช่วงค่าความยากของข้อสอบนั้น ๆ การกำหนดเวลานี้ผู้วิจัยได้ทำการศึกษาเปรียบเทียบกับกลุ่มตัวอย่าง ดังปรากฏรายละเอียดในวิธีการดำเนินการวิจัย ในบทที่ 3

5. รูปแบบของข้อสอบที่มีสื่อประสมต่างกัน หมายถึง รูปแบบของข้อสอบที่รวมเอาสื่อประสมภาพกราฟฟิก ภาพวีว เสียง และคำถาม คำตอบเข้าไว้ด้วยกัน และรูปแบบของข้อสอบแบบปกติที่ไม่รวมเอาสื่อประสมดังกล่าวเข้าไว้ในการเสนอข้อสอบแต่ละข้อ การคัดเลือกสื่อประสมทั้งภาพกราฟฟิก สีและเสียง รวมเข้าไว้ด้วยกัน ผู้วิจัยได้อ้างอิงผลการวิจัยของ ศิริพงษ์ ภูพินนา (2546) ที่ได้พัฒนาคอมพิวเตอร์มัลติมีเดียแบบฝึกและคอมพิวเตอร์มัลติมีเดียเพื่อใช้ในการทดสอบวัดผลสัมฤทธิ์ของนักเรียนซึ่งคอมพิวเตอร์มัลติมีเดียสามารถทำงานอย่างมีประสิทธิภาพตามวัตถุประสงค์ ดังนั้นผู้วิจัยจึงพัฒนาคอมพิวเตอร์มัลติมีเดียเพื่อใช้ในการทดสอบวัดผลสัมฤทธิ์ของ ศิริพงษ์ ภูพินนา บางส่วนมาใช้ในการทดสอบแบบปรับเหมาะในครั้งนี้

6. ค่าฟังก์ชันสารสนเทศของข้อสอบ หมายถึง ค่าความแม่นยำในการประมาณค่าความสามารถของผู้สอบจากการตอบข้อสอบในข้อนั้น ๆ ซึ่งเป็นค่าที่คำนวณได้จากอัตราส่วนระหว่างกำลังสองค่าอนุพันธ์ของโอกาสในการตอบถูกต้องกับผลคูณของโอกาสในการตอบถูกต้องและตอบผิดในระดับความสามารถนั้น ๆ ของผู้สอบ เขียนเป็นสมการ ดังนี้ (Hambleton and Swaminathan. 1985 : 105 ; Baker, Frank B. 1992.)

$$I_i(\theta) = \frac{[P_i'(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

เมื่อ $I_i(\theta)$ หมายถึง สารสนเทศของข้อสอบ i ที่ระดับความสามารถ θ
 $P_i'(\theta)$ หมายถึง ค่าอนุพันธ์ของโอกาสในการตอบข้อสอบข้อที่ i ถูก
 $P_i(\theta)$ หมายถึง โอกาสในการตอบข้อสอบข้อที่ i ถูก ที่ระดับความสามารถ θ
 $Q_i(\theta)$ หมายถึง โอกาสในการตอบข้อสอบข้อที่ i ผิด ที่ระดับความสามารถ θ

7. ค่าฟังก์ชันสารสนเทศของแบบทดสอบ หมายถึง ค่าความแม่นยำในการประมาณค่าความสามารถจริง ของผู้สอบจากการตอบแบบทดสอบฉบับนั้น ๆ ซึ่งเป็นค่าที่คำนวณได้จากผลรวมของค่าฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อที่รวมกันเป็นค่าฟังก์ชันสารสนเทศของแบบทดสอบทั้งฉบับ เขียนเป็นสมการ ดังนี้ (Hambleton and Swaminathan. 1985 : 104 ; Baker. 1992.)

$$I(\theta) = \sum_{i=1}^n \frac{[P_i'(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

เมื่อ $I(\theta)$ หมายถึง ค่าฟังก์ชันสารสนเทศของแบบทดสอบ
 $I_i(\theta)$ หมายถึง สารสนเทศของข้อสอบ i ที่ระดับความสามารถ θ

8. ประสิทธิภาพของโปรแกรม หมายถึง ประสิทธิภาพการทำงานของโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ที่ผู้วิจัยพัฒนาขึ้น ประสิทธิภาพของโปรแกรมการทดสอบนี้พิจารณาจากผู้ใช้โปรแกรม 2 กลุ่มคือ

8.1 ผู้เชี่ยวชาญด้านโปรแกรมการทดสอบ โดยพิจารณาจากความเห็นของผู้เชี่ยวชาญว่าโปรแกรมสามารถทำงานได้อย่างมีประสิทธิภาพ เป็นไปตามเกณฑ์ของตัวแปรที่ผู้วิจัยศึกษาหรือไม่จากผู้เชี่ยวชาญจำนวน 3 ท่าน ที่สำเร็จการศึกษาในระดับปริญญาโทหรือปริญญาเอก ทางด้านการวัดผลการศึกษา ที่มีความสามารถในการใช้โปรแกรมคอมพิวเตอร์

8.2 ผู้ใช้โปรแกรมการทดสอบ ประกอบด้วย

8.2.1 ครูผู้สอน หรือผู้ดำเนินการสอบ ได้แก่ครูผู้สอน และศึกษานิเทศก์ ของสำนักงานเขตพื้นที่การศึกษาปทุมธานีเขตที่ 1 จำนวน 5 ท่าน เป็นผู้ทดลองใช้ และตรวจสอบประสิทธิภาพการทำงานของโปรแกรมด้านความสะดวกรวดเร็ว ความถูกต้องแม่นยำ และความทนทานต่อความผิดพลาดขณะใช้โปรแกรม

8.2.2 นักเรียนกลุ่มตัวอย่าง โดยทดลองใช้โปรแกรมการทดสอบกับนักเรียนกลุ่มตัวอย่างเพื่อพิจารณาประสิทธิภาพของโปรแกรมด้านความพึงพอใจในการสอบในปีจจัยที่เกี่ยวข้องกับรูปแบบการนำเสนอข้อสอบ เวลาที่ใช้ในการตอบข้อสอบแต่ละข้อ กระบวนการและบรรยากาศในการดำเนินการสอบ และการใช้สื่อประสมประกอบการทดสอบ

9. เวลาที่ใช้ในการทดสอบ หมายถึง เวลาทั้งหมดที่นักเรียนแต่ละคนใช้ในการทำแบบทดสอบ ซึ่งเวลาดังกล่าวจะเริ่มตั้งแต่ เริ่มทำข้อสอบข้อแรก จนกระทั่งถึงข้อสุดท้ายที่ยุติการสอบ

10. จำนวนข้อของแบบทดสอบ หมายถึง จำนวนข้อสอบทั้งหมดที่นักเรียนใช้ในการทำการทดสอบ ซึ่งเริ่มตั้งแต่ข้อแรก ถึงข้อสุดท้ายที่ยุติการทดสอบ

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การศึกษาเปรียบเทียบเวลาที่ใช้ในการทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ ด้วยการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ที่มีการจัดชุดข้อสอบให้ตอบในแต่ละชั้น การกำหนดเวลาที่ใช้ในการตอบแต่ละข้อ และรูปแบบของข้อสอบที่มีสื่อประสมแตกต่างกัน ผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับหัวข้อข้างต้นตามลำดับ ดังนี้

1. การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์
2. ทฤษฎีการตอบสนองข้อสอบ
3. งานวิจัยที่เกี่ยวข้อง

1. การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

1.1 ปรัชญาและลักษณะทั่วไปของการทดสอบ

การบริหารการทดสอบที่ใช้คอมพิวเตอร์ช่วยในการดำเนินการสอบเพื่อให้ผู้สอบได้รับข้อสอบที่เหมาะสมกับระดับความสามารถของตนเองนั้นเราเรียกการทดสอบลักษณะนี้ว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized Adaptive Testing; CAT) การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ผู้สอบแต่ละคนจะทำแบบทดสอบเฉพาะ (Tailored) ซึ่งออกแบบมาเฉพาะสำหรับระดับความสามารถของผู้สอบแต่ละคนข้อสอบที่ใช้ในการทดสอบของแต่ละบุคคลจะถูกเลือกมาจากธนาคารข้อสอบ ซึ่งธนาคารข้อสอบจะมีข้อสอบที่เหมาะสมกับระดับความสามารถของผู้สอบในระดับต่างๆ และมีเนื้อหาที่เกี่ยวข้องกับการทดสอบในแต่ละครั้งนั้น ๆ และข้อสอบแต่ละข้อจะกระจายออกไปตามเนื้อหาและระดับความยากของข้อสอบ โดยข้อสอบจะถูกตรวจสอบความยากง่ายโดยใช้ทฤษฎีการตอบสนองข้อสอบ และมีความรวดเร็วในการระบุความสามารถของผู้สอบที่มีความแม่นยำด้วยข้อสอบเพียงไม่กี่ข้อ (<http://www.ascp.org>)

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive Testing)

มีเป้าหมายเพื่อที่จะดำเนินการสอบด้วยการคัดเลือกข้อสอบให้เหมาะสม (Tailoring) กับความสามารถของผู้สอบแต่ละคน โดยอาศัยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory; IRT) มาใช้เป็นพื้นฐานในการคำนวณค่าสารสนเทศของข้อสอบ ซึ่งจะมีการคัดเลือกข้อสอบที่สามารถให้สารสนเทศอันเป็นประโยชน์เกี่ยวกับความสามารถของผู้สอบแต่ละคน ทำให้การทดสอบเป็นไปอย่างมีประสิทธิภาพ ใช้จำนวนข้อสอบที่น้อยลง ประหยัดเวลาของการสอบมีความคล่องตัว ผลที่ได้มีความคลาดเคลื่อนต่ำ และสามารถนำผลมาเปรียบเทียบกันได้อย่างมีความหมาย (Weiss, 1973 ; ศิริชัย กาญจนวาสี, 2538) ในระยะเริ่มต้นของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ มีการดำเนินการสอบโดยให้ผู้สอบเขียนตอบลงในกระดาษคำตอบ ซึ่งการคัดเลือกข้อสอบข้อต่อไปให้เหมาะ

กับความสามารถของผู้สอบ เป็นเรื่องยุ่งยาก ไม่สะดวก ต่อมาวิทยาการทางคอมพิวเตอร์พัฒนาก้าวหน้ามากขึ้น ดังนั้นการดำเนินการสอบแบบปรับเหมาะจึงได้นำคอมพิวเตอร์เข้ามาใช้ เนื่องจากคอมพิวเตอร์รักษาความลับของข้อสอบได้เป็นอย่างดี การบันทึกคำตอบ การแสดงผลการทดสอบ สามารถทำได้ง่ายรวดเร็วกว่าการใช้กระดาษคำตอบและการตรวจข้อสอบแบบเดิม การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบจะต้องมีคลังข้อสอบขนาดใหญ่ ข้อสอบจะต้องมีค่าพารามิเตอร์ประจำตัวของข้อสอบแต่ละข้อ ได้แก่ ค่าอำนาจจำแนก (a) ค่าความยากง่าย (b) ค่าการเดา (c) และค่าสารสนเทศของข้อสอบ (Item Information Function) เป็นพารามิเตอร์แสดงถึงคุณภาพของข้อสอบไว้เพื่อความสะดวกในการเลือกใช้ข้อสอบต่อไป (รังสรรค์ มณีเล็ก, 2540; อ้างอิงจาก Hambleton and Swaminathan, 1985.)

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (Computerized Adaptive Testing; CAT) เป็นการรวมพัฒนาการของศาสตร์ 2 ด้าน คือ พัฒนาการทางทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) และพัฒนาการทางเทคโนโลยีซิลิคอนชิป (Silicon Chip Technology) เข้าด้วยกัน เวียเนอร์กล่าวว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เป็นการสร้างแบบทดสอบโดยการนำคอมพิวเตอร์เข้ามาช่วยในการดำเนินการสอบเพื่อให้ได้แบบทดสอบที่เหมาะสมกับระดับความสามารถของผู้เข้าสอบ ซึ่งเป็นความพยายามของผู้สร้างแบบทดสอบที่จะเลียนแบบเทคนิคการสร้างแบบทดสอบโดยผู้เชี่ยวชาญด้านการสอบ กล่าวคือ ถ้าผู้เชี่ยวชาญด้านการสอบถามคำถามยากเกินไปสำหรับผู้สอบในคำถามแรก คำถามข้อถัดไปผู้เชี่ยวชาญจะพิจารณาคำถามที่ง่ายขึ้น (Straetmans and Eggen, 1998; อ้างอิงจาก Wainer, 1990)

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เป็นการทดสอบที่มีวัตถุประสงค์เพื่อปรับข้อสอบให้เหมาะกับความสามารถของผู้สอบโดยผู้สอบจะใช้แป้นพิมพ์ของคอมพิวเตอร์เพื่อตอบคำถามที่แสดงอยู่บนหน้าจอคอมพิวเตอร์หลังจากนั้นคอมพิวเตอร์จะประมาณความสามารถของผู้สอบ และทำการปรับข้อสอบให้เหมาะสมกับระดับความสามารถหรือผลสัมฤทธิ์ของผู้เข้าสอบแต่ละคน โดยหลักการแล้วโปรแกรมคอมพิวเตอร์จะเลือกข้อสอบที่ยากขึ้นหลังจากผู้สอบตอบคำถามถูกต้องหรือจะเลือกคำถามง่ายขึ้นถ้าผู้สอบตอบคำถามผิด การทดสอบแบบปรับเหมาะนี้ สามารถใช้วัดผลทางจิตวิทยาได้อย่างมีประสิทธิภาพมีความแม่นยำในการประมาณค่าความสามารถของผู้สอบ และรายงานผลการสอบได้ทันที โดยใช้ข้อสอบจำนวนน้อยลงและใช้เวลาในการสอบลดลงประมาณครึ่งหนึ่งของการทดสอบแบบมาตรฐานเดิมแต่บรรลุผลในระดับเดียวกัน ทั้งนี้เนื่องมาจากลักษณะของการทดสอบแบบปรับเหมาะจะเลือกเฉพาะข้อสอบที่มีความสอดคล้องเหมาะสมกับความสามารถของผู้สอบและจะตัดข้อสอบส่วนใหญ่ที่มีความยากง่ายไม่เหมาะสมกับผู้สอบออกไป (Meijer and Michael, 1999; ศิริชัย กาญจนวาสี, 2538)

1.2 ลักษณะการทำงานของ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

สตรีทแมนและเอจเจน (Straetmans and Eggen, 1998) กล่าวว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (CAT) เป็นการสร้างแบบทดสอบให้เหมาะสมกับระดับความสามารถของผู้เข้าสอบ ซึ่งเป็นความพยายามที่จะพัฒนาเทคนิคการทดสอบที่ชาญฉลาด

ซึ่งในช่วงปี 1980-1989 ได้มีการรวมพัฒนาการศาสตร์ 2 ด้าน นั่นคือ ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) และเทคโนโลยี Silicon Chip (Silicon Chip Technology) เข้าด้วยกันเพื่อใช้ประโยชน์ในการทดสอบ การนำคอมพิวเตอร์มาใช้ในการทดสอบให้ประโยชน์ดังนี้

1. คอมพิวเตอร์สามารถนำเสนอข้อสอบที่ตรงกับความสามารถของผู้สอบได้อย่างถูกต้องรวดเร็ว
2. คอมพิวเตอร์สามารถนำเสนอวิธีการทดสอบที่รวมเอาภาพกราฟฟิก เสียง วีดิโอ เคลื่อนไหว และเนื้อหาสาระเข้าไว้ด้วยกัน จะคล้ายคลึงกับวิถีชีวิตจริง
3. คอมพิวเตอร์จะเป็นผู้แสดงผลคะแนนการสอบของผู้เข้าสอบและรายงานผลขั้นสุดท้ายทันทีที่สิ้นสุดการทดสอบจะประหยัดทั้งเวลาและแรงงาน
4. คอมพิวเตอร์จะทำการจัดชุดแบบทดสอบให้กับผู้เข้าสอบจึงเป็นการแบ่งเบาภาระแก่ครูผู้สอน
5. คอมพิวเตอร์สามารถประมวลผลที่มีความรวดเร็ว ถูกต้อง และมีประสิทธิภาพมาก
6. การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์สามารถประเมินผลคุณลักษณะทางจิตวิทยาได้เที่ยงตรง น่าเชื่อถือโดยใช้จำนวนข้อสอบและเวลาตอบน้อยกว่าเมื่อเทียบกับการทดสอบแบบมาตรฐานเดิม

ส่วนขั้นตอนของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (CAT) มีดังนี้

1. คอมพิวเตอร์ ทำการคัดเลือกข้อสอบจากธนาคารข้อสอบ
2. คอมพิวเตอร์นำเสนอข้อสอบให้นักเรียนตอบบนหน้าจอคอมพิวเตอร์
3. นักเรียนตอบคำถามโดยการพิมพ์หรือเลือกคำตอบบนหน้าจคอมพิวเตอร์
4. คอมพิวเตอร์ทำการประเมินคำตอบว่าตอบถูกหรือตอบผิด
5. ถ้านักเรียนตอบผิด คอมพิวเตอร์จะคัดเลือกข้อสอบข้อต่อไปง่ายขึ้น ถ้าตอบถูก ข้อสอบก็จะยากขึ้น

6. การทดสอบจะสิ้นสุดลงเมื่อเป็นไปตามกฎเกณฑ์ของการยุติการสอบ

กระบวนการของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จะต้องมียังประกอบสำคัญ ดังนี้ (Straetmans and Eggen. 1998)

1. ธนาคารข้อสอบ (A Calibrated Item Bank)

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (CAT) มีความจำเป็นอย่างยิ่งที่จะต้องมียังธนาคารข้อสอบ ซึ่งเป็นระบบการจัดข้อสอบที่มีคุณภาพ ธนาคารข้อสอบจะเป็นที่รวมของข้อสอบที่สามารถค้นหาได้ง่าย (Accessible) ข้อสอบจะถูกแยกประเภทหรือจัดระบบตามลักษณะที่จะนำไปใช้งานได้ง่ายในการทดสอบแต่ละครั้ง ข้อสอบในธนาคารข้อสอบโดยปกติแล้วจะจัดแยกประเภท ตามเนื้อหา ลักษณะข้อคำถาม ลักษณะการปฏิบัติการเชื่อมโยงกับข้อสอบข้ออื่น ๆ หรือจัดแยกตามวัสดุสื่อ ผู้เขียน ประวัติการทดสอบ และลักษณะทางจิตวิทยา รวมทั้งระดับความยากของข้อสอบ สำหรับข้อสอบที่รวบรวมไว้จะมีสารสนเทศเกี่ยวกับระดับความยากของข้อสอบ ซึ่งเป็นเรื่องที่สำคัญมากในทฤษฎีมาตรฐานเดิม (Classical Test Theory) คำดัชนีที่ใช้อธิบายความยาก

ของข้อสอบคือค่า p ซึ่งเป็นสัดส่วนของกลุ่มผู้เข้าสอบกับผู้ที่ตอบข้อสอบนี้ถูก ค่า p ต่ำ หมายถึง ข้อสอบยาก และค่า p สูง หมายถึงข้อสอบง่าย อย่างไรก็ตาม ค่า p จะขึ้นอยู่กับความสามารถของกลุ่มที่สอบข้อสอบนั้น ค่า p ของกลุ่มข้อสอบที่ใช้กับนักเรียนกลุ่มที่มีความสามารถสูง จะแตกต่างกันอย่างมากกับค่า p ของข้อสอบชุดเดียวกันที่ใช้สอบกับนักเรียนกลุ่มที่มีความสามารถต่ำด้วยเหตุนี้เอง ทำให้ค่า p ตามทฤษฎีมาตรฐานเดิม ไม่เหมาะสมสำหรับการอธิบายระดับความยากของข้อสอบในธนาคารข้อสอบ ในขณะที่เดียวกันก็ไม่สามารถรับรองได้ว่า ระดับความสามารถของกลุ่มที่ทำแบบทดสอบแต่ละชุดนั้นจะสอดคล้องกับระดับความสามารถของกลุ่มตัวอย่างที่น่าจะนำมาใช้คำนวณค่า p ดังนั้นทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) สามารถแก้ปัญหานี้ได้ โดยการประมาณค่าพารามิเตอร์ความยาก ซึ่งไม่ขึ้นอยู่กับระดับความสามารถของกลุ่มที่นำมาพัฒนา (Weiss and Others, 1990) ในทฤษฎีการตอบสนองข้อสอบ ค่าพารามิเตอร์ความยากของข้อสอบแต่ละชุดนั้นเป็นอิสระในแต่ละกลุ่ม โอกาสของการตอบถูกแตกต่างกันไปตามระดับความสามารถของแต่ละบุคคล นั่นคือ โอกาสของผู้ที่มีความสามารถสูงในการตอบข้อสอบได้ถูกต้องจะสูงกว่า ผู้ที่มีความสามารถต่ำ

2. ขั้นตอนในการทดสอบ (Testing Algorithm)

กระบวนการในการทดสอบเป็นหัวใจสำคัญของการทดสอบแบบปรับเหมาะ ซึ่งจะต้องทราบว่าจะเริ่มต้น ดำเนินการสอบและสิ้นสุดการทดสอบอย่างไร

2.1 การเริ่มต้นการทดสอบ (Starting)

การทดสอบแบบปรับเหมาะ จะให้ความสำคัญกับค่าความยากของข้อสอบข้อต่อไปที่จะนำมาเสนอแก่ผู้สอบ การพิจารณาความยากของข้อสอบจะพิจารณาจากการประมาณค่าความสามารถของผู้เข้าสอบในขณะนั้น เมื่อผู้เข้าสอบตอบข้อสอบในแต่ละครั้งความสามารถของผู้สอบจะถูกประมาณค่าใหม่ทุกครั้งเสมอ เมื่อเริ่มต้นการทดสอบมักจะไม่มีความรู้เกี่ยวกับความสามารถของผู้เข้าสอบ ดังนั้นการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ จึงคัดเลือกข้อสอบที่มีความยากปานกลางเสนอแก่ผู้เข้าสอบเป็นข้อแรกในการเริ่มต้นการสอบ หรืออาจคัดเลือกข้อสอบที่ง่ายเป็นข้อสอบในการเริ่มต้นสอบเพื่อลดความวิตกกังวลและสร้างแรงจูงใจในการสอบของผู้เข้าสอบ

2.2 การดำเนินการสอบ (Continuing)

การดำเนินการสอบด้วยการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จะต้องคำนึงถึงประสิทธิภาพของการทดสอบ กล่าวคือ การทดสอบจะต้องใช้ข้อสอบน้อย ข้อและลดเวลาที่ใช้ในการทดสอบลง เพื่อลดความเครียดและความเหนื่อยล้าจากการทดสอบ แต่ผลการทดสอบสามารถประมาณค่าความสามารถของผู้สอบได้ตรงกับระดับความสามารถที่แท้จริง ดังนั้นใน ขั้นตอนของการดำเนินการสอบ จะต้องกระทำอย่างรัดกุม กระบวนการคัดเลือกข้อสอบจะต้องสามารถคัดเลือกข้อสอบที่มีสารสนเทศ โดยเฉพาะค่าความยากจะต้องสอดคล้องกับค่าความสามารถของผู้สอบที่ถูกประมาณค่าครั้งสุดท้ายมากที่สุด การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์สามารถคัดเลือกข้อสอบเสนอแก่ผู้สอบได้ถูกต้องรวดเร็ว การทดสอบแบบปรับเหมาะจึงเป็นกระบวนการทดสอบที่มีประสิทธิภาพสูง

2.3 การสิ้นสุดการทดสอบ (Stopping)

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบจะสิ้นสุดลงเมื่อผลการสอบที่ผ่านมาเป็นไปตามเกณฑ์ยุติการทดสอบ ซึ่งเกณฑ์การยุติการทดสอบนั้นอาจจะกำหนดโดยใช้เกณฑ์จำนวนข้อสอบ เกณฑ์เวลาที่ใช้ในการสอบ เกณฑ์คะแนนจุดตัด (Cut off score) เพื่อแยกระดับความสามารถของผู้สอบออกเป็นกลุ่มรอบรู้และกลุ่มไม่รอบรู้ และเกณฑ์ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบเป็นไปตามเกณฑ์ยุติการทดสอบที่ตั้งไว้ การทดสอบของแต่ละคนก็จะสิ้นสุดลง

1.3 ความแตกต่างระหว่างการทดสอบแบบเดิมที่ใช้กระดาษเขียนตอบและการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

สตรีทแมนและเอจเจน (Straetmans and Eggen, 1998) กล่าวว่า แบบทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แตกต่างจากแบบทดสอบที่ใช้กระดาษเขียนตอบ พอสรุปได้ดังตาราง 1

ตาราง 1 ความแตกต่างระหว่างแบบทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับแบบทดสอบที่ใช้กระดาษเขียนตอบ

ประเด็น	แบบทดสอบที่ใช้กระดาษเขียนตอบ	แบบทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
ความยาก	ปรับความสามารถของผู้เข้าสอบให้อยู่ในระดับปานกลาง	ปรับความสามารถของผู้เข้าสอบเป็นรายบุคคล
ความยาวของแบบทดสอบ	ระบุเจาะจงและข้อสอบมีจำนวนมากข้อ	ไม่ระบุเจาะจงและข้อสอบจำนวนน้อยข้อ
ความปลอดภัย	มีปัญหาเรื่องความปลอดภัยเมื่อผู้เข้าสอบไม่ได้สอบในเวลาเดียวกัน	ป้องกันความปลอดภัยได้ดีกว่า เพราะผู้เข้าสอบแต่ละคนจะใช้แบบทดสอบที่ต่างกัน
ความสมดุลย์ของเนื้อหา	ไม่มีปัญหา	ไม่ได้้อยู่ภายใต้การควบคุมโดยปราศจากการวัดผลเป็นพิเศษ
วิธีการควบคุม	ผู้เข้าสอบอยู่ในการควบคุม	คอมพิวเตอร์อยู่ในการควบคุม
การบริหารจัดการ	ทำงานหนัก	กระจายการทำงาน

1.3.1 ความยาก (Difficulty)

แบบทดสอบที่ใช้กระดาษเขียนตอบโดยเฉพาะข้อสอบแบบเลือกตอบ มีการใช้กันอย่างแพร่หลายในปัจจุบัน นักเรียนทุกคนจะต้องทำการทดสอบในเวลาเดียวกันด้วยข้อสอบชุดเดียวกัน ระดับความยากของแบบทดสอบจะถูกปรับตามค่าเฉลี่ยความสามารถของกลุ่มผู้เข้าสอบ ซึ่งมีผลทำให้ผู้เข้าสอบที่มีความสามารถสูงหรือต่ำกว่าค่าเฉลี่ยเกิดความสะเพร่าหรือมีความเครียดเป็นอย่างมาก ผู้เข้าสอบจะเผชิญกับข้อสอบที่ยากหรือง่ายเกินไป ตามหลักจิตวิทยาแล้วความถูกต้อง

ของการวัดผลจากการสอบจะลดลง ส่วนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ผู้เข้าสอบแต่ละคนจะได้รับแบบทดสอบที่แตกต่างกัน ซึ่งเป็นแบบทดสอบที่ไม่ยากหรือง่ายเกินไป ดังนั้นแบบทดสอบจึงท้าทายความรู้สึกของผู้เข้าสอบ ผู้เข้าสอบไม่วิตกกังวล ความถูกต้องของการวัดของผู้เข้าสอบแต่ละคนจะมีค่าสูง

1.3.2 ความยาวของแบบทดสอบ (Test length)

การทดสอบที่ใช้กระดาษเขียนตอบจะกำหนดความยาวของแบบทดสอบตายตัว ความยาวของแบบทดสอบนั้นจะถูกกำหนดจากกลุ่มผู้เข้าสอบทั้งหมด ผู้เข้าสอบทุกคนจะต้องใช้ข้อสอบชุดเดียวกัน ถ้าความสามารถของกลุ่มผู้เข้าสอบกระจายเป็นโค้งปกติแล้ว แบบทดสอบที่ดีจะต้องมีความยากกระจายไปให้ครอบคลุมความสามารถของกลุ่มผู้เข้าสอบด้วย ดังนั้นจึงมีข้อสอบที่ไม่เหมาะสมกับความสามารถของผู้เข้าสอบแต่ละคนรวมอยู่ด้วย ส่วนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ผู้เข้าสอบแต่ละคนจะได้รับแบบทดสอบที่แตกต่างกัน ซึ่งจะถูกปรับตามความสามารถของผู้เข้าสอบหลังจากทำ ข้อสอบในแต่ละข้อแล้วความสามารถของผู้เข้าสอบแต่ละคนจะถูกประมาณค่าใหม่ทุกครั้ง และเมื่อค่าความสามารถของผู้สอบจากการประมาณครั้งหลังสุดเป็นไปตามเกณฑ์ยุติการสอบ การทดสอบก็จะสิ้นสุดลง

1.3.3 ความปลอดภัย (Security)

การทดสอบที่ใช้กระดาษเขียนตอบ ผู้เข้าสอบจะต้องทำแบบทดสอบชุดเดียวกันในเวลาเดียวกันด้วยจึงจะทำให้ข้อสอบไม่รั่ว ถ้าทำการทดสอบในเวลาที่แตกต่างกัน ความเชื่อมั่นของการทดสอบจะลดลง แต่การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สามารถแก้ปัญหานี้ได้ โดยการให้ผู้เข้าสอบแต่ละคนทดสอบโดยใช้ข้อสอบที่ต่างกัน ในเวลาที่ต่างกันได้ด้วย แต่ผลการสอบยังคงเปรียบเทียบกันได้

1.3.4 ความสมดุลย์ของเนื้อหา (Balancing of Content)

การสร้างแบบทดสอบที่ใช้กระดาษเขียนตอบจะสร้างตามแผนการออกข้อสอบ ทั้งนี้เพื่อให้เกิดความมั่นใจว่า ข้อสอบที่เขียนออกมานั้นเป็นตัวแทนของความรู้และทักษะที่กำหนดในจุดประสงค์ของหลักสูตร แต่ข้อสอบในแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จะถูกคัดเลือกโดยยุทธศาสตร์การคัดเลือกจากสารสนเทศสูงสุด เนื้อหาในแบบทดสอบจึงถูกกำหนด โดยการพิจารณาตามหลักการทางจิตวิทยา ดังนั้นผู้เข้าสอบจะได้รับข้อสอบที่มีความสมดุลย์ ถึงแม้ว่าการทดสอบจะสิ้นสุดลงหลังจากที่ทำแบบทดสอบที่มีจำนวนข้อสอบเพียงไม่กี่ข้อก็ตาม

1.3.5 วิธีการควบคุม (Locus of Control)

การทดสอบที่ใช้กระดาษเขียนตอบ ผู้ใช้ข้อสอบจะต้องตัดสินใจว่า คำตอบที่ได้จากแบบทดสอบจะนำไปทำอะไร มีความเป็นไปได้ที่จะมีคำตอบที่แตกต่างกันในคำถามเดียวกัน จึงอาจมีการทดสอบโดยให้เปิดหนังสือตอบและมีการทบทวนคำตอบต่าง ๆ ซึ่งปกติแล้ว ผู้บริหารการสอบจะเป็นผู้อนุญาตให้ทำในลักษณะใดลักษณะหนึ่ง ส่วนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ คอมพิวเตอร์จะเป็นผู้ตัดสินใจว่า คำถามข้อต่อไปควรจะเป็นอะไร การควบคุมกระบวนการทดสอบจะรัดกุมมากยิ่งขึ้น

1.3.6 การบริหารจัดการ (Logistics)

การทดสอบที่ใช้กระดาษเขียนตอบจะต้องมีครูผู้ควบคุมการสอบหรือผู้ที่เกี่ยวข้องกับสอบจำนวนมาก แบบทดสอบจะนำมาจากธนาคารข้อสอบและถ่ายทำสำเนา ต้องมีการใช้ห้องสอบและจ้างผู้ควบคุมการสอบ หลังจากทำการทดสอบจะมีการเก็บรวบรวมกระดาษคำตอบ และ ข้อสอบพร้อมทั้งทำเครื่องหมายต่าง ๆ ไว้ ส่วนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ คอมพิวเตอร์จะตั้งงานต่าง ๆ เหล่านี้มาจากครูและลดอุปสรรคในการทดสอบลง

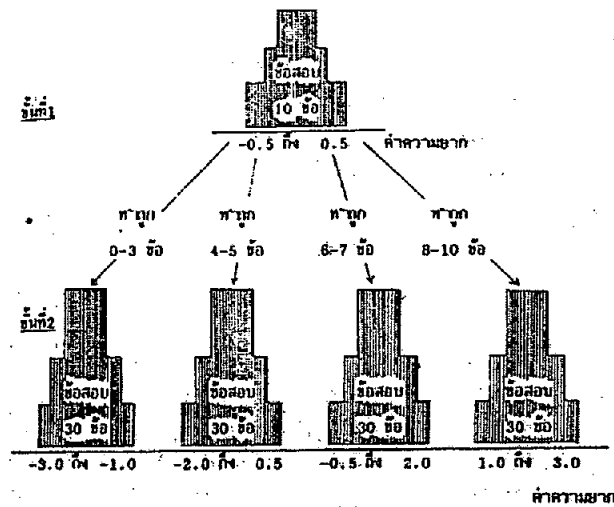
1.4 ประเภทของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ

ไวส์ แฮมเบิลตันและสวามินาทาน (Weiss, 1974 ; Hambleton and Swaminathan, 1985) จำแนกการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบออกเป็น 2 ประเภทใหญ่ ๆ คือ

- 1) การทดสอบสองขั้นตอน (two-stage strategies)
- 2) การทดสอบหลายขั้นตอน (multi-stage strategies)
 - (1) แบบแยกทางคงที่ (fixed branching model)
 - (1.1) รูปพีระมิตขนาดขั้นคงที่ (constant step size pyramid)
 - (1.2) รูปพีระมิตขนาดขั้นแปรผัน (variable step size pyramid)
 - (1.3) รูปพีระมิตข้างตัด (truncated pyramid)
 - (1.4) รูปพีระมิตที่มีหลายข้อในแต่ละขั้น (multiple item pyramid)
 - (1.5) รูปพีระมิตที่ให้น้ำหนักแก่ตัวเลือก (differential response option branching)
 - (1.6) แบบทดสอบเฟล็กซ์ซิเลเวล (flexilevel test)
 - (2) แบบแยกทางแปรผัน (variable branching model)

1.4.1 การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบสองขั้นตอน

การทดสอบสองขั้นตอน เป็นการทดสอบโดยใช้ข้อสอบ 2 ชุด การดำเนินการสอบเริ่มด้วยให้ ผู้สอบทำแบบทดสอบเพื่อแยกทาง (routing test) ผลจากการทดสอบนี้จะเห็นตัวชี้วัดว่าผู้สอบจะได้รับแบบทดสอบชุดใด ในขั้นที่ 2 ซึ่งเรียกแบบทดสอบในขั้นนี้ว่า แบบทดสอบเพื่อวัดผล (Measurement test) โดยหลักการแล้วผู้สอบที่ได้คะแนนน้อยจากการทดสอบแยกทางถือว่าเป็นผู้ที่มีความสามารถต่ำจึงจะได้รับข้อสอบเพื่อวัดผลชุดที่ง่าย ส่วนผู้ที่ได้คะแนนจากการทดสอบเพื่อแยกทางมากหรือมีความสามารถสูง จะได้รับข้อสอบเพื่อวัดผลชุดที่ยากขึ้น ดังภาพประกอบ 2



ภาพประกอบ 2 โครงสร้างของการทดสอบเฉพาะบุคคลแบบสองขั้นตอน

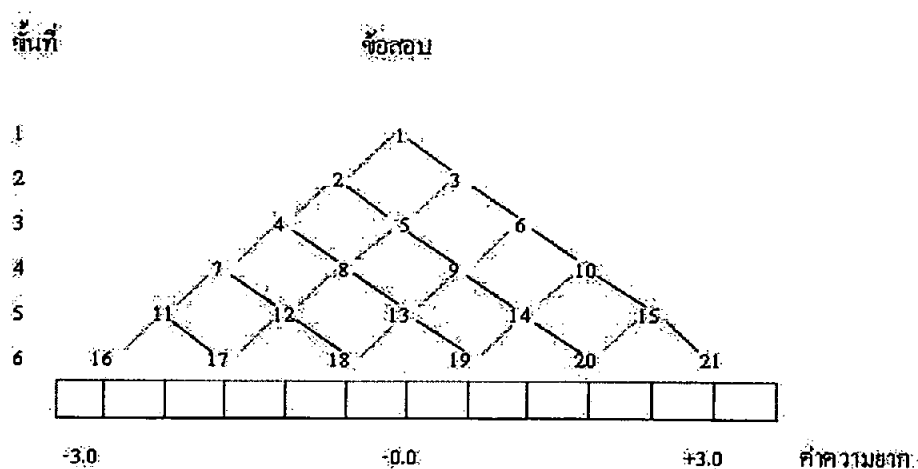
จากภาพประกอบ 2 จะเห็นว่า มีข้อสอบเพื่อแยกทางในขั้นที่ 1 จำนวน 10 ข้อ ถ้าผู้สอบตอบคำถามถูก 0 – 3 ข้อ จะได้ทำข้อสอบเพื่อวัดผลในขั้นที่ 2 ในชุดที่ง่ายที่สุดจำนวน 30 ข้อ ส่วนผู้ที่มีความสามารถสูงขึ้นก็จะได้รับข้อสอบในชุดที่ยากขึ้นตามลำดับ

1.4.2 การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบหลายขั้นตอน

การทดสอบหลายขั้นตอน เป็นการทดสอบที่มีโครงสร้างของแบบทดสอบเป็นรูปพีระมิดโดยใช้ค่าความยากง่ายของข้อสอบเป็นเกณฑ์ ในการจัดเรียงข้อสอบที่มีความยากง่ายปานกลางจะอยู่จุดยอดของพีระมิด การดำเนินการสอบจะเริ่มด้วยให้ผู้สอบตอบข้อสอบข้อแรกที่มีความยากง่ายปานกลาง ถ้าผู้สอบตอบคำถามถูกข้อสอบข้อต่อไปจะยากขึ้น และถ้าผู้สอบตอบผิดข้อสอบข้อต่อไปก็จะง่ายลง การทดสอบจะดำเนินการเช่นนี้ไปเรื่อย ๆ จนกระทั่งถึงเกณฑ์ยุติการสอบตามที่กำหนดไว้การสอบจึงสิ้นสุดลง การทดสอบหลายขั้นตอนสามารถแยกเป็นประเภทย่อย ๆ ได้ดังนี้

(1) แบบแยกคงที่

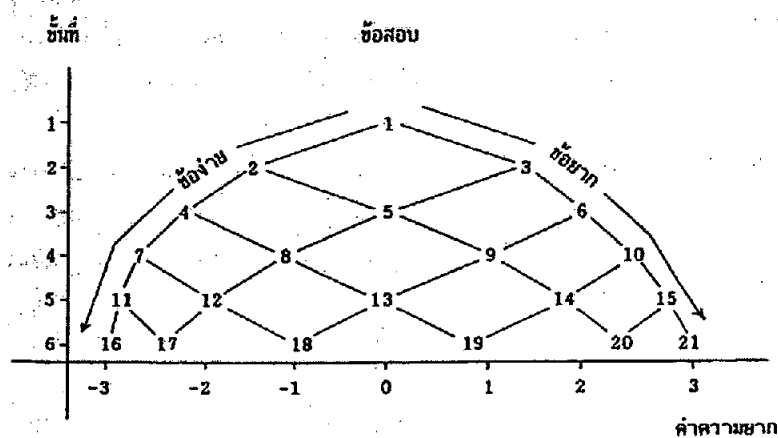
(1.1) รูปพีระมิดขนาดขั้นคงที่เป็นการจัดชุดแบบทดสอบเป็นรูปสามเหลี่ยม โดยมีจำนวนข้อสอบในแต่ละขั้นเท่ากับลำดับที่ของขั้น ผู้สอบจะต้องตอบข้อสอบขั้นละ 1 ข้อ โดยเริ่มตอบข้อ 1 ก่อน ถ้าตอบถูกก็จะไปตอบข้อที่ 3 ซึ่งยากกว่าข้อ 1 แต่ถ้าตอบผิดก็จะไปตอบข้อ 2 ซึ่งง่ายกว่าข้อ 1 ดังภาพประกอบ 3



ภาพประกอบ 3 โครงสร้างของการทดสอบเฉพาะบุคคลรูปแบบพีระมิดขนาดชั้นคงที่

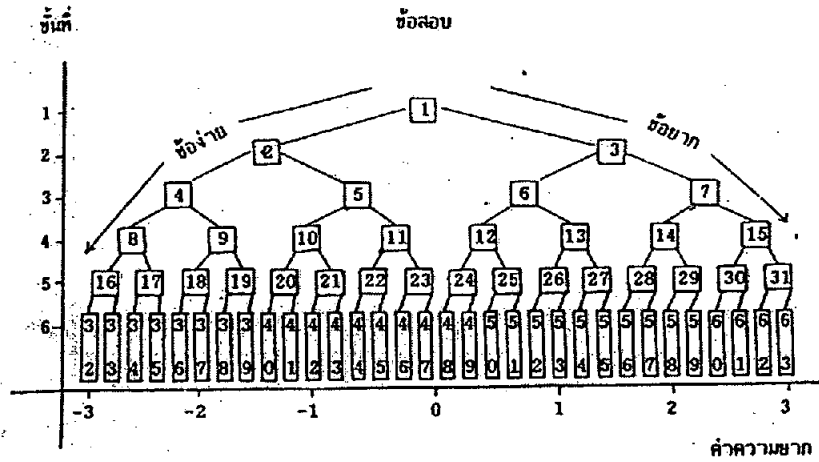
(1.2) รูปพีระมิดขนาดชั้นแปรผัน

การจัดชุดแบบทดสอบลักษณะนี้ ได้กำหนดช่วงห่างของค่าความยากง่ายของข้อสอบในชั้นเดียวกันไม่เท่ากัน โดยข้อสอบที่มีความยากง่ายปานกลางจะมีช่วงห่างของค่าความยากง่ายมากกว่าข้อสอบที่ยากหรือข้อสอบที่ง่าย ดังภาพประกอบ 4



ภาพประกอบ 4 โครงสร้างของแบบทดสอบรูปพีระมิดขนาดชั้นแปรผัน

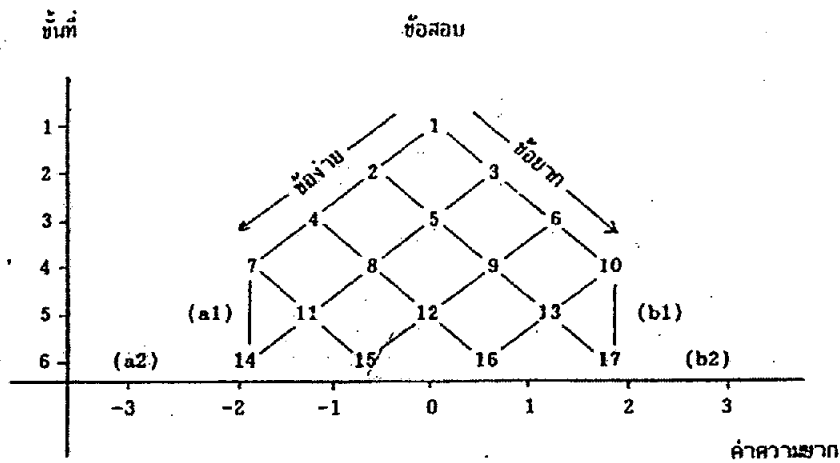
ต่อมาได้มีการพัฒนาโครงสร้างของแบบทดสอบรูปพีระมิดขนาดชั้นแปรผันใหม่ โดยมีการเพิ่มจำนวนข้อสอบในแต่ละชั้นให้มีจำนวนเป็น 2 เท่าของจำนวนข้อสอบในชั้นที่ผ่านมา รูปแบบใหม่นี้มีชื่อว่า โรบินส์มอนโร (Robbins Monro) ซึ่งมีโครงสร้างดังภาพประกอบ 5



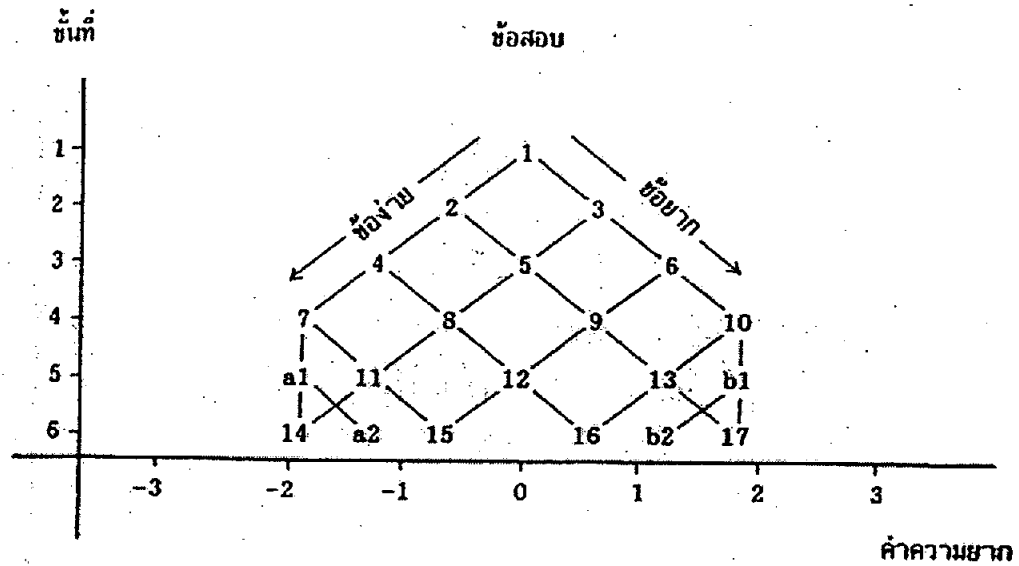
ภาพประกอบ 5 โครงสร้างของแบบทดสอบรูปพีระมิดแบบโรบินส์มอนโร

(1.3) รูปพีระมิดแบบข้างตัด

เป็นรูปแบบที่มุสลีโอ (Weiss, 1974; Citing Mussio, 1973) ที่สร้างขึ้นเพื่อลดจำนวนข้อสอบจากโครงสร้างแบบทดสอบรูปพีระมิดแบบขนาดชั้นคงที่ ซึ่งจำแนกเป็นแบบสกัดการสะท้อนกลับ (reflecting barrier) และรักษาการสะท้อนกลับ (retaining barrier) ดังภาพประกอบ 6 และ 7 ตามลำดับ



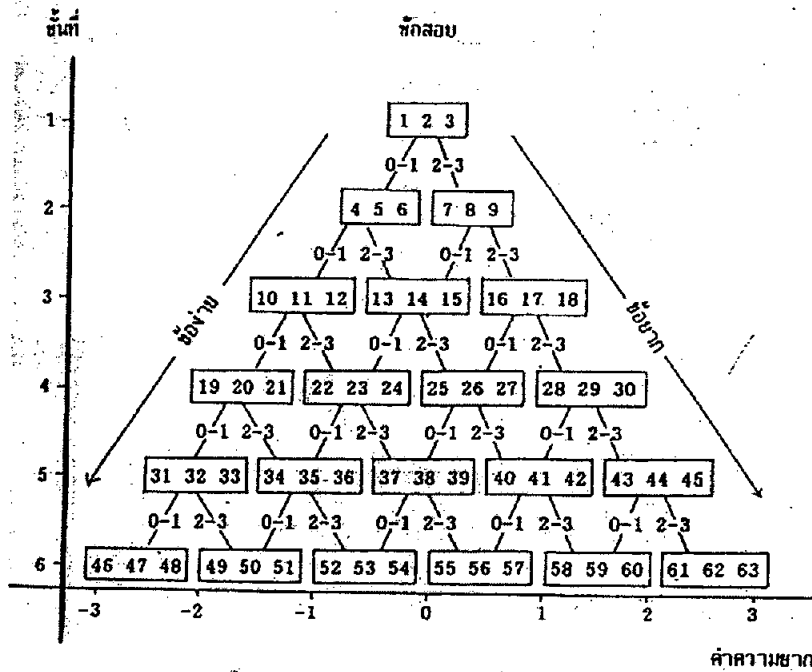
ภาพประกอบ 6 โครงสร้างของแบบทดสอบรูปพีระมิดแบบสกัดการสะท้อนกลับ



ภาพประกอบ 7 โครงสร้างของแบบทดสอบรูปพีระมิดแบบรักษาการสะท้อนกลับ

(1.4) รูปพีระมิดแบบมีหลายข้อในแต่ละชั้น

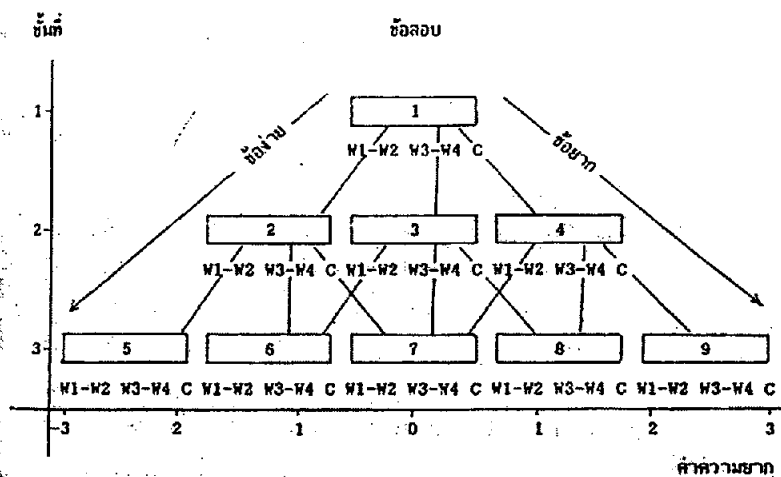
รูปแบบนี้กำหนดให้มีข้อสอบหลายข้อในแต่ละชั้น ทั้งนี้เพื่อให้การแยกทางมีความถูกต้องมากยิ่งขึ้น ซึ่งมีโครงสร้างของการจัดข้อสอบดังภาพประกอบ 8



ภาพประกอบ 8 โครงสร้างของแบบทดสอบรูปพีระมิดแบบมีหลายข้อในแต่ละชั้น

(1.5) รูปพีระมิดแบบให้น้ำหนักแก่ตัวเลือก

รูปแบบนี้คล้ายกับรูปแบบที่มีหลายข้อในแต่ละชั้น ต่างกันตรงที่จะมีการกำหนดน้ำหนักให้กับตัวเลือกแต่ละข้อ ดังภาพประกอบ 9



ภาพประกอบ 9 โครงสร้างของแบบทดสอบรูปพีระมิดแบบให้น้ำหนักแก่ตัวเลือกของข้อสอบเพื่อแยกทาง

(1.6) การทดสอบเฟล็กซิเบิล

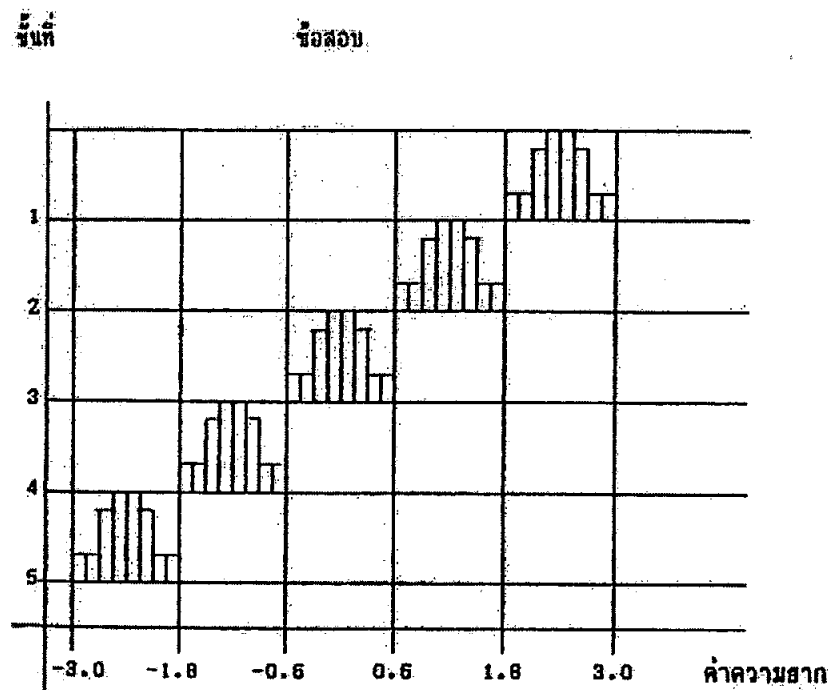
ลอร์ด (Lord, 1980 : 114-117) เป็นผู้นำเสนอการทดสอบลักษณะนี้ขึ้นมา การจัดเรียงข้อสอบจะเรียงจากง่ายไปหายาก ข้อที่มีความยากง่ายปานกลางจะถูกนำเสนอแก่ผู้สอบก่อน เช่น ถ้าข้อสอบมีทั้งหมด 75 ข้อ โครงสร้างของการจัดข้อสอบลักษณะนี้สามารถแสดงได้ดังตาราง 2

ตาราง 2 โครงสร้างของการจัดแบบทดสอบเฟล็กซิเบิล

ข้อ 0. (มีความยากอยู่กึ่งกลางระหว่างข้อสอบกลุ่มง่ายกับกลุ่มยาก)	
(ข้อสอบกลุ่มง่าย)	(ข้อสอบกลุ่มยาก)
(พิมพ์สีแดง)	(พิมพ์สีน้ำเงิน)
ข้อ 1. (ข้อที่ยากที่สุดในกลุ่มนี้)	ข้อ 1. (ข้อที่ง่ายที่สุดในกลุ่มนี้)
ข้อ 2.	ข้อ 2.
ข้อ 3.	ข้อ 3.
:	:
ข้อ 37. (ข้อที่ง่ายที่สุดในกลุ่มนี้)	ข้อ 37. (ข้อที่ยากที่สุดในกลุ่มนี้)

กระดาษคำตอบของการทดสอบลักษณะนี้จะทำให้ผู้สอบทราบได้ทันทีหลังจากการตอบแต่ละข้อว่า ถ้าตอบผิดจุดแดงจะขึ้นมา (ข้อสอบง่ายกว่าเดิม) ถ้าตอบถูกต้องจุดน้ำเงินจะขึ้นมา (ข้อสอบยากกว่าเดิม) ซึ่งอาจใช้วิธีการเจาะหรือแกะกระดาษที่หับตัวเลือกแต่ละข้อ ผู้สอบจะต้องตอบข้อสอบครึ่งหนึ่งของข้อสอบทั้งหมด การทดสอบจึงจะยุติ วิธีการง่าย ๆ เพื่อให้ผู้สอบสามารถยุติการสอบได้ด้วยตนเอง ก็คือ จัดเรียงข้อสอบด้านยากใหม่โดยกลับหัวจากง่ายลงไปหายากที่สุด เป็นจากยากที่สุดลงไปหาง่าย (เหมือนการจัดข้อสอบด้านง่าย) การตอบก็จะมีลักษณะเหมือนเดิม คือตอบข้อสอบที่ยากง่ายปานกลางก่อน ถ้าตอบผิดก็จะไปตอบข้อที่ง่ายกว่าเดิม ถ้าตอบถูกก็จะไปตอบข้อที่ยากกว่าเดิม ตอบไปจนกระทั่งพบว่าในแต่ละแถวของข้อสอบ (มี 2 ข้อ) ถูกตอบครบทุกแถว แถวละ 1 ข้อ จึงยุติการสอบ

ต่อมาได้มีการปรับปรุงโครงสร้างของการทดสอบลักษณะนี้ โดยแบ่งความยากง่ายของข้อสอบออกเป็นช่วง ๆ ในแต่ละช่วงความยากง่ายจะมีข้อสอบหลายข้อ การจัดแบบทดสอบลักษณะนี้เรียกว่า แบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบแบบแบ่งชั้น (stradaptive test) ซึ่งมีโครงสร้างของการจัดข้อสอบ ดังภาพประกอบ 10



ภาพประกอบ 10 โครงสร้างของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบแบบแบ่งชั้น

(2) การทดสอบแบบหลายชั้นตอนรูปแบบแยกทางแปรผัน

การทดสอบแบบแยกทางแปรผันนี้ต่างจากการทดสอบแบบแยกทางคงที่ตรงที่ไม่สามารถกำหนดเส้นทางของการตอบข้อสอบได้แน่นอนตายตัว ข้อสอบที่จะตอบข้อต่อไปนั้นขึ้นอยู่กับผลการตอบข้อสอบข้อที่ผ่านมา ซึ่งต้องมีการประมาณค่าความสามารถของผู้สอบทุกครั้ง

หลังจากตอบข้อสอบแต่ละข้อ จึงเป็นเรื่องที่ยุงยากที่จะจัดแบบทดสอบเป็นแบบเขียนตอบ (paper and pencil test) ดังนั้น จึงมีการนำไมโครคอมพิวเตอร์เข้ามาใช้เพื่อความสะดวกในการดำเนินการสอบและบันทึกผลการสอบ การทดสอบจะเริ่มต้นโดยให้ผู้สอบนั่งที่หน้าจอคอมพิวเตอร์ให้ข้อมูลของตนเองตามที่ผู้ดำเนินการสอบต้องการ เช่น เลขที่ ชื่อ ชั้นเรียน ห้องเรียน เป็นต้น มีการแนะนำวิธีการตอบข้อสอบ หลังจาก que ผู้สอบเข้าใจวิธีการตอบดีแล้ว ข้อสอบข้อแรกที่จะถูกนำเสนอแก่ผู้สอบ ซึ่งตามปกติจะใช้ข้อสอบที่มีความยากง่ายปานกลางหรือใช้ข้อสอบที่ยากง่ายเหมาะกับผู้สอบ หากทราบสภาพภูมิหลังของผู้สอบมาก่อนข้อสอบถูกนำเสนอที่หน้าจอคอมพิวเตอร์ครั้งละ 1 ข้อ หลังจากการตอบข้อสอบแต่ละข้อผ่านไปข้อสอบข้อใหม่ก็จะปรากฏขึ้นมาที่หน้าจอตามเกณฑ์การคัดเลือกข้อสอบที่กำหนดไว้ การทดสอบจะดำเนินไปเรื่อย ๆ จนกระทั่งถึงเกณฑ์ยุติการสอบ การทดสอบก็จะสิ้นสุดลง

2. ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)

ทฤษฎีการทดสอบแบบมาตรฐานเดิม (Classical Test Theory) เป็นทฤษฎีที่ไม่เหมาะสมกับการทดสอบเฉพาะบุคคล เนื่องจากดัชนีความเที่ยงตรง ความเชื่อมั่น และคุณภาพของข้อสอบแบบมาตรฐานเดิมนั้นจะสอดคล้องกันเมื่อผู้สอบทุกคนได้ทำข้อสอบชุดเดียวกัน กล่าวคือ ค่าความยาก และค่าอำนาจจำแนกของข้อสอบไม่คงที่ แต่จะแปรเปลี่ยนไปตามกลุ่มตัวอย่าง (Hambleton and Swaminathan, 1985) และคะแนนของผู้สอบก็แปรเปลี่ยนไปตามค่าความยากของข้อสอบ จึงจำเป็นต้องใช้ทฤษฎีที่เหมาะสมกับการทดสอบเฉพาะบุคคล (Tailored Testing) ที่มีการจัดข้อสอบให้เหมาะสมกับระดับความสามารถของผู้สอบโดยที่ผู้สอบที่ต่างความสามารถกันจะได้รับชุดข้อสอบที่แตกต่างกัน และมีค่าความยากหรือง่ายไม่เท่ากัน ทฤษฎีที่เหมาะสมสามารถนำมาใช้กับการทดสอบเฉพาะบุคคลได้อย่างมีประสิทธิภาพ ก็คือ ทฤษฎีการตอบสนองข้อสอบ (IRT) ซึ่งจะกล่าวถึงทฤษฎีนี้เฉพาะในส่วนที่เกี่ยวกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ดังนี้

2.1 คุณสมบัติของรูปแบบการตอบสนองข้อสอบ

แฮมเบิลตันและสวามินาธาน (Hambleton and Swaminathan, 1985) ได้สรุปคุณสมบัติของรูปแบบการตอบสนองข้อสอบไว้ดังนี้

1. เป็นรูปแบบที่อธิบายเกี่ยวกับปฏิบัติการของผู้สอบในการทำแบบทดสอบ ว่าสามารถอธิบายได้ในคุณลักษณะหนึ่งหรือหลายคุณลักษณะ ซึ่งเรียกว่าคุณลักษณะแฝง
2. รูปแบบการตอบข้อสอบเป็นความสัมพันธ์ระหว่างปฏิบัติการของผู้สอบ ในการตอบข้อสอบทั้งที่สังเกตได้และคุณลักษณะแฝงหรือความสามารถที่ถือว่าเป็นปฏิบัติการหลักของพฤติกรรมในการทำแบบทดสอบ
3. รูปแบบการตอบข้อสอบที่ดีจะต้องให้วิธีทางในการประมาณคะแนนของผู้สอบได้ตรงกับคุณลักษณะแฝงนั้น
4. คุณลักษณะแฝงอธิบายได้จากปฏิบัติการที่สังเกตได้ของผู้สอบในการตอบข้อสอบ ความสัมพันธ์ระหว่างปริมาณของการทำข้อสอบหรือคะแนนของผู้สอบ (Test Performance) กับปริมาณความสามารถของผู้สอบ (Ability) สามารถเขียนในรูปของความสัมพันธ์ทั่ว ๆ ไปดังนี้ (ตรึงใจ พูลผลอำนวย, 2534 ; อ้างอิงจาก Lord, 1982)

$$P = f(U_i | \theta_1, \theta_2, \theta_3, \dots, \theta_k; \beta_j)$$

เมื่อ P แทน ผลการสอบ

f แทน ฟังก์ชัน

U_i แทน ผลการตอบข้อสอบ ตอบถูก $U_i = 1$ ตอบผิด $U_i = 0$

$\theta_1, \theta_2, \theta_3, \dots, \theta_k$ แทน ความสามารถ (Ability หรือ Trait) ที่ 1, 2, 3, ..., k

β_j แทน ค่าพารามิเตอร์ ของข้อสอบข้อที่ j

จากความสัมพันธ์ดังกล่าวนี้เป็นฟังก์ชันความสัมพันธ์ในลักษณะทั่ว ๆ ไป นักวัดผล การศึกษาจะต้องหา โมเดลทางคณิตศาสตร์ (Mathematical Model) ที่เหมาะสมเพื่อใช้แทน ฟังก์ชันความสัมพันธ์ดังกล่าว โดยอาศัยข้อตกลงเบื้องต้นต่าง ๆ ของทฤษฎีการตอบสนองข้อสอบ

2.2 ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

1. ความเป็นมิติเดียว (Unidimension) หมายถึง ข้อสอบแต่ละข้อในแบบทดสอบ จะต้องวัดความสามารถหรือคุณลักษณะเดียวกัน (unidimensionality) หรือ ข้อสอบเหล่านี้มีความเป็น เอกพันธ์ (Homogeneous) การกำหนดเช่นนี้ก็เพื่อให้รูปแบบของทฤษฎีนี้มีความหมายซับซ้อน น้อยลงและง่ายแก่การแปลความหมายของคะแนนที่ได้จากแบบทดสอบ

วิธีการตรวจสอบว่าแบบทดสอบนั้นวัดในมิติเดียวหรือไม่นั้น ทำได้หลายวิธี เช่น โดยการวิเคราะห์องค์ประกอบ (Factor Analysis) แล้วสังเกตค่าไอเกน (Eigen value) ค่าสูงสุด ว่า แตกต่างจากค่าอื่น ๆ อย่างชัดเจนหรือไม่ (Hambleton and Cook, 1977) ได้เสนอแนะว่าการที่จะทำให้เกิดความเป็นมิติเดียว (Unidimension) ได้ง่ายวิธีหนึ่งคือ การเลือกข้อสอบที่มีค่าอำนาจจำแนกสูง เนื่องจากข้อสอบที่มีค่าอำนาจจำแนกสูงจะมีสหพันธ์กับคะแนนแบบทดสอบทั้งหมดสูง ยูรี (Urry, 1977) เสนอไว้ว่าควรเลือกข้อสอบที่มีค่าอำนาจจำแนกตั้งแต่ .8 ขึ้นไป

2. ความเป็นอิสระในการตอบข้อสอบ (Local Independence) หมายถึง โอกาสในการตอบข้อสอบแต่ละข้อได้ถูกต้องเป็นอิสระจากกัน นั่นคือ การตอบข้อสอบข้อใดข้อหนึ่งได้ถูกหรือ ผิดจะไม่มีผลต่อการตอบข้ออื่น ๆ ด้วย (Lord, 1980 ; Hambleton and Swaminathan, 1985) และ แอมเบิลตันกับสวามินาธาน (Hambleton and Swaminathan, 1985) มีความเห็นตรงกันว่า ถ้า มีความเป็นมิติเดียวกันแล้ว ความเป็นอิสระในการตอบข้อสอบก็จะมีเอง

3. โค้งลักษณะข้อสอบ (Item Characteristic Curve) เป็นฟังก์ชันทางคณิตศาสตร์ ที่แสดงถึงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบนั้นได้ถูกต้องกับระดับความสามารถที่วัดได้ โดยชุดของข้อสอบหรือ แบบทดสอบนั้น (Hambleton and Swaminathan, 1985) ซึ่งจะเห็นว่า โอกาสผู้สอบตอบข้อสอบถูก (Probability) จะขึ้นอยู่กับโค้งลักษณะของข้อสอบ (ICC) ในแต่ละข้อมี คุณสมบัติไม่แปรเปลี่ยน (Invariant) ไปตามกลุ่มตัวอย่างหรือตัวผู้สอบ (Hambleton and Cook, 1977) ฉะนั้นโอกาสในการตอบข้อสอบถูกในแต่ละข้อจึงไม่แปรเปลี่ยนด้วย

เนื่องจากมีความเชื่อหลายความเชื่อ เกี่ยวกับความสัมพันธ์ระหว่าง โอกาสในการตอบข้อสอบได้ถูกต้องกับระดับความสามารถที่วัดได้โดยใช้ชุดของข้อสอบ หรือแบบทดสอบนั้น จึงทำให้เกิดมีรูปแบบของโค้งลักษณะข้อสอบ (Item Response Model) ขึ้นมาหลายรูปแบบ เช่น Guttman Perfect Scale, Latent Distance Model, Linear model, Normal orgive Model, Nominal Response Model และ Grade Response Model (จักรกฤษณ์ สำราญใจ. 2531) แต่รูปแบบที่นิยมใช้กันแพร่หลายคือ รูปแบบโลจิสติก (Logistic Model) รูปแบบนี้แบ่งออกเป็น 4 รูปแบบ ตามจำนวนพารามิเตอร์ของโค้งลักษณะข้อสอบ (Hambleton and Swaminathan. 1985)

2.3 โมเดลตามทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบเป็นทฤษฎีที่อธิบายความสัมพันธ์ระหว่างปริมาณของการกระทำหรือคะแนนของผู้สอบกับปริมาณความสามารถของผู้สอบ ซึ่งแสดงได้ดังสมการ (สำเร็จบุญเรืองรัตน์. 2529)

$$P = f(\theta)$$

- เมื่อ P แทน ผลการสอบ (performance)
 θ แทน ความสามารถหรือคุณลักษณะ (ability หรือ trait)
 f แทน ฟังก์ชัน (function)

เนื่องจากข้อมูลที่ได้รับจากผู้สอบนั้นมีหลายลักษณะ ดังนั้นจึงได้มีการพัฒนาโมเดลที่แสดงความสัมพันธ์นี้ในลักษณะต่าง ๆ กัน เพื่อให้แต่ละโมเดลมีความเหมาะสมกับข้อมูลที่ได้ โมเดลของทฤษฎีการตอบสนองข้อสอบสามารถจำแนกออกได้เป็น 3 กลุ่มใหญ่ ๆ ตามลักษณะของข้อมูลที่ได้จากการสอบ 3 ลักษณะ คือ คำตอบที่เป็นข้อมูลแบบแบ่งสอง (dichotomous) ข้อมูลแบบพหุ (multichotomous) และข้อมูลแบบต่อเนื่อง (contineous) โมเดลที่มีความเหมาะสมกับลักษณะข้อมูลดังกล่าว มีดังนี้ (รังสรรค์ มณีเล็ก. 2540 ; อ้างอิงจาก Hambleton and Swaminathan. 1985 : 33 – 52)

ตาราง 3 ลักษณะของข้อมูล โมเดลที่ใช้วิเคราะห์และผู้เสนอโมเดล

ลักษณะของข้อมูล	โมเดลที่ใช้วิเคราะห์	ผู้เสนอโมเดล
Dichotomous	Latent Linear	Lazarsfeld & Henry (1968)
	Latent Distance	
	Perfect scale	Guttman (1944)
	One, Two, Three Parameter	Lord (1952)
	Normal Ogive	
	One, Two, Three Parameter	Birnbaum (1957, 1958a,
	Logistic	1958b 1968), Lord & Novick (1968) Wright & Stone (1979)
Multicategory Scoring	Four Parameter Logistic	McDonald (1957)
		Barton & Lord (1981)
	Nominal Response	Bock (1972)
	Graded Response	Samejima (1969)
Continuous	Partial Credit Model	Master (1982)
	Continuous Response	Samejima (1972)

จากตารางที่ 3 ในแต่ละโมเดลก็มีฟังก์ชันทางคณิตศาสตร์แตกต่างกันออกไป ซึ่งสามารถแสดงได้ตามลำดับของการคิดค้นโมเดล ดังนี้

2.3.1 โมเดลนอร์มอลโอโงอิจแบบ 2 พารามิเตอร์ (two parameter normal ogive model) ลอร์ด (Lord.) เป็นผู้เสนอโมเดลนี้ โด่งลักษณะข้อสอบจะได้จากสมการ

$$P_i(\theta) = \int_{-\infty}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

เมื่อ $P_i(\theta)$ แทน ความน่าจะเป็นที่ผู้สอบที่มีความสามารถ θ จะตอบข้อที่ i ได้ถูกต้อง

b แทน ค่าความยากของข้อสอบข้อที่ i

a แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ i

Z แทน ความเบี่ยงเบนปกติ จากการแจกแจงที่ค่าเฉลี่ย b และความเบี่ยงเบนมาตรฐาน $1/a$

2.3.2 โมเดลโลจิสติกแบบ 2 พารามิเตอร์ (two parameter logistic model) เบิร์น บอม (Birnbau) เป็นผู้เสนอโมเดลนี้ ซึ่งโค้งลักษณะข้อสอบจะได้จากฟังก์ชันการแจกแจง ดังนี้

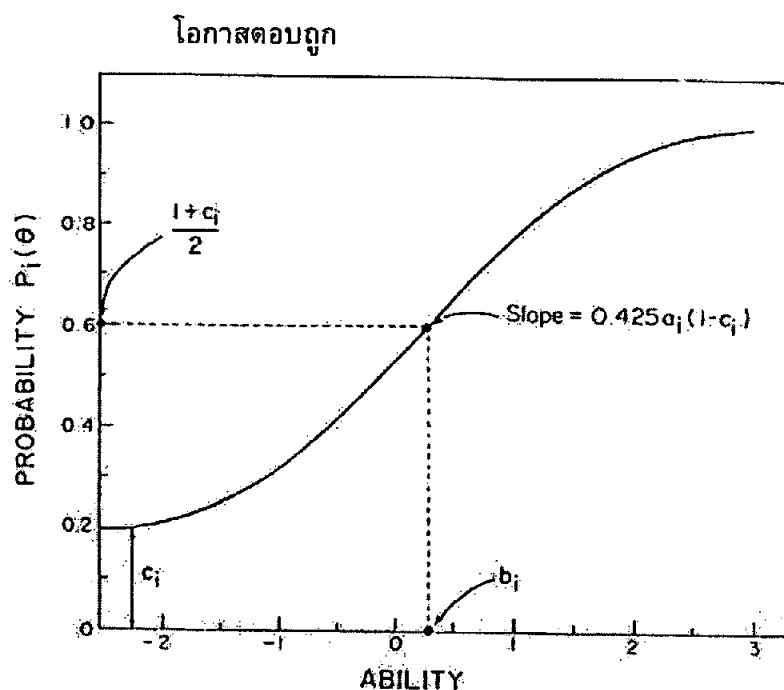
$$P_i(\theta) = \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}} \quad (i=1, 2, \dots, n).$$

ค่าของ $P_i(\theta)$, a , b และ θ มีความหมายเหมือนเดิม ส่วน D เป็นค่าที่ใช้ในการปรับแก้ (Scaling factor) ซึ่งถ้ากำหนดให้ $D = 1.7$ แล้ว ค่าของ $P(\theta)$ จากโมเดลนอร์มอลโอไฮโอและโมเดลโลจิสติกแบบ 2 พารามิเตอร์จะต่างกันน้อยกว่า .01 ในทุก ๆ ค่าของ θ ที่ประมาณได้ และโมเดลโลจิสติกนั้นนำไปใช้ได้สะดวกกว่าโมเดลนอร์มอลโอไฮโอในการประมาณค่า $P(\theta)$

2.3.3 โมเดลโลจิสติกแบบ 3 พารามิเตอร์ (three parameter logistic model) โมเดลนี้พัฒนามาจากโมเดลโลจิสติกแบบ 2 พารามิเตอร์ โดยเพิ่มค่าพารามิเตอร์การเดา (c) เข้าไปในฟังก์ชันอีก 1 ค่า ทำให้ได้ฟังก์ชันของการแจกแจง ดังนี้

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

พารามิเตอร์ c เป็นค่าความน่าจะเป็นของผู้สอบที่มีความสามารถต่ำมาก ๆ แต่สามารถตอบข้อสอบได้ถูกต้อง ค่าความยากง่ายของข้อสอบ (b) จะอยู่บนสเกลความสามารถ (θ) ตรงจุดที่ค่าความชันของโค้งมีค่าสูงสุด ซึ่งค่าความชันของโค้งที่สูงสุดมีค่า เท่ากับ $0.425a(1-c)$ โค้งลักษณะข้อสอบที่มีค่าอำนาจจำแนกสูงจะมีความชันมากกว่าโค้งลักษณะข้อสอบที่มีค่าอำนาจจำแนกต่ำ ความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้สอบที่มีความสามารถเท่ากับค่าความยากง่ายของข้อสอบ จะเท่ากับ $(1 + c)/2$ ซึ่งโค้งแสดงลักษณะของข้อสอบตามโมเดลนี้มีลักษณะ ดังภาพประกอบ 11

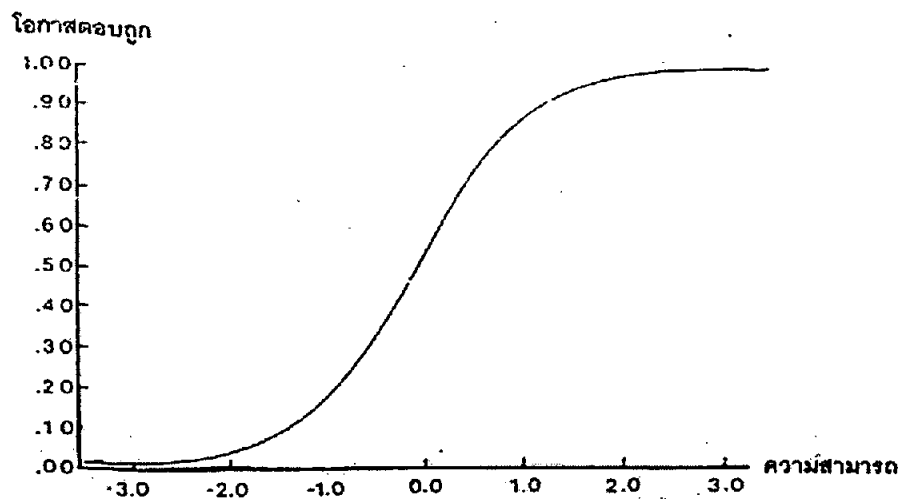


ภาพประกอบ 11 โค้งแสดงลักษณะของข้อสอบตามโมเดลโลจิสติกแบบ 3 พารามิเตอร์

2.3.4 โมเดลโลจิสติกแบบ 1 พารามิเตอร์หรือราสช์โมเดล (one parameter logistic model, rasch model) ราสช์ (Rasch) ซึ่งเป็นนักคณิตศาสตร์ชาวเดนมาร์กเป็นผู้เสนอโมเดลนี้ ซึ่งเป็นโมเดลที่เป็นกรณีเฉพาะของโมเดลโลจิสติก 3 พารามิเตอร์ของเบอร์นบอม (Birnbaum) ที่มี ข้อตกลงว่าข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากันและมีค่าการเดาน้อยมาก ๆ ทำให้ได้ฟังก์ชันของการแจกแจง ดังนี้

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

เมื่อ a แทน ค่าอำนาจจำแนกของข้อสอบ ซึ่งกำหนดให้มีค่าเท่ากันทุกข้อ ซึ่งมีโค้งแสดงลักษณะของข้อสอบ ดังแสดงในภาพประกอบ 12



ภาพประกอบ 12 โค้งแสดงลักษณะของข้อสอบตามโมเดลโลจิสติกแบบ 1 พารามิเตอร์

2.3.5 โมเดลโลจิสติกแบบ 4 พารามิเตอร์ (four parameter logistic model)

แมคโดนัลด์ (McDonald.) บาร์ตันและลอร์ด (Barton and Lord.) เป็นผู้นำเสนอโมเดลนี้โดยมีความเชื่อว่าผู้สอบที่มีความสามารถสูง ๆ นั้นบางครั้งอาจตอบข้อสอบผิดได้ด้วยความสะดวกหรือการเฉลยข้อสอบที่ผิดของผู้สร้างข้อสอบ ทำให้ได้ฟังก์ชันของการแจกแจง ดังนี้

$$P_i(\theta) = c_i + (\gamma - c_i) \frac{e^{D a_i(\theta - b_i)}}{1 + e^{D a_i(\theta - b_i)}}$$

เมื่อ γ เป็นค่าพารามิเตอร์ความสะดวก ซึ่งมีค่าต่ำกว่า 1 เล็กน้อย โมเดลนี้เป็นเพียงโมเดลที่ศึกษาตามทฤษฎีเท่านั้น ส่วนการปฏิบัติจริงนั้นยังไม่ค่อยได้นำมาใช้

2.3.6 โมเดลนอร์มอลโอโงอิจ์แบบ 1 3 และ 4 พารามิเตอร์ (one three and four parameter normal ogive model)

ลอร์ด (Lord.) เป็นผู้นำเสนอโมเดลนี้ ซึ่งมีฟังก์ชันของการแจกแจง ดังนี้

$$P_i(\theta) = \int_{-\alpha}^{\theta - b_i} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

$$P_i(\theta) = c + (1 - c) \int_{-\alpha}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

$$P_i(\theta) = c + (\gamma - c) \int_{-\alpha}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

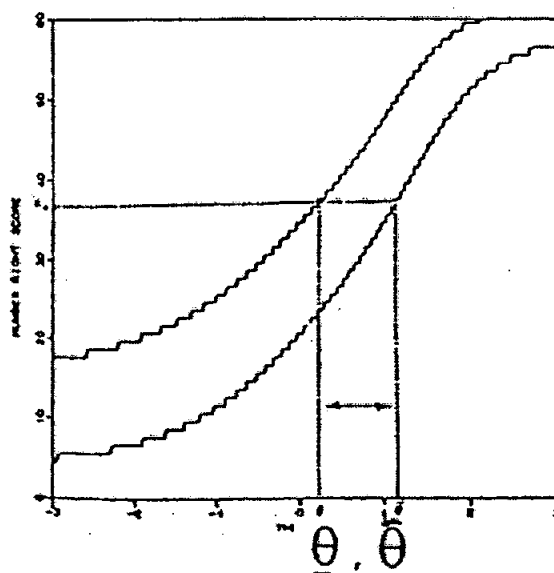
เมื่อเปรียบเทียบกับโมเดลโลจิสติกแล้ว พบว่า สมการจะมีลักษณะคล้าย ๆ กัน แต่สมการของโมเดลโลจิสติกจะง่ายกว่า ซึ่งในทางปฏิบัติแล้วไม่ค่อยได้ใช้โมเดลนอร์มอลโอใจผีเท่าใดนัก เนื่องจากมีข้อจำกัดในเรื่องความยุ่งยากในการประมาณค่าและจะใช้โมเดลโลจิสติกแทน สำหรับโมเดลนอร์มอลโอใจผีนั้นเป็นการศึกษาตามทฤษฎีเท่านั้น

2.4 ฟังก์ชันสารสนเทศ (information function)

ในสถานการณ์ทั่วไป ถ้ามั่นใจค่อนข้างมากกว่าเหตุการณ์อย่างหนึ่งจะเกิดขึ้นแสดงว่ามีข่าวสารข้อมูลหรือสารสนเทศเกี่ยวกับเหตุการณ์นั้นมากพอสมควร ในทางกลับกันถ้าไม่มีข่าวสารข้อมูลเกี่ยวกับเหตุการณ์นั้นหรือมีน้อยความมั่นใจก็จะมีน้อยตามไปด้วย ในการอ้างอิงเชิงสถิติ ความแม่นยำของการประมาณค่าพารามิเตอร์ของกลุ่มประชากรอาจดูได้จากช่วงกว้างของค่าประมาณ ถ้าไม่มีสารสนเทศใด ๆ เกี่ยวกับประชากรเลยก็อาจจะต้องประมาณค่าเป็นค่าใด ๆ ในช่วง $-\infty$ ถึง $+\infty$ แต่ถ้ามีสารสนเทศเกี่ยวกับประชากรบ้าง ช่วงของค่าประมาณจะแคบเข้า นั่นคือความแม่นยำในการประมาณเริ่มมีมากขึ้น ตามปกติความแม่นยำของการประมาณค่าพารามิเตอร์จะแสดงได้ด้วยค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า กล่าวคือ ถ้าความคลาดเคลื่อนมาตรฐานของการประมาณค่ามีมาก ความแม่นยำของการประมาณค่าก็จะมีน้อย เพราะช่วงของค่าประมาณจะกว้าง ในทางกลับกัน ถ้าความคลาดเคลื่อนมาตรฐานของการประมาณค่ามีน้อย ความแม่นยำของการประมาณค่าก็จะมีมาก ช่วงของค่าประมาณจะแคบ แสดงว่าค่าสารสนเทศมีความสัมพันธ์กับความคลาดเคลื่อนมาตรฐานของการประมาณค่า โดยมีความสัมพันธ์ในทิศทางกลับกัน คือความคลาดเคลื่อนมาตรฐานของการประมาณค่าเท่ากับ $1/\sqrt{\text{สารสนเทศ}}$ ในทฤษฎีการตอบสนองข้อสอบจะใช้ผลการตอบแบบทดสอบประมาณค่าความสามารถของผู้สอบ การประเมินคุณภาพของแบบทดสอบดูได้จากความถูกต้องแม่นยำในการประมาณค่าความสามารถโดยใช้คะแนนจากแบบทดสอบ ค่าสารสนเทศจากแบบทดสอบจะเป็นดัชนีบ่งชี้ถึงความถูกต้องแม่นยำของการประมาณค่า (จักรกฤษณ์ ส้าราญใจ. 2531 : 40 ; อ้างอิงจาก Bimbaum. 1968 : 418) เนื่องจากฟังก์ชันลักษณะของข้อสอบ $P_i(\theta)$ ขึ้นอยู่กับพารามิเตอร์ของข้อสอบข้อที่ i คือ a_i , b_i , และ c_i ดังนั้นค่าพารามิเตอร์ของข้อสอบจึงเป็นตัวกำหนดค่าสารสนเทศของข้อสอบและแบบทดสอบจึงอาจกล่าวได้ว่า ค่าสารสนเทศของข้อสอบหรือแบบทดสอบเป็นดัชนีผสม (Composite Index) ที่สร้างจากดัชนีบอกคุณลักษณะของข้อสอบหลายลักษณะรวมเป็นดัชนีเพียงตัวเดียวเพื่อชี้ถึงคุณภาพของข้อสอบหรือแบบทดสอบ อีกทั้งคุณสมบัติความไม่แปรเปลี่ยนของค่าพารามิเตอร์ของข้อสอบ ค่าสารสนเทศจึงเหมาะที่จะใช้เป็นดัชนีบอกคุณภาพของข้อสอบหรือแบบทดสอบได้ดีกว่าค่าสถิติหรือค่าดัชนีอื่น ๆ ตามแนวคิดของทฤษฎีการวัดแบบมาตรฐานเดิม (จักรกฤษณ์ ส้าราญใจ. 2531 : 42)

ทฤษฎีการตอบสนองข้อสอบได้กล่าวถึงฟังก์ชันสารสนเทศไว้ โดยแบ่งออกเป็นฟังก์ชันสารสนเทศของคะแนน (score information function) ฟังก์ชันสารสนเทศของข้อสอบ (item information function) และฟังก์ชันสารสนเทศของแบบทดสอบ (test information function) ซึ่งแต่ละเรื่องมีรายละเอียด ดังนี้ (รังสรรค์ มณีเล็ก. 2541 ; อ้างอิงจาก Hambleton and Swaminathan. 1985 : 124)

2.4.1 ฟังก์ชันสารสนเทศของคะแนน (score information function) เมื่อเรานำคะแนนผลการสอบของผู้สอบแต่ละคนที่มีความสามารถต่าง ๆ กันมาแจกแจงความถี่และสร้างเป็นแผนภูมิ ก็พบว่า คะแนนในแต่ละระดับความสามารถ ส่วนใหญ่จะตกอยู่ในช่วง 95% ของการแจกแจงในแต่ละระดับความสามารถ ดังภาพประกอบ 13



ภาพประกอบ 13 ช่วงความเชื่อมั่น (confidence interval ; $\underline{\theta}, \bar{\theta}$) ของความสามารถที่ได้จากการประมาณค่า ($\hat{\theta}$)

เนื่องจากคะแนนนั้นเป็นจำนวนนับที่ไม่ต่อเนื่อง จึงทำให้ภาพที่เห็นเป็นรูปขั้นบันได จากภาพ กล่าวได้ว่า ผู้สอบที่ได้คะแนน x_0 จะมีความสามารถอยู่ในช่วง $(\underline{\theta}, \bar{\theta})$ ด้วยความเชื่อมั่น 95%

ถ้าให้ $z = x/n$ ดังนั้น การถดถอยของ z ลงบน θ จะได้

$$\mu_{z|\theta} = \frac{1}{n} \sum_{i=1}^n P_i(\theta) = \xi$$

ถึงแม้ว่า n จะมากขึ้นแต่ก็ไม่ทำให้การถดถอยนี้เปลี่ยนไป ความแปรปรวนของ z สำหรับ θ ที่คงที่ ก็คือ

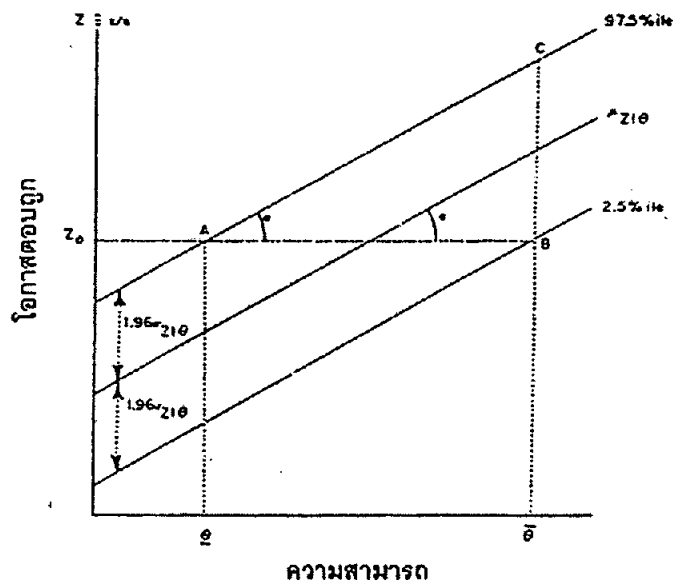
$$\sigma^2_{z|\theta} = \frac{1}{n^2} \sum_{i=1}^n P_i Q_i = \left(\frac{1}{n}\right) (\overline{PQ} - \sigma^2_{P|\theta})$$

ซึ่งเมื่อ n มีค่ามากขึ้น σ^2 ก็จะมีค่าเข้าใกล้ 0

เมื่อนำค่า z มาเขียนเป็นแผนภูมิของการแจกแจง ลักษณะของโค้งการแจกแจงก็จะมีลักษณะเหมือนกับ ภาพประกอบ 13 เมื่อลากเส้นถดถอยของ z ลงบน θ เส้นถดถอยนี้จะอยู่ระหว่าง

กลางฟังก์ชันแบบขั้นบันได 2 ฟังก์ชัน ซึ่งอยู่ที่ตำแหน่งเปอร์เซ็นต์ไทล์ที่ 2.5 และ 97.5 ของการแจกแจง z เมื่อกำหนด θ ให้ และเมื่อ n มีค่าสูงขึ้นการถดถอยก็จะมีค่าเปลี่ยนแปรปรวน จะลดลงเข้าหา 0 เมื่อฟังก์ชันขั้นบันไดลู่เข้าหากัน จนกระทั่งการแจกแจงของ z เป็นโค้งปกติ ตำแหน่งเปอร์เซ็นต์ไทล์ที่ 2.5 และ 97.5 จะอยู่ห่างจากเส้นถดถอยประมาณ $1.96\sigma_{z|\theta}$ หรือ

$$\frac{1.96}{n} \sqrt{\sum_{i=1}^n P_i Q_i} \quad \text{ดังภาพประกอบ 14}$$



ภาพประกอบ 14 ช่วงแห่งความเชื่อมั่นแอสิมโทติก 95% ($\underline{\theta}, \bar{\theta}$) สำหรับความสามารถ θ

$$\tan \alpha = \frac{\overline{CB}}{\overline{AB}} = \frac{2(1.96 \sigma_{z|\theta})}{\overline{AB}}$$

เมื่อ \overline{AB} แทน ช่วงของความเชื่อมั่นสำหรับความสามารถ θ

(asymtotic confidence interval ; $\underline{\theta}, \bar{\theta}$)

\overline{CB} แทน ค่าเบี่ยงเบนมาตรฐานของ Z

ดังนั้น

$$\frac{1}{(\overline{AB})^2} = \frac{(d\mu_{z|\theta})^2}{(3.92)^2 \text{Var}(Z|\theta)}$$

เมื่อ $\frac{1}{(\overline{AB})^2}$ แทน ฟังก์ชันสารสนเทศของคะแนน ใช้สัญลักษณ์ $\{|\theta, y\}$

ซึ่งเบอร์นบอม (ผจงจิต อินทสุวรรณ. 2528 : 100 ; อ้างอิงจาก Birnbaum. 1968) ได้นิยาม ฟังก์ชันสารสนเทศของคะแนน y ใน θ ดังนี้

$$|\{\theta, y\} = \frac{(d\mu_{y|\theta})^2}{\text{Var}(y|\theta)}$$

ค่าสารสนเทศของคะแนนนี้จะแปรเปลี่ยนไปตามระดับความสามารถ ซึ่งสรุปได้ว่า

(1) ที่ความสามารถระดับเดียวกัน เส้นถดถอยที่มีความชันมากกว่าจะให้สารสนเทศของคะแนนสูงกว่าเส้นถดถอยที่มีความชันน้อยกว่า

(2) ค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ต่ำจะให้ค่าสารสนเทศของคะแนนสูง

2.4.2 ฟังก์ชันสารสนเทศของแบบทดสอบ (test information function) ค่าความสามารถโดยประมาณที่เป็นไปได้สูงสุด (maximum likelihood, θ) ในสภาวะปกติจะมีการแจกแจงในลักษณะเอซิมโทติก มีค่าเฉลี่ยเท่ากับ 0 และค่าความแปรปรวนเท่ากับ (ริงสรรค์ มณีเล็ก. 2540 ; อ้างอิงจาก Lord, Frederic M. 1980)

$$\begin{aligned} \text{Var}(\hat{\theta}|\theta_0) &= \frac{1}{\zeta \left[\left(\frac{d \ln L}{d\theta} \right)^2 \theta_0 \right]} \\ \frac{1}{\text{Var}(\hat{\theta}|\theta_0)} &= \zeta \left\{ \left[\sum_{i=1}^n (u_i - P_i) P_i' / P_i Q_i \right]^2 \mid \theta_0 \right\} \\ &= \zeta \left\{ \left[\sum_{i=1}^n (u_i - P_i) P_i' / P_i Q_i \right] \left[\sum_{j=1}^n (u_j - P_j) P_j' / P_j Q_j \right] \mid \theta_0 \right\} \\ &= \sum_{i=1}^n \sum_{j=1}^n \frac{P_i' P_j'}{P_i P_j Q_i Q_j} \zeta [(u_i - P_i)(u_j - P_j) \mid \theta_0] \\ \frac{1}{\text{Var}(\hat{\theta}|\theta_0)} &= \sum_{i=1}^n \frac{P_i'^2}{P_i^2 Q_i^2} \text{Var}(u_{i0} / \theta_0) = \sum_{i=1}^n \frac{P_i'^2}{P_i^2 Q_i^2} P_i Q_i \\ \text{Var}(\hat{\theta}|\theta_0) &= \frac{1}{\sum_{i=1}^n \frac{P_i'^2}{P_i Q_i}} \\ I(\theta) = I(\hat{\theta}) &= \sum_{i=1}^n \frac{P_i'^2}{P_i Q_i} \end{aligned}$$

เมื่อ

- L แทน ฟังก์ชันความเป็นไปได้ (likelihood function)
- $I\{\theta\}$ แทน ฟังก์ชันสารสนเทศของแบบทดสอบ
- $P_i Q_i$ แทน ความน่าจะเป็นในการตอบถูก และตอบผิดตามลำดับ
- P_i' แทน ค่าความลาดชันของเส้นโค้ง

หรือเขียนในรูปอื่นได้อีกเช่น

$$I(\theta) = \sum_{i=1}^n \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

ลักษณะของฟังก์ชันสารสนเทศของแบบทดสอบ สรุปได้ดังนี้

- (1) ฟังก์ชันสารสนเทศของแบบทดสอบได้ถูกกำหนดขึ้นสำหรับชุดของข้อสอบที่แต่ละจุดของสเกลความสามารถ
- (2) ค่าของสารสนเทศของแบบทดสอบเป็นผลมาจากคุณภาพและจำนวนของข้อสอบ
- (3) ที่ความสามารถระดับเดียวกัน เส้นถดถอยที่มีความชันมากกว่าจะให้ค่าสารสนเทศของแบบทดสอบสูงกว่าเส้นถดถอยที่มีความชันน้อยกว่า
- (4) ข้อสอบที่มีค่าความแปรปรวนต่ำจะส่งผลให้ค่าสารสนเทศของแบบทดสอบมีค่าสูง
- (5) ค่าสารสนเทศของแบบทดสอบจะไม่ขึ้นกับการจัดหมู่เฉพาะของข้อสอบ ข้อสอบแต่ละข้อเป็นอิสระจากกัน
- (6) ค่าสารสนเทศของแบบทดสอบมีความสัมพันธ์แบบผกผันกับค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ

2.4.3 ฟังก์ชันสารสนเทศของข้อสอบ (item information function) ข้อสอบแต่ละข้อจะมีฟังก์ชันแสดงสารสนเทศ ได้แก่ (Lord, 1980 ; Baker, 1992)

$$I_i\{\theta\} = \frac{P_i'^2}{P_i Q_i}$$

หรือ เขียนในรูปอื่นได้อีก เช่น

$$I_i(\theta) = \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

$$I_i\{\theta\} = D^2 a_i^2 \frac{Q_i}{P_i} \left(\frac{P_i - c_i}{1 - c_i} \right)^2$$

$$I_i\{\theta\} = \frac{D^2 a_i^2 (1 - c_i)}{(c_i + e^{DL_i}) (1 - e^{-DL_i})^2}$$

$$\text{เมื่อ } L_i = a_i(\theta - b_i)$$

ฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อจะรวมกันเป็นฟังก์ชันสารสนเทศของแบบทดสอบ ซึ่งค่าฟังก์ชันสารสนเทศของข้อสอบจะขึ้นอยู่กับค่าความชันของฟังก์ชันการตอบข้อสอบและค่าความ

แปรปรวนที่มีเงื่อนไขที่แต่ละระดับของความสามารถ ค่าความชันสูง ๆ ค่าความแปรปรวนต่ำ ๆ จะทำให้ค่าสารสนเทศของข้อสอบมีค่าสูง ซึ่งจะทำให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดมีค่าต่ำ การแจกแจงของฟังก์ชันสารสนเทศของข้อสอบมีลักษณะเป็นรูปประฆังคว่ำ ค่าสารสนเทศที่สูงที่สุดจะอยู่ที่จุด b บนสเกลความสามารถสำหรับโมเดลการตอบโลจิสติกแบบ 1 และ 2 พารามิเตอร์ ส่วนโมเดล 3 พารามิเตอร์นั้น ค่าสารสนเทศของข้อสอบข้อที่ i จะสูงสุดที่จุด θ_{\max}

$$\theta_{\max} = b_i + \frac{1}{Da_i} \ln \left[\frac{1}{2} + \frac{\sqrt{1+8c_i}}{2} \right]$$

สำหรับโมเดล 1 พารามิเตอร์นั้นค่าสูงสุดของค่าสารสนเทศของข้อสอบจะคงที่ สำหรับโมเดล 2 พารามิเตอร์ ค่าสูงสุดของค่าสารสนเทศของข้อสอบจะเป็นสัดส่วนโดยตรงกับกำลังสองของค่าอำนาจจำแนก ถ้าค่าอำนาจจำแนกของข้อสอบสูงก็จะทำให้ค่าสารสนเทศของข้อสอบมีค่ามาก ส่วนโมเดล 3 พารามิเตอร์นั้น ค่าสารสนเทศของข้อสอบสูงสุดจะมีค่าเท่ากับ (Hambleton and Swaminathan, 1985 : 105 ; Lord 1980a : 152)

$$I_i\{\theta\}_{\max} = \frac{D^2 a_i^2}{8(1-c_i)^2} [1 - 20c_i - 8c_i^2 + (1+8c_i)^{3/2}]$$

ลักษณะของสารสนเทศของข้อสอบ

- (1) ผลรวมค่าสารสนเทศของข้อสอบทุกข้อ คือ ค่าสารสนเทศของแบบทดสอบ
- (2) ค่าฟังก์ชันสารสนเทศของข้อสอบจะขึ้นอยู่กับค่าความชันของฟังก์ชันการตอบสนองข้อสอบและค่าความแปรปรวนที่มีเงื่อนไขที่แต่ละระดับของความสามารถ ถ้าค่าความชันมาก ๆ และค่าความแปรปรวนต่ำ ๆ จะทำให้ค่าสารสนเทศของข้อสอบมีค่าสูงซึ่งจะทำให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดมีค่าต่ำ

2.5 ประสิทธิภาพสัมพัทธ์ (relative efficiency)

หากต้องการเปรียบเทียบแบบทดสอบ 2 ฉบับที่วัดคุณลักษณะเดียวกันว่าฉบับใดจะมี ประสิทธิภาพสูงกว่ากัน ก็สามารถทำได้โดยการใช้ประสิทธิภาพสัมพัทธ์ ซึ่งค่าประสิทธิภาพสัมพัทธ์ของคะแนนจากแบบทดสอบ y กับคะแนนจากแบบทดสอบ x เป็นอัตราส่วนของฟังก์ชันสารสนเทศของคะแนน ดังนี้ (Lord, 1980)

$$RE\{y, x\} = \frac{I\{\theta, y\}}{I\{\theta, x\}}$$

- ถ้า $RE\{y, x\} = 1$ แสดงว่า แบบทดสอบทั้งสองฉบับมีความแม่นยำในการประมาณค่าความสามารถเท่า ๆ กัน
- $RE\{y, x\} > 1$ แสดงว่า แบบทดสอบ Y มีความแม่นยำในการประมาณค่าความสามารถมากกว่าแบบทดสอบ X

$RE\{y, x\} < 1$ แสดงว่า แบบทดสอบ Y มีความแม่นยำในการ
ประมาณค่าความสามารถน้อยกว่าแบบทดสอบ X

คะแนน x และ y อาจเป็นคะแนนจากแบบทดสอบสองฉบับที่ต่างกันแต่วัดความสามารถ
เดียวกันหรืออาจเป็นคะแนนจากแบบทดสอบฉบับเดียวกัน แต่ให้คะแนนด้วยวิธีที่ต่างกัน ค่า
ประสิทธิภาพสัมพัทธ์ของคะแนนจะแปรเปลี่ยนไปในแต่ละระดับความสามารถ ซึ่งสามารถเขียนให้อยู่
ในรูปของการแจกแจงคะแนนจริงได้ดังนี้

$$RE\{y, x\} = \frac{(\sigma^2_{x|\xi})p^2(\xi)}{\sigma^2_{y|\eta(\xi)}q^2[\eta(\xi)]}$$

ถ้าใช้คะแนนจากการสอบแทนคะแนนจริง จะได้

$$RE\{y, x\} = \frac{n_y x (n_x - x) f_x^2}{n_x y (n_y - y) f_y^2}$$

เมื่อ $p^2(\xi), q^2[\eta(\xi)]$ แทน การแจกแจงความถี่ของคะแนนจริง ξ และ
 η ตามลำดับ

$\sigma^2_{x|\xi}, \sigma^2_{y|\eta(\xi)}$ แทน ค่าความแปรปรวนของคะแนนจากแบบทดสอบ x และ y
ตามลำดับ

x, y แทน คะแนนที่ตอบถูกจากแบบทดสอบ x และ y ที่ตำแหน่ง
เปอร์เซ็นต์ไทล์เดียวกัน

f_x, f_y แทน ความถี่สัมพัทธ์ระหว่างกลุ่ม x และ y ตามลำดับ

n_x, n_y แทน จำนวนข้อสอบในแบบทดสอบ x และ y ตามลำดับ

2.6 การประมาณค่าความสามารถของผู้สอบ

ทฤษฎีการตอบสนองข้อสอบมีความเชื่อว่าสิ่งที่อยู่เบื้องหลังและส่งผลต่อการทดสอบ
ก็คือความสามารถของผู้สอบ (ability) ซึ่งอาจเป็นความสามารถเพียงอย่างเดียว หรือหลาย ๆ อย่างต่อ
ผลการสอบในครั้งหนึ่ง ๆ ก็ได้ ความสามารถดังกล่าวถูกกำหนดขึ้นเพื่อใช้ในการระบุคุณลักษณะ
ภายใน (trait หรือ characteristic) ที่ต้องการวัดจากผู้สอบ ความสามารถนั้นเป็นสิ่งที่ถูกพรรณนาว่า
เป็นสิ่งที่วัดได้ด้วยชุดของข้อสอบ อาจนิยามความสามารถอย่างกว้าง ๆ ว่าเป็นตัวแปรด้านความถนัด
หรือผลสัมฤทธิ์ทางการเรียน เช่น ความเข้าใจในการอ่าน ความสามารถทางจำนวน หรือ อาจนิยามใน
วงแคบ เช่น ความสามารถในการคูณจำนวน หรือ ตัวแปรที่เกี่ยวข้องกับบุคลิกภาพ เช่น สังกัปแห่งตน
แรงจูงใจ ความสามารถเหล่านี้มิใช่เป็นสิ่งที่มียุติตายตัว แต่สามารถเปลี่ยนแปลงได้ตลอดเวลา
โดยเฉพาะเมื่อได้รับการเรียนการสอนเพิ่มขึ้น ในการแปลความหมายของความสามารถนั้น ก็จะใช้ผล
ที่ได้จากการศึกษาความเที่ยงตรงเชิงโครงสร้าง (Hambleton and Swaminathan. 1985 : 55)

ความสามารถ (ability, θ) และคะแนนจริง (true score, ξ) นั้นเป็นสิ่งเดียวกันแต่ใช้มาตรวัดต่างกัน มาตรวัดคะแนนจริงขึ้นอยู่กับข้อสอบทั้งหลายในแบบทดสอบ ส่วนมาตรวัดความสามารถนั้นจะเป็นอิสระจากข้อสอบทั้งหลายในแบบทดสอบ ดังนั้นค่าของความสามารถจึงมีประโยชน์มากกว่าคะแนนจริงเมื่อต้องการเปรียบเทียบแบบทดสอบต่างฉบับที่วัดความสามารถเดียวกัน ลอร์ดและโนวิก (Lord, 1980 : 45 – 46) กล่าวว่า คะแนนจริงเป็นค่าที่คาดหวังของคะแนนที่วัดได้ ผู้สอบทุกคนที่มีความสามารถเท่ากับ θ จะมีคะแนนจริงของจำนวนคำตอบถูกเท่ากับ คือ

$$\xi = \sum_{i=1}^n P_i(\theta)$$

ขั้นตอนในการประมาณค่าความสามารถของผู้สอบนั้นมี ดังนี้

- (1) รวบรวมข้อมูลการตอบของผู้สอบ
- (2) เปรียบเทียบความเหมาะสม (fit) ระหว่างโมเดลการวิเคราะห์ข้อมูลกับข้อมูลที่ได้มา และเลือกใช้โมเดลสำหรับวิเคราะห์ข้อมูลที่เหมาะสม
- (3) ประมาณค่าพารามิเตอร์ต่าง ๆ
- (4) กำหนดสเกลและแปลงค่าความสามารถให้อยู่ในรูปของสเกลง่าย ๆ

ในการประมาณค่าความสามารถของผู้สอบนั้นมีอยู่หลายวิธีด้วยกัน แต่พอจะจำแนกออกเป็น 2 กลุ่มใหญ่ ๆ ได้แก่ วิธีความเป็นไปได้สูงสุด (maximum likelihood) และวิธีของเบส์ (bayesian) แต่การศึกษาวิจัยในครั้งนี้ผู้วิจัยเลือกวิธีการประมาณค่าความสามารถตามวิธีของเบส์ ซึ่งมีรายละเอียดดังนี้ (รังสรรค์ มณีเล็ก, 2540 ; อ้างอิงจาก Hambleton and Swaminathan, 1985 : 76 – 97)

การประมาณค่าความสามารถของผู้สอบ ถ้าผู้สอบตอบข้อสอบถูกทุกข้อหรือผิดหมดทุกข้อ การประมาณค่าความสามารถด้วยวิธีความเป็นไปได้สูงสุดก็จะไม่เหมาะสม เว้นแต่จะนำเอาผลการตอบของผู้สอบเหล่านั้นออกไปจากกลุ่มผู้สอบก่อน หากทราบค่าสารสนเทศภายนอก (prior information) การประมาณค่าความสามารถด้วยวิธีของเบส์ก็จะเกิดประสิทธิภาพมากกว่าค่าสารสนเทศภายนอกที่มักนิยมใช้กันก็คือความสามารถของผู้สอบที่มีการแจกแจงเป็นรูปโค้งปกติ มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 (Hambleton and Swaminathan, 1985 : 92)

การกำหนดรูปร่างของการแจกแจงเป็นรูปโค้งปกตินับว่าค่อนข้างสะดวก เบอร์นบวม (Birnbbaum) เชื่อว่า การแจกแจงภายนอกของความสามารถเป็นฟังก์ชันโลจิสติก (logistic density function) ซึ่งมีสมการ ดังนี้

$$f(\theta) = \frac{\exp.(\theta)}{[1 + \exp.(\theta)]^2}$$

หัวใจของวิธีการประมาณค่าความสามารถตามวิธีการของเบส์ คือ ทฤษฎีของเบส์ (bayes theorem) ซึ่งมีความสัมพันธ์กับ condition probability และ marginal probability

$$P(B|A) = \frac{P(A|B) P(B)}{P(A)}$$

ในการประมาณค่าความสามารถของผู้สอบ A อาจถูกกำหนดให้เป็น θ_a และ B เป็นชุดของคำตอบของข้อสอบ n ข้อ ได้คะแนน u แล้วสมการ $P(B | A)$ เขียนใหม่ได้เป็น

$$P(\theta_a | u) = \frac{P(u | \theta_a) P(\theta_a)}{P(u)}$$

เนื่องจาก θ_a เป็นตัวแปรต่อเนื่อง ดังนั้นจึงถูกแปลความว่าเป็นฟังก์ชันความหนาแน่น (density function) เพื่อเป็นการป้องกันการกำสับสนระหว่าง $P(\theta_a)$ กับ $P(\theta)$ ซึ่งเป็นฟังก์ชันการตอบสนองข้อสอบ (item response function) จึงใช้ $f(\theta_a)$ แทน $P(\theta)$ ดังนั้น จึงเขียนสมการใหม่ได้ ดังนี้

$$f(\theta_a | u) = \frac{f(u | \theta_a) f(\theta_a)}{f(u)}$$

เมื่อ $f(u)$ แทน ค่าคงที่
 $f(\theta_a | u)$ แทน ความหนาแน่นภายหลัง (posterior density) ของ θ_a
 $f(\theta_a)$ แทน ความหนาแน่นภายหลัง (prior density) ของ θ_a

แต่เนื่องจาก $f(u | \theta_a)$ เป็นฟังก์ชันความเป็นไปได้ (likelihood function) ของคะแนน ดังนั้น จึงสามารถเขียนสมการ $f(u | \theta_a)$ ให้อยู่ในรูปของการแปรผันได้ ดังนี้

$$f(\theta_a | u) \propto L(u | \theta_a) f(\theta) \text{ หรือ Posterior} \propto \text{Likelihood} \times \text{Prior}$$

วิธีการประมาณค่าความสามารถของผู้สอบตามวิธีการของเบส์นี้ ได้ถูกนำไปประยุกต์เป็นวิธีต่าง ๆ อีกหลายวิธี แต่วิธีที่ได้รับความนิยมมาก คือ วิธีของเบส์ที่ปรับใหม่ (bayesian updating) ซึ่งเสนอโดยโอเวน (รังสรรค์ มณีเล็ก. 2540 ; อ้างอิงจาก Owen. 1969 ; 1975 : 353) ซึ่งมีวิธีการประมาณค่า ดังนี้

ถ้าตอบถูกต้อง

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

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

ถ้าตอบผิด

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

$$\sigma_{m+1}^2 = \sigma_m^2 \left[1 - \left[\frac{\Theta(D)}{1 + \frac{1}{a^2 \sigma_m^2}} \right] \left[\frac{\Theta(D)}{\Phi(D)} - D \right] / [\Phi(D)] \right]$$

- เมื่อ
- ๑ (D) แทน ค่าออร์ดิเนท (ordinate) ของโค้งปกติที่จุด D
 - ๑ (D) แทน พื้นที่ใต้โค้งปกติ จากค่า D ต่ำสุด จนถึงจุด D ที่มีค่า เท่ากับ $\frac{b - \theta_m}{\sqrt{\frac{1}{a^2} + \sigma_m^2}}$ และ $A = c + (1 - c)\Phi(-D)$
 - θ_m แทน ความสามารถของผู้สอบที่ประมาณค่าได้ก่อนตอบข้อสอบข้อที่ $m + 1$ ซึ่งตามปกติแล้ว ถ้าไม่ทราบค่าความสามารถเบื้องต้นของผู้สอบ ก็มักกำหนดให้เท่ากับ 0 ซึ่งจากสูตรเดิมใช้สัญลักษณ์ M_0
 - σ_m^2 แทน ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบก่อนตอบข้อที่ $m + 1$ ถ้าไม่ทราบค่าความแปรปรวนดังกล่าวมาก่อน ก็มักกำหนดให้เท่ากับ 1 ซึ่งจากสูตรเดิมใช้สัญลักษณ์ V_0
 - θ_{m+1} แทน ค่าความสามารถของผู้สอบโดยประมาณหลังจากที่ตอบข้อที่ $m + 1$ ซึ่งจากสูตรเดิมใช้สัญลักษณ์ $E(\theta|1)$
 - σ_{m+1}^2 แทน ค่าความแปรปรวนในการประมาณค่าความสามารถของผู้สอบ เมื่อตอบข้อสอบข้อที่ $m + 1$ ซึ่งจากสูตรเดิมใช้สัญลักษณ์ Var
 - a แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ $m + 1$ ซึ่งจากสูตรเดิมใช้สัญลักษณ์ p
 - b แทน ค่าความยากง่ายของข้อสอบข้อที่ $m + 1$ ซึ่งจากสูตรเดิมใช้สัญลักษณ์ d
 - c แทน ค่าการเดาของข้อสอบข้อที่ $m + 1$ ซึ่งจากสูตรเดิมใช้สัญลักษณ์ g

การประมาณค่าความสามารถของผู้สอบด้วยวิธีเบส์ที่ปรับเปลี่ยน (Bayesian updating) ซึ่งเสนอโดยโอเวน (Owen) เป็นวิธีการประมาณค่าความสามารถของผู้สอบที่ให้ค่าเที่ยงตรงเชิงสภาพสูง (ริงสรรค มณีเล็ก. 2540) ดังนั้นการศึกษาวิจัยในครั้งนี้ ผู้วิจัยจึงเลือกใช้วิธีการประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์ที่ปรับเปลี่ยน ดังกล่าว

3. งานวิจัยที่เกี่ยวข้อง

3.1 งานวิจัยเกี่ยวกับการสร้างแบบทดสอบ

รังสรรค์ มณีเล็ก (2540) ได้ทำการสร้างเครื่องมือที่ใช้สำหรับสร้างคลังข้อสอบในการทำวิจัยเรื่อง ผลของตัวแปรบางตัวต่อความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ได้แก่ แบบทดสอบวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 มีลักษณะเป็นแบบทดสอบวัดความเข้าใจเรื่องเศษส่วน เนื่องจากแบบทดสอบดังกล่าว มีกระบวนการสร้างอย่างเป็นระบบ มีคุณภาพที่น่าเชื่อถือ ผู้วิจัยขอนำมากล่าวถึงกระบวนการสร้าง โดยสรุปดังต่อไปนี้

3.1.1 กระบวนการสร้างแบบทดสอบ

กระบวนการสร้างแบบทดสอบนั้น รังสรรค์ มณีเล็ก ได้เริ่มจากการวิเคราะห์หลักสูตร กล่าวคือศึกษาจุดประสงค์ เนื้อหาและเวลาของการเรียนการสอนคณิตศาสตร์ระดับประถมศึกษา จากหลักสูตรประถมศึกษา และคู่มือครูคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 กระทรวงศึกษาธิการ ศึกษาสมรรถภาพที่ใช้วัดทางคณิตศาสตร์พบว่าสมรรถภาพด้านความเข้าใจเรื่องเศษส่วนเป็นสมรรถภาพที่ค่อนข้างมีปัญหาสำหรับนักเรียน จึงได้นำสมรรถภาพด้านนี้มาสร้างแบบทดสอบ จากนั้นกำหนดลักษณะของเครื่องมือวัดโดยกำหนดรูปแบบของข้อสอบเป็นแบบปรนัยชนิดเลือกตอบ 4 ตัวเลือก ให้คะแนนแบบ 0, 1 กำหนดวัดจุดประสงค์ย่อย สร้างฟาเซต (Facet) ของคำถาม และตัวเลือก เพื่อเป็นกรอบในการสร้างข้อสอบ แล้วทำการสร้างข้อสอบเพื่อวัดสมรรถภาพความเข้าใจเรื่องเศษส่วน จำนวน 290 ข้อ โดยแยกออกเป็น 7 ฉบับ ฉบับละ 50 ข้อ เป็นข้อสอบคู่ขนานกัน เพราะเป็นข้อสอบที่สร้างจากฟาเซตเดียวกัน แต่ละฉบับมีข้อสอบต่างจากฉบับอื่น 40 ข้อ เหมือนกันทุกฉบับอยู่ 10 ข้อเพื่อใช้เป็น ข้อสอบร่วม (Anchor item) สำหรับเทียบคะแนนของข้อสอบแต่ละฉบับให้อยู่บนสเกลเดียวกัน จำนวน ข้อสอบรวมอย่างน้อยต้องมีจำนวนร้อยละ 20 ของข้อสอบทั้งฉบับตามข้อเสนอแนะของ ลอร์ด (Lord)

3.1.2 การตรวจสอบคุณภาพของเครื่องมือ

รังสรรค์ มณีเล็ก นำข้อสอบไปให้ผู้เชี่ยวชาญด้านหลักสูตร การสอน การวัดและประเมินผลการศึกษา ตรวจสอบคุณภาพด้านความเที่ยงตรงเชิงเนื้อหาและความเที่ยงตรงเชิงโครงสร้างตามเกณฑ์พิจารณา 3 เกณฑ์ คือ ตรง ไม่ตรง และไม่แน่ใจว่าตรงตามเนื้อหา และความเที่ยงตรงตามโครงสร้างที่ใช้วัดหรือไม่ ของโรวินลลี และแฮมเบิลตัน (Rovinelli and hambleton) ซึ่งเป็นข้อสอบวัดเนื้อหาเรื่องเศษส่วน วัดโครงสร้างความเข้าใจ แล้วจึงปรับปรุงข้อสอบตามคำแนะนำของผู้เชี่ยวชาญซึ่งเป็นการปรับเกี่ยวกับภาษาที่ใช้และรูปภาพที่ใช้แทนเศษส่วน จากนั้นนำข้อสอบไปเก็บข้อมูลกับนักเรียนชั้นประถมศึกษาปีที่ 6 ในโรงเรียนที่ได้รับการสุ่มเป็นกลุ่มตัวอย่าง คือโรงเรียนในสังกัดกรุงเทพมหานคร และสังกัดสำนักงานการประถมศึกษากรุงเทพมหานคร จำนวนกลุ่มตัวอย่างทั้งสิ้น 8,720 คน วิธีการสอบจัดทำโดยการแจกแบบทดสอบทั้ง 7 ฉบับให้กับนักเรียนอย่างเป็นระบบ กล่าวคือ คนที่ 1 – 7 จะได้แบบทดสอบฉบับที่ 1 – 7 ตามลำดับ คนที่ 8 – 14 จะได้รับแบบทดสอบฉบับที่ 1 – 7 เช่นเดียวกัน และกระทำเช่นนี้ไปเรื่อย ๆ จนครบนักเรียนทุกคนที่อยู่ในห้อง แล้วนำผลการสอบมาวิเคราะห์เพื่อหาคุณภาพของข้อสอบ โดยวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ

เพื่อตรวจสอบความเป็นมิติเดียวด้วยการวิเคราะห์องค์ประกอบ (Factor analysis) หมุนแกนโดยวิธี แวริแมกซ์ (Varimax) สกัดตัวประกอบโดยวิเคราะห์องค์ประกอบหลัก (Principal component) พิจารณาค่าไอเกนของตัวประกอบที่ 1 ตั้งแต่ 3 เท่าขึ้นไป และค่าไอเกนของตัวประกอบที่ 2 มีค่าใกล้เคียงกับตัวประกอบที่เหลือ หรือพิจารณาตัวประกอบที่สามารถอธิบายความแปรปรวนร่วมได้ร้อยละ 70 ขึ้นไป วิเคราะห์คุณภาพของข้อสอบรายข้อทั้ง 7 ฉบับเพื่อหาค่าพารามิเตอร์ของข้อสอบ คือค่าอำนาจจำแนก (a) ค่าความยากง่าย (b) และค่าการเดา (c) โดยใช้โมเดลโลจิสติกแบบ 3 พารามิเตอร์ โดยการใช้โปรแกรม BILOG3 แปลงค่าพารามิเตอร์ของข้อสอบ แต่ละฉบับให้อยู่บนสเกลเดียวกัน (Common scale) โดยเทียบคะแนนตามแบบแผนการใช้ข้อสอบร่วม (Anchor item design) หาค่าคงที่สำหรับการแปลงคะแนนด้วยวิธีค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐาน (Mean and Sigma method) ทำการคัดเลือกข้อสอบที่มีคุณภาพตามเกณฑ์ของยูริ (Urry, 1977) คือข้อสอบที่มีค่าอำนาจจำแนกตั้งแต่ 0.8 ขึ้นไป ค่าความยากง่ายอยู่ในช่วง -2.0 ถึง 2.0 ค่าการเดาน้อยกว่า 0.3 ได้ข้อสอบทั้งหมด 150 ข้อ แต่เพื่อให้ได้ข้อสอบครอบคลุมความสามารถของผู้สอบที่เป็นกลุ่มตัวอย่างในการวิจัย จึงได้คัดเลือกข้อสอบที่เหลือเพิ่มอีก 50 ข้อ ดังนั้น ค่าพารามิเตอร์ค่าอำนาจจำแนกจึงเริ่มที่ 0.4 ขึ้นไป ค่าความยากอยู่ในช่วง -3.0 ถึง 3.0 และค่าการเดา ต่ำกว่า 0.3 ซึ่งทำให้ได้ข้อสอบทั้งหมด 200 ข้อ เพื่อจัดเก็บไว้ในคลังข้อสอบ โดยใช้โปรแกรมคอมพิวเตอร์สำหรับการสร้างคลังซึ่งจัดเก็บไว้ในไฟล์ 3 ไฟล์ คือ ไฟล์ตัวอักษร (test.dat) ไฟล์รูปภาพ (t.pic) และไฟล์พารามิเตอร์ (compute.dat) นอกจากนี้ยังได้เก็บค่าความสามารถที่ทำให้ข้อสอบมีค่าสารสนเทศสูงสุด (Qmax) ของแต่ละข้อไว้ด้วย เพื่อนำไปใช้ในการประมาณค่าความสามารถของผู้สอบและใช้สำหรับการคัดเลือกข้อสอบจากคลังมาเสนอแก่ผู้สอบบนหน้าจอคอมพิวเตอร์

3.1.3 คุณภาพของเครื่องมือวัด

คุณภาพของแบบทดสอบวิชาคณิตศาสตร์ วัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 ที่ยังไม่ได้ปรับสเกลการวัดนั้นมีค่าพารามิเตอร์ของข้อสอบโดยเฉลี่ยค่าใกล้เคียงกัน กล่าวคือค่าอำนาจจำแนกประมาณ 1.0 ซึ่งอยู่ในเกณฑ์ดี ค่าความยากประมาณ 0.8 ซึ่งอยู่ในเกณฑ์ปานกลางค่อนข้างไปทางยาก และค่าการเดามีค่าประมาณ 0.2 ซึ่งอยู่ในเกณฑ์ดี ส่วนผลการตรวจสอบความเป็นมิติเดียวของข้อสอบวัดความเข้าใจเรื่องเศษส่วนทั้ง 7 ฉบับ โดยการวิเคราะห์ตัวประกอบปรากฏว่า ค่าไอเกน (Eigen value) ของตัวประกอบที่ 1 สูงกว่าค่าไอเกนของตัวประกอบอื่น ๆ ที่เหลือ และตัวประกอบที่เหลือดังกล่าวมีค่าไอเกนใกล้เคียงกัน จึงอนุมานได้ว่า ข้อสอบวัดความเข้าใจเรื่องเศษส่วนชั้นประถมศึกษาปีที่ 6 ของรังสรรค์ มณีเล็ก มีความเป็นมิติเดียว ตรงตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

3.2 งานวิจัยเกี่ยวกับการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะ

งานวิจัยเกี่ยวกับการพัฒนาโปรแกรมสำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์นั้น ได้มีผู้พัฒนาโปรแกรมไว้เป็นที่น่าสนใจ 2 โปรแกรม ผู้วิจัย ขอนำมากล่าวถึงตัวโปรแกรม คุณภาพของโปรแกรม และผลการศึกษาจากการใช้โปรแกรม โดยสรุป ดังต่อไปนี้

3.2.1 โปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบของรั้งสรรพค์ มณีเล็ก (2540) ซึ่งทำการสร้างโปรแกรมคอมพิวเตอร์ขึ้น และศึกษาผลของตัวแปรคือค่าความยากง่ายของข้อสอบข้อแรก เกณฑ์การคัดเลือกข้อสอบจากคลังข้อสอบ วิธีการประมาณค่าความสามารถของผู้สอบ และตัวแปรเกณฑ์ยุติการทดสอบ ต่อความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการทดสอบ การสร้างโปรแกรมการทดสอบโดยใช้ภาษาซี (C Language) ในการเขียนโปรแกรม เนื่องจากภาษาซีมีความยืดหยุ่น ง่ายต่อการใช้คำสั่งและเกิดการผิดพลาดน้อย โปรแกรมสามารถทำงานได้บนระบบปฏิบัติการ DOS ซึ่งมีขั้นตอนในการสร้างโปรแกรม ดังนี้

1) เขียนผังงานแสดงขั้นตอนการทำงานของคอมพิวเตอร์ในการดำเนินการสอบรวมทั้งการบันทึกผลการสอบ มีทั้งหมด 18 ผังงาน ตามจำนวนที่ได้จากการจัดหมู่ตัวแปรอิสระ 3 ตัวแปร คือ ความยากง่ายของข้อสอบข้อแรก 3 ระดับ เกณฑ์การคัดเลือกข้อสอบจากคลังข้อสอบ 3 เกณฑ์ วิธีการประมาณค่าความสามารถของผู้สอบ 2 วิธี นอกจากนั้นยังได้เขียนโปรแกรมสั่งให้คอมพิวเตอร์บันทึก ค่าของตัวแปรตาม จำแนกตามเกณฑ์สำหรับยุติการทดสอบ 3 เกณฑ์

2) การเขียนโปรแกรมในครั้งนี้ผู้วิจัยได้เลือกใช้ภาษาซี (C language) ในการเขียนโปรแกรม ทั้งนี้เนื่องจากภาษาซีมีความยืดหยุ่น ง่ายต่อการใช้คำสั่งและเกิดการผิดพลาดน้อยและนำโปรแกรม BTRIEVE ของ NETWARE เข้ามาช่วยในการจัดระบบข้อมูลด้วย

3) เขียนโปรแกรมคอมพิวเตอร์ตามผังงานที่ได้กำหนดไว้ 18 ผังงาน ซึ่งโปรแกรมที่เขียนขึ้นตามโครงสร้างของตัวแปรที่ศึกษานั้น มีบางส่วนของโปรแกรมที่ต้องมีการประมาณค่าต่างๆ ได้แก่

3.1) ความน่าจะเป็นในการตอบข้อสอบได้ถูกต้อง

3.2) ค่าความสามารถของผู้สอบ วิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข

(conditional maximum likelihood) และ วิธีของเบส์ (Bayesian) โดยเลือกใช้วิธีของโอเวน (Owen)

4) ทดลองใช้โปรแกรมคอมพิวเตอร์สำหรับดำเนินการสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์กับนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ของปีการศึกษา 2538 ประมาณ 600 คน จากโรงเรียนวัดพลับพลาชัย โรงเรียนดาราธรรม โรงเรียนราชวินิต โรงเรียนวัดจันทน์ โรงเรียนการเคหะท่าทรายและโรงเรียนประชานิเวศน์ ปรับปรุงแก้ไขข้อบกพร่องของโปรแกรมคอมพิวเตอร์ให้สมบูรณ์มากยิ่งขึ้น

ส่วนรูปแบบของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ในการวิจัยครั้งนี้ผู้วิจัยเลือกรูปแบบของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แบบหลายขั้นตอน (multi-stage strategies) แบบแยกทางแปรผัน (variable branching model) ซึ่งในการทดสอบตามแบบนี้ไม่สามารถกำหนดจำนวนขั้นของการตอบได้อย่างแน่นอนว่าผู้สอบจะต้องตอบกี่ขั้นก็ข้อ ทั้งนี้เนื่องจากผู้วิจัยได้ใช้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบเป็นเกณฑ์สำหรับยุติการทดสอบ

สำหรับผลของตัวแปรที่ศึกษาพบว่า การเริ่มต้นสอบด้วยข้อสอบที่ง่าย ปานกลาง หรือยากไม่ทำให้ความเที่ยงตรงเชิงสภาพแตกต่างกัน แต่ควรเริ่มต้นสอบด้วยข้อสอบที่มีความยากง่ายปานกลาง เพราะจะใช้จำนวนข้อสอบน้อยกว่า

การคัดเลือกข้อสอบข้อต่อไปในคลังข้อสอบ 3 วิธี คือ ใช้ค่าความสามารถของผู้สอบจากการทำข้อสอบที่ผ่านไปแล้วเทียบกับค่าความยากง่ายของข้อสอบในคลัง วิธีใช้ค่าความสามารถของผู้สอบจากการทำข้อสอบที่ผ่านไปแล้วเทียบกับค่าความสามารถที่ทำให้ข้อสอบมีสารสนเทศสูงสุด และใช้ค่าสารสนเทศของข้อสอบที่สูงที่สุด พบว่า การคัดเลือกข้อสอบจาก 3 วิธี ไม่ทำให้ค่าความเที่ยงตรงเชิงสภาพแตกต่างกัน แต่วิธีที่ 2 และ 3 จะใช้ข้อสอบน้อยกว่าวิธีแรก และพบว่าวิธีที่ 2 เหมาะสำหรับคลังข้อสอบขนาดเล็ก วิธีที่ 3 เหมาะกับคลังข้อสอบขนาดใหญ่

การประมาณค่าความสามารถของผู้สอบจากการศึกษาเปรียบเทียบวิธีการประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์และวิธีความเป็นไปได้สูงสุดพบว่าวิธีการประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์มีความเที่ยงตรงเชิงสภาพสูงกว่าการประมาณค่าความสามารถด้วยวิธีความเป็นไปได้สูงสุด

สำหรับเกณฑ์ยุติการทดสอบนั้น พบว่าเกณฑ์ยุติการสอบต่างกันทำให้ความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการสอบต่างกัน ถ้าใช้เกณฑ์ค่าความคลาดเคลื่อนมาตรฐานน้อยกว่าหรือเท่ากับ .30 จะใช้จำนวนข้อสอบมากกว่าการใช้เกณฑ์ค่าความคลาดเคลื่อนมาตรฐานเท่ากับหรือน้อยกว่า .40 อยู่ถึง 2 เท่า ดังนั้น ถ้าเป็นการทดสอบโดยทั่วๆ ไปจึงควรใช้เกณฑ์ค่าความคลาดเคลื่อนมาตรฐาน .40 จะได้ค่าความเชื่อมั่นในการวัดเท่ากับ .9165 แต่ใช้ข้อสอบจำนวนน้อยลงกว่าครึ่ง

3.2.2 โปรแกรมสำหรับการทดสอบแบบปรับเหมาะสำหรับระดับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ของสายชล อบทม (2539) ซึ่งได้พัฒนาโปรแกรมนี้ขึ้นนั้น เป็นโปรแกรมที่สามารถจัดพิมพ์ข้อสอบ รูปแบบเลือกตอบ ที่ผ่านการวิเคราะห์ตามแนวทฤษฎีการตอบสนองข้อสอบ (IRT) ที่มีค่าอำนาจจำแนก (a) ในช่วง 0.0 ถึง 2.5 ค่าความยากอยู่ในช่วง -3.0 ถึง 3.0 และ ค่าการเดา (c) ในช่วง 0 ถึง 0.3 เป็นโปรแกรมการทดสอบแบบปรับเหมาะสำหรับระดับความสามารถของผู้สอบที่ใช้รูปแบบทางแยกแปรผัน ประมาณค่าความสามารถของผู้สอบโดยวิธีของเบส์ ยุติการสอบเมื่อค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบต่ำกว่า 0.3 และโปรแกรมสามารถใช้กับระบบการปฏิบัติงานของวินโดว 95 ซึ่งมีขั้นตอนการพัฒนาพอสรุปได้ดังนี้

ขั้นตอนที่ 1 การกำหนดปัญหา เนื่องจากการทดสอบแบบมาตรฐานเดิมมักมีปัญหาในเรื่องค่าสถิติของข้อสอบ และของแบบทดสอบที่ไม่คงที่แต่จะแปรเปลี่ยนไปตามกลุ่มผู้สอบจึงทำให้การประมาณค่าความสามารถของผู้สอบให้ใกล้เคียงตามความเป็นจริงได้ยาก การทดสอบแบบปรับเหมาะสำหรับระดับความสามารถของผู้สอบที่อาศัยทฤษฎีการตอบสนองข้อสอบ มาเป็นพื้นฐานในการคำนวณค่าพารามิเตอร์ต่าง ๆ ของข้อสอบและแบบทดสอบ สามารถประมาณค่าความสามารถของผู้สอบได้ใกล้เคียงตามความเป็นจริงมากกว่า จึงได้มีการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ขึ้นใช้

ขั้นตอนที่ 2 วิเคราะห์ปัญหาและสิ่งที่ต้องการ สายชล อบทม ได้ศึกษารูปแบบการทดสอบแบบปรับเหมาะสำหรับระดับความสามารถของผู้สอบที่พัฒนาอยู่ทั่วไป และศึกษาเพื่อกำหนดขอบเขตของพารามิเตอร์ของข้อสอบ และของผู้ทดสอบเพื่อให้แต่ละค่าที่กำหนดเป็นมาตรฐาน

จากการศึกษาได้พบว่า มีการพัฒนาโปรแกรมคอมพิวเตอร์เพื่อใช้ในการทดสอบแบบปรับเหมาะโดยปฏิบัติการบนระบบ DOS สายชล ออบทม จึงทำการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะให้สามารถปฏิบัติการบนระบบ Windows และกำหนดขอบเขตของพารามิเตอร์ของข้อสอบและของผู้ทดสอบไว้ดังนี้ คือ การกำหนดจำนวนตัวเลือกตอบสามารถกำหนดใช้ตัวเลือกตอบได้ 3 ถึง 5 ตัวเลือก รูปแบบของตัวเลือกมี 3 ลักษณะกล่าวคือแบบ ก, ข, ค แบบ a, b, c และแบบ 1,2,3 กำหนดค่าอำนาจจำแนก (a) ที่นำมาใช้กับโปรแกรมได้ตั้งแต่ 0.0 ถึง 2.5 ค่าความยาก (b) อยู่ในช่วง -3.0 ถึง 3.0 ค่าการเดาค่ากว่า 0.3 ค่าถามและคำตอบแต่ละข้อสามารถพิมพ์ได้ถึง 1,000 ตัวอักษร นอกจากนี้ยังสามารถพิมพ์ชื่อผู้เข้าสอบ ระดับชั้น โรงเรียน วิชาที่สอบ และรหัสผ่านไว้ด้วย

ขั้นตอนที่ 3 การออกแบบโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะ การออกแบบโปรแกรมสำหรับการทดสอบแบบปรับเหมาะประกอบด้วยโลโก้ของโปรแกรม ออกแบบหน้าจอโปรแกรม หน้าจอของโปรแกรมย่อยการพิมพ์ข้อสอบซึ่งมีกลุ่มของคำสั่งให้โปรแกรมทำงาน เช่นคำสั่งเพิ่มข้อสอบ ข้อความ รูปภาพ วิธีใช้ กลุ่มคำสั่งเปิดบันทึกปิด ยกเลิกการทำงาน กลุ่มคำสั่งค่าพารามิเตอร์ของข้อสอบและแบบทดสอบ ออกแบบหน้าจอโปรแกรมย่อยการทดสอบ ซึ่งมีกลุ่มคำสั่งต่าง ๆ ที่ใช้ในการทดสอบเช่นเดียวกันกับที่กล่าวข้างต้น และการสร้างผังงานเพื่อกำหนดขอบเขตการทำงานของโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะให้เป็นไปตามมาตรฐานที่กำหนด

ขั้นตอนที่ 4 เลือกภาษาคอมพิวเตอร์ โปรแกรมการทดสอบแบบปรับเหมาะกับระดับความสามารถของผู้สอบ ของสายชล ออบทม ได้เลือกใช้ภาษาปาสคาล (Pascal) ในการเขียนโปรแกรม เนื่องจากเป็นภาษาที่มีโครงสร้างชัดเจน ง่ายต่อการใช้งาน สามารถใช้กับเครื่องคอมพิวเตอร์ขนาดใหญ่ และเครื่องไมโครคอมพิวเตอร์ได้ เป็นภาษาที่สามารถเล่นกราฟฟิคได้ และสามารถเขียนสูตรสมการหรือฟังก์ชันคณิตศาสตร์ได้ โปรแกรมการทดสอบนี้ใช้กับรูปแบบการทดสอบแบบทางแยก แปรผัน ประมาณค่าพารามิเตอร์ตามความสามารถของผู้สอบด้วยวิธีเบย์ส์ (Bayesian Estimation) โปรแกรมจะยุติการสอบเมื่อค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบต่ำกว่า 0.3 และคัดเลือกข้อสอบข้อต่อไปโดยวิธีดูค่าสารสนเทศของข้อสอบที่สูงที่สุดมาเป็นข้อสอบข้อต่อไป

ขั้นตอนที่ 5 การทดสอบและแก้ไขโปรแกรม การทดสอบความคลาดเคลื่อนในการทำงานของโปรแกรมที่อาจเกิดขึ้นจากการเขียนภาษา โดยมีการทดสอบการทำงานของโปรแกรมกับเครื่องคอมพิวเตอร์แล้วทำการปรับปรุงแก้ไขจนไม่มีความคลาดเคลื่อนในการทำงาน และทำการทดสอบว่าโปรแกรมทำงานเป็นไปตามวัตถุประสงค์หรือไม่อีกด้วย

ขั้นตอนที่ 6 จัดทำเอกสารประกอบการใช้โปรแกรม จัดทำเอกสารอธิบายลักษณะของโปรแกรม การติดตั้ง การเรียกใช้ การพิมพ์ข้อสอบและการทดสอบ ตลอดจนการอภิปรายผลการสอบ

ขั้นตอนที่ 7 ประเมินคุณภาพของโปรแกรม การประเมินคุณภาพของโปรแกรม ทำการประเมินจากผู้พัฒนาโปรแกรมเป็นผู้ประเมินเอง และให้ผู้ใช้โปรแกรมเป็นผู้ทำการประเมินอีกทางหนึ่งด้วย

คุณภาพของโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะของสายชล ออบทม โปรแกรมการทดสอบนี้สามารถทำงานร่วมกับเครื่องไมโครคอมพิวเตอร์ ที่ปฏิบัติการบนระบบวินโดวส์

95 โดยโปรแกรมสามารถใช้โปรแกรมย่อยพิมพ์ข้อสอบแบบเลือกตอบ 3 – 5 ตัวเลือก วิเคราะห์คุณภาพของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ ค่าพารามิเตอร์ของข้อสอบจะต้องมีค่าความยาก (b) อยู่ระหว่าง -3.0 ถึง 3.0 ค่าอำนาจจำแนก (a) อยู่ระหว่าง 0 ถึง 2.5 และค่าการเดา (c) ต่ำกว่า 0.3 ใช้กับรูปแบบการทดสอบแบบทางแยกแปรผัน ประเมินค่าความสามารถของผู้สอบด้วยวิธีของเบส์ ยุติการสอบเมื่อค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบต่ำกว่า 0.3 และผลการประเมินจากผู้ใช้โปรแกรมพบว่าโปรแกรมมีความสะดวกในการใช้งาน รูปแบบของโปรแกรมมีความเหมาะสมดี การทดสอบมีความทันสมัย นักเรียนมีความกระตือรือร้นและให้ความสนใจการทดสอบมากยิ่งขึ้น

เนื่องจากผลการศึกษาวิจัย ของรังสรรค์ มณีเล็ก ที่ค้นพบผลของตัวแปรที่ส่งผลต่อคุณภาพของการทดสอบแบบปรับเหมาะ กล่าวคือถ้าเริ่มต้นสอบด้วยข้อสอบที่มีความยากง่ายปานกลางจะทำให้ใช้จำนวนข้อสอบน้อยกว่าการเริ่มต้นจากข้อยากหรือข้อง่าย วิธีคัดเลือกข้อสอบข้อต่อไปในคลังข้อสอบ ควรใช้ค่าสารสนเทศของข้อสอบที่สูงที่สุดจะทำให้ใช้ข้อสอบน้อยลง และเหมาะกับคลังข้อสอบขนาดใหญ่ การประมาณค่าความสามารถของผู้สอบด้วยวิธีเบส์ให้ค่าความเที่ยงตรงเชิงสภาพสูงกว่า และเกณฑ์การยุติการสอบ ควรใช้เกณฑ์ค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.3 และจากการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบของ สายชล ออบทม ซึ่งได้พัฒนาโปรแกรมอย่างมีขั้นตอน เป็นระบบ โปรแกรมสามารถรองรับการทำงานในระบบปฏิบัติการวินโดวส์ 95 และผลจากการประเมินการใช้โปรแกรมแสดงถึงคุณภาพของโปรแกรมอยู่ในเกณฑ์ที่มีคุณภาพ และมีประสิทธิภาพ กล่าวคือ โปรแกรมการทดสอบสามารถรองรับข้อสอบแบบเลือกตอบ ที่มีตัวเลือก 3 ถึง 5 ตัวเลือก รองรับข้อสอบที่ผ่านการวิเคราะห์ค่าพารามิเตอร์แบบ 3 พารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบ (IRT) และมีกลุ่มของคำสั่งต่าง ๆ ของโปรแกรมเพื่อใช้ในการทำงานได้อย่างสะดวก มีความทันสมัย ผู้สอบมีความสนใจในการทดสอบมากยิ่งขึ้น ดังนั้นผู้วิจัยจึงเลือกข้อค้นพบของรังสรรค์ มณีเล็ก และนำแนวทางการพัฒนาโปรแกรมการทดสอบของสายชล ออบทม มาพัฒนาค้นคลังข้อสอบ และโปรแกรมการทดสอบที่ใช้ในการค้นคว้าของผู้วิจัย ซึ่งผู้วิจัยได้สร้างข้อกำหนดคุณลักษณะของข้อสอบใหม่ที่คาดว่าจะส่งผลต่อคุณภาพของการทดสอบแบบปรับเหมาะยิ่งขึ้น คือ การจัดชุดข้อสอบให้ตอบในแต่ละครั้งต่างกัน การใช้เวลาในการตอบแต่ละข้อต่างกัน และการสร้างรูปแบบของข้อสอบที่มีสื่อประสมกับไม่มีสื่อประสม

3.3 งานวิจัยเกี่ยวกับการพัฒนาโปรแกรมคอมพิวเตอร์มัลติมีเดีย

สำหรับการสร้างรูปแบบของข้อสอบที่มีสื่อประสมประกอบการทดสอบในครั้งนี้ ผู้วิจัยได้คัดเลือกสื่อประสมที่ใช้กับแบบทดสอบวัดผลการเรียนหลังการใช้คอมพิวเตอร์มัลติมีเดียแบบฝึกของศิริพงษ์ ภูพันนา (2546) ที่ได้พัฒนาบทเรียนคอมพิวเตอร์มัลติมีเดียแบบฝึก และได้พัฒนาคอมพิวเตอร์มัลติมีเดียสำหรับการทดสอบหลังการใช้แบบฝึกขึ้นพร้อม ๆ กันนี้ด้วย ซึ่งมีขั้นตอนในการพัฒนาพอสรุปได้ดังต่อไปนี้

ขั้นที่ 1 การออกแบบ (Design)

เป็นการศึกษาค้นคว้าและวิเคราะห์ข้อมูลจากเอกสารงานวิจัยทั้งในประเทศและต่างประเทศ สรุปเป็นหลักการ แนวคิด ทฤษฎี แล้วนำเสนอผู้เชี่ยวชาญตรวจสอบความเหมาะสมและความเป็นไปได้ของแบบจำลองบทเรียนคอมพิวเตอร์มัลติมีเดียแบบฝึกและคอมพิวเตอร์มัลติมีเดียสำหรับการทดสอบหลังการใช้แบบฝึก โดยขอความคิดเห็นจากผู้เชี่ยวชาญด้านเนื้อหาและด้านเทคโนโลยีทางการศึกษา จากนั้นผู้วิจัยได้นำเอาหลักการ แนวคิด ทฤษฎี และความคิดเห็นของผู้เชี่ยวชาญสังเคราะห์เป็นแบบจำลองบทเรียนคอมพิวเตอร์มัลติมีเดียแบบฝึกและคอมพิวเตอร์มัลติมีเดียสำหรับการทดสอบ ดังกล่าว

ขั้นที่ 2 การสร้างและพัฒนา (Production and Development)

นำเอาแบบจำลองบทเรียนคอมพิวเตอร์มัลติมีเดียแบบฝึกและคอมพิวเตอร์มัลติมีเดียสำหรับการทดสอบหลังการใช้แบบฝึก มาดำเนินการสร้างและพัฒนาโดยใช้เนื้อหาวิชาคณิตศาสตร์ เรื่อง ความรู้เบื้องต้นเกี่ยวกับเรขาคณิตวิเคราะห์ มาสร้างเป็นต้นแบบบทเรียนคอมพิวเตอร์มัลติมีเดียแบบฝึกดังกล่าวและนำไปผ่านการตรวจสอบและประเมินความเหมาะสมจากผู้เชี่ยวชาญทั้งด้านเนื้อหาวิชาและด้านเทคโนโลยีทางการศึกษา ดำเนินการปรับปรุงและแก้ไขตามที่ผู้เชี่ยวชาญแนะนำ แล้วจึงนำไปทดลองครั้งที่ 1 โดยทดลองใช้กับนักเรียนที่มีระดับความสามารถทางการเรียนต่างกัน จำนวน 3 คน ต่อบทเรียนแต่ละรูปแบบ รวม 9 คน สังเกตและวิเคราะห์หาข้อบกพร่องและความเหมาะสมของเวลาในการนำเสนอ ดำเนินการทดลองครั้งที่ 2 กับนักเรียนที่มีระดับความสามารถทางการเรียนต่างกัน จำนวน 9 คน ต่อบทเรียนแต่ละรูปแบบ รวม 27 คน ทำการวิเคราะห์หาแนวโน้มประสิทธิภาพของบทเรียน และดำเนินการทดลองครั้งที่ 3 กับนักเรียนที่มีความสามารถทางการเรียนต่างกัน จำนวน 36 คน ต่อบทเรียนแต่ละรูปแบบ รวม 108 คน และทำการวิเคราะห์หาประสิทธิภาพของบทเรียน

ขั้นที่ 3 การทดลอง (Implementation)

นำบทเรียนแต่ละรูปแบบไปทดลองใช้กับกลุ่มตัวอย่างที่มีความสามารถทางการเรียนต่างกัน 3 กลุ่ม กลุ่มละ 15 คน รวมเป็นรูปแบบละ 45 คน รวมทั้งสิ้น 135 คน โดยใช้แบบแผนการวิจัยเชิงทดลอง (Experimental Research) แบบ 3 x 3 factorial design ดำเนินการเก็บรวบรวมข้อมูลเพื่อการวิเคราะห์และรายงานผลการทดลอง

การวิเคราะห์ข้อมูล

1. วิเคราะห์ประสิทธิภาพของต้นแบบบทเรียนคอมพิวเตอร์มัลติมีเดียแบบฝึกที่มีการควบคุมการเรียนต่างกัน 3 รูปแบบ โดยการวิเคราะห์หาประสิทธิภาพตามเกณฑ์ E_1/E_2
2. วิเคราะห์ความแปรปรวนของผลสัมฤทธิ์ทางการเรียนและความคงทนในการเรียนรู้ตามแบบแผนการวิจัยเชิงทดลอง 3 x 3 factorial design โดยใช้การวิเคราะห์ความแปรปรวนแบบ 2 ทาง (Two way Analysis of Variance)
3. วิเคราะห์เพื่อเปรียบเทียบความแตกต่างเป็นรายคู่ ในกรณีที่พบความแตกต่างของค่าคะแนนเฉลี่ยในแต่ละกลุ่ม โดยใช้การวิเคราะห์ผลต่างอย่างมีนัยสำคัญน้อยที่สุด (LSD)

สำหรับคุณภาพของคอมพิวเตอร์มัลติมีเดียและคอมพิวเตอร์มัลติมีเดียสำหรับทดสอบ หลังการใช้แบบฝึก มีดังนี้

1. ผู้เชี่ยวชาญด้านเนื้อหาวิชามีความเห็นว่า บทเรียนคอมพิวเตอร์มัลติมีเดียแบบฝึก และคอมพิวเตอร์มัลติมีเดียสำหรับการทดสอบหลังการใช้แบบฝึกมีคุณภาพอยู่ในเกณฑ์ดี
2. ผู้เชี่ยวชาญด้านเทคโนโลยีทางการศึกษา มีความเห็นว่า บทเรียนคอมพิวเตอร์มัลติมีเดียแบบฝึกและคอมพิวเตอร์มัลติมีเดียสำหรับการทดสอบหลังการใช้แบบฝึกมีคุณภาพอยู่ในเกณฑ์ดีมาก
3. การหาประสิทธิภาพตามเกณฑ์ E_1/E_2 ของบทเรียนคอมพิวเตอร์มัลติมีเดียแบบฝึกและคอมพิวเตอร์มัลติมีเดียสำหรับการทดสอบหลังการใช้บทเรียน ปรากฏว่ามีประสิทธิภาพตามเกณฑ์ E_1/E_2 (90.73/90.63)

เนื่องจากผลการวิจัยดังกล่าวได้พัฒนาสื่อประสมโดยผ่านการพิจารณาและประเมินผลจากผู้เชี่ยวชาญด้านเทคโนโลยีหลายท่าน ผู้วิจัยจึงเชื่อว่าสื่อประสมดังกล่าวมีคุณภาพและสามารถนำมาพัฒนาใช้กับการทดสอบแบบปรับเหมาะของผู้วิจัยได้ ดังนั้นการวิจัยในครั้งนี้ ผู้วิจัยจึงเลือกสื่อประสมจากผลการวิจัยดังกล่าวข้างต้นมาพัฒนาสร้างรูปแบบข้อสอบแบบปรับเหมาะที่มีสื่อประสมประกอบการทดสอบต่อไป

3.4 งานวิจัยเกี่ยวกับการทดสอบแบบปรับเหมาะ

จิราพร ไกรสรศิวเวท (2529) ได้ศึกษาความสัมพันธ์ระหว่างรูปแบบบางประการของการทดสอบเฉพาะบุคคลรูปแบบปิรามิดกับความสามารถทางการเรียน โดยใช้การทดสอบเฉพาะบุคคลรูปแบบปิรามิดที่มี 8 ชั้น 3 แบบ คือ แบบที่ให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ได้ทำ ให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ทำถูก และให้คะแนนทุกข้อ โดยทำการศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 348 คน ผลการศึกษาพบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากการทดสอบเฉพาะบุคคลรูปแบบปิรามิดทั้ง 3 แบบ โดยให้คะแนนทั้ง 3 วิธี กับคะแนนจากการทดสอบแบบมาตรฐานเดิม มีความสัมพันธ์กันในทางบวกอย่างมีนัยสำคัญทางสถิติทุกค่า เมื่อพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ภายใน ระหว่างคะแนนที่ได้จากการทดสอบเฉพาะบุคคลรูปแบบปิรามิดทั้ง 3 แบบ โดยให้คะแนนทั้ง 3 วิธี ปรากฏว่า มีค่าสัมพันธ์เป็นบวกอย่างมีนัยสำคัญทางสถิติทุกค่า

นันทิยา พึ่งคำ (2531) ได้ศึกษาเปรียบเทียบคุณภาพของการทดสอบแบบซี เอ ที (Computerized Adaptive Testing : CAT) กับแบบทดสอบมาตรฐานเดิม (Conventional Test) ในการวัดความสามารถด้านคำศัพท์ภาษาอังกฤษโดยทำการศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 120 คน การทดสอบแบบซี เอ ที ใช้กลุ่มข้อสอบ (Item Pool) จำนวน 361 ข้อ โดยมีค่าความยากอยู่ระหว่าง -0.90552 ถึง 3.61860 มีค่าอำนาจจำแนกอยู่ระหว่าง 0.80334 ถึง 2.0000 และมีค่าการเดาอยู่ระหว่าง 0.04421 ถึง 0.29886 การประมาณค่าความสามารถของผู้สอบโดยใช้วิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข (Conditional Maximum Likelihood) ส่วนการทดสอบแบบมาตรฐานเดิมใช้ข้อสอบทั้งหมด 40 ข้อ โดยมีค่าความยากอยู่ระหว่าง -0.76488 ถึง 2.97884 มีค่าอำนาจ

จำแนกอยู่ระหว่าง 0.94069 ถึง 2.0000 และมีค่าการเดาอยู่ระหว่าง 0.03462 ถึง 0.28620 ผลการศึกษาพบว่า การทดสอบทั้งสองแบบมีความตรงเชิงเกณฑ์สัมพันธ์ค่อนข้างสูง และมีความเที่ยงตรงไม่แตกต่างกัน ส่วนค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ (SEE) ของการทดสอบแบบซี เอ ที่ ต่ำกว่าการทดสอบแบบมาตรฐานเดิมอย่างมีนัยสำคัญทางสถิติ และจำนวนข้อสอบของการทดสอบแบบซี เอ ที่ น้อยกว่าครึ่งหนึ่งของการทดสอบแบบมาตรฐานเดิม

ตริงใจ พูนผลอำนวย (2534) ได้พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความระดับความสามารถของผู้สอบ โดยใช้รูปแบบพีรามิด ขนาดชั้นคงที่ 10 ชั้น โดยใช้คอมพิวเตอร์ ผลปรากฏว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ดังกล่าวให้ความสะดวกในการทดสอบมากยิ่งขึ้น ผู้สอบให้ความสนใจ และสร้างความดึงดูดใจทำให้ผู้สอบใส่ใจในการทำข้อสอบเพิ่มมากขึ้น

รังสรรค์ มณีเล็ก (2540) ทำการสร้างโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความความสามารถของผู้สอบและศึกษาผลของตัวแปรค่าความยากง่ายของข้อสอบข้อแรก เกณฑ์การคัดเลือกข้อสอบจากคลังข้อสอบ วิธีการประมาณค่าความสามารถของผู้สอบ และตัวแปรเกณฑ์ยุติการทดสอบ ต่อความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการทดสอบ การสร้างโปรแกรมการทดสอบเป็นโปรแกรมที่ใช้ภาษาซี (C Language) ในการเขียนโปรแกรม เนื่องจากภาษาซีมีความยืดหยุ่น ง่ายต่อการใช้คำสั่งและเกิดการผิดพลาดน้อย สำหรับผลของตัวแปรที่ศึกษาพบว่า การเริ่มต้นสอบด้วยข้อสอบที่ง่าย ปานกลาง หรือยาก ไม่ทำให้ความเที่ยงตรงเชิงสภาพแตกต่างกัน แต่ควรเริ่มต้นสอบด้วยข้อสอบที่มีความยากง่ายปานกลาง เพราะจะใช้จำนวนข้อสอบน้อยกว่า การคัดเลือกข้อสอบข้อต่อไปในคลังข้อสอบ 3 วิธี คือ ใช้ค่าความสามารถของผู้สอบจากการทำข้อสอบที่ผ่านไปแล้ว เทียบกับค่าความยากง่ายของข้อสอบในคลัง วิธีใช้ค่าความสามารถของผู้สอบจากการทำข้อสอบที่ผ่านไปแล้วเทียบกับค่าความสามารถที่ทำให้ข้อสอบมีสารสนเทศสูงสุด และใช้ค่าสารสนเทศของข้อสอบที่สูงที่สุด พบว่า การคัดเลือกข้อสอบจาก 3 วิธี ไม่ทำให้ค่าความเที่ยงตรงเชิงสภาพแตกต่างกัน แต่วิธีที่ 2 และ 3 จะใช้ข้อสอบน้อยกว่าวิธีแรก และพบว่าวิธีที่ 2 เหมาะสำหรับคลังข้อสอบขนาดเล็ก วิธีที่ 3 เหมาะกับคลังข้อสอบขนาดใหญ่

การประมาณค่าความสามารถของผู้สอบจากการศึกษาเปรียบเทียบวิธีการประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์และวิธีความเป็นไปได้สูงสุดพบว่าวิธีการประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบส์มีความเที่ยงตรงเชิงสภาพสูงกว่าการประมาณค่าความสามารถด้วยวิธีความเป็นไปได้สูงสุด

สำหรับเกณฑ์ยุติการทดสอบนั้น พบว่า เกณฑ์ยุติการสอบต่างกันทำให้ความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการสอบต่างกัน ถ้าใช้เกณฑ์ค่าความคลาดเคลื่อนมาตรฐานน้อยกว่าหรือเท่ากับ .30 จะใช้จำนวนข้อสอบมากกว่า ใช้เกณฑ์ค่าความคลาดเคลื่อนมาตรฐานเท่ากับหรือน้อยกว่า .40 อยู่ถึง 2 เท่า ดังนั้น ถ้าเป็นการทดสอบโดยทั่ว ๆ ไปจึงควรใช้เกณฑ์ค่าความคลาดเคลื่อนมาตรฐาน .04 จะได้ค่าความเชื่อมั่นในการวัดเท่ากับ .9165 แต่ใช้ข้อสอบจำนวนน้อยลงกว่าครึ่ง

ส่วนตัวแปรความสามารถของผู้สอบ พบว่า ผู้สอบที่มีความสามารถต่างกัน จะใช้ข้อสอบมากน้อยต่างกัน และมีความเที่ยงตรงเชิงสภาพต่างกันด้วย ผู้สอบที่มีความสามารถสูงกว่าจะใช้ข้อสอบน้อยกว่า และความเที่ยงตรงเชิงสภาพสูงกว่าด้วยเช่นกัน

สายชล อบทม (2539) ทำการพัฒนาโปรแกรมสำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ และทำการประเมินผลการใช้โปรแกรม ผลปรากฏว่าได้โปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์รูปแบบทางแยกแปรผันประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบสส์ ใช้ภาษาปาสคาลในการเขียนโปรแกรม โปรแกรมปฏิบัติงานบนวินโดวส์ 95 และยุติแบบทดสอบเมื่อความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบต่ำกว่า 0.3 ส่วนการใช้โปรแกรมการทดสอบนั้น พบว่า นักเรียนมีความกระตือรือร้นที่จะทำการทดสอบ มีเจตคติที่ดีต่อการทดสอบ และอยากให้มีการทดสอบแบบนี้ในวิชาอื่น ๆ อีกด้วย

เบย์รอฟและซีเลย์ (Weiss and Others. 1974 ; citing Bayroff and Seeley. 1967) ดำเนินการสอบโดยใช้แบบทดสอบเฉพาะบุคคล รูปแบบปิรามิด 8 ชั้น 2 ฉบับ โดยให้ผู้สอบจำนวน 102 คน ทดสอบกับคอมพิวเตอร์ ผู้สอบจะต้องทำข้อสอบที่ทดสอบแบบมาตรฐานเดิม ซึ่งวัดด้านภาษา 50 ข้อ ด้านตัวเลข 40 ข้อ พบว่า สหสัมพันธ์ระหว่างการทดสอบเฉพาะบุคคลและการทดสอบแบบมาตรฐานเดิมมีค่าเท่ากับ 0.87 และ 0.79 เมื่อเปรียบเทียบกับค่าสหสัมพันธ์โดยประมาณระหว่างการทดสอบแบบมาตรฐานเดิมโดยใช้ข้อสอบ 8 ข้อ กับการทดสอบแบบมาตรฐานเดิมโดยใช้ข้อสอบ 40 และ 50 ข้อ พบว่า ค่าสหสัมพันธ์ที่ได้มีค่าเท่ากับ .75 และ .67 ดังนั้น จึงพบว่า การทดสอบโดยใช้รูปแบบปิรามิดมีค่าสหสัมพันธ์สูงกว่าการทดสอบแบบมาตรฐานเดิมเมื่อใช้ข้อสอบจำนวนเท่ากัน

ลินน์และคนอื่นๆ (Weiss and Others. 1974 ; citing Linn and others. 1969) ได้ศึกษาความแตกต่างของวิธีการแยกทาง 7 วิธี โดยใช้คอมพิวเตอร์ศึกษากับนักเรียน 4,885 คน แบบทดสอบมาตรฐานเดิมมี 190 ข้อ ในแต่ละวิธีการคัดเลือกข้อสอบที่นำมาใช้จะเลือกเฉพาะข้อสอบที่เหมาะสมกับผู้สอบแต่ละคน วิธีการแยกทาง 5 วิธีแรกเป็นแบบสองชั้น (Two Stages) อีก 2 วิธีเป็นแบบปิรามิด ซึ่งปิรามิดแบบแรกเป็นแบบ 10 ชั้นที่มีความแตกต่างระหว่างค่าความยากของข้อสอบในชั้นที่อยู่ติดกันประมาณ .02 และปิรามิดรูปแบบที่ 2 เป็นแบบ 5 ชั้น โดยที่แต่ละชั้นมีข้อสอบ 5 ข้อ นั่นคือ ผู้สอบแต่ละคนจะต้องทำข้อสอบในทางแยกใดทางแยกหนึ่งจำนวน 25 ข้อ แบบทดสอบรูปแบบปิรามิดทั้งสอง จะใช้วิธีแยกทางแบบเพิ่ม-ลดเท่ากัน แบบทดสอบรูปแบบปิรามิดจะถูกนำมาเปรียบเทียบกับแบบทดสอบมาตรฐานเดิมขนาดสั้น 5 ชุด ซึ่งมีจำนวนข้อสอบ 10 ถึง 50 ข้อ ผลการศึกษาพบว่า แบบทดสอบรูปปิรามิดชนิด 10 ชั้น มีความสัมพันธ์กับคะแนนรวมของแบบทดสอบ ซึ่งสอดคล้องกับความสัมพันธ์กับแบบทดสอบมาตรฐานเดิมที่มีข้อสอบ 35 ข้อ นอกจากนั้นลินน์และคณะได้ทำการศึกษากับแบบทดสอบวัดผลสัมฤทธิ์ 2 ชุด โดยใช้กลุ่มตัวอย่างเดียวกันเพื่อใช้เป็นเกณฑ์การวัด พบว่า แบบทดสอบรูปแบบปิรามิดชนิด 10 ชั้น มีความสัมพันธ์กับเกณฑ์การวัดสูงกว่าความสัมพันธ์กับแบบทดสอบมาตรฐานเดิมที่มีความยาวเท่ากัน ในทำนองเดียวกันนี้ แบบทดสอบรูปแบบปิรามิดชนิด 5 ชั้น ที่มีข้อสอบ 25 ข้อ จะมีความสัมพันธ์กับ

แบบทดสอบอิงเกณฑ์ทั้ง 2 ชุดสูงกว่าแบบทดสอบมาตรฐานเดิมที่มีข้อสอบ 50 ข้อ ผลการศึกษาแสดงให้เห็นว่า แบบทดสอบรูปแบบปิรามิดที่มีข้อสอบจำนวนน้อยข้อ จะมีความตรงเมื่อเปรียบเทียบกับแบบทดสอบมาตรฐานเดิม

ไบร์สัน (ตริงใจ พูลผลอำนาย. 2534 ; อ้างอิงจาก Bryson. 1971) ได้ศึกษาเปรียบเทียบแบบทดสอบรูปแบบปิรามิดชนิด 5 ชั้น 2 ชุด กับแบบทดสอบมาตรฐานเดิม 2 ชุด ๑ ละ 5 ข้อ สร้างจากข้อสอบที่สัมพันธ์กันจำนวน 100 ข้อ แบบมาตรฐานเดิมดำเนินการสอบโดยวิธีเขียนตอบ ขณะที่แบบทดสอบรูปแบบปิรามิดดำเนินการสอบโดยใช้เครื่องคอมพิวเตอร์ แบบทดสอบรูปแบบปิรามิดชุดหนึ่งใช้วิธีเลือกลำดับข้อสอบที่ขึ้นอยู่กับความแตกต่างของข้อสอบที่มากที่สุด สำหรับผู้สอบที่ตอบได้ถึงจุดที่กำหนดไว้ในโครงสร้างของรูปแบบปิรามิด ขณะที่แบบทดสอบรูปแบบปิรามิดอีกชุดหนึ่งใช้กระบวนการเลือกข้อสอบโดยการนำค่าสูงสุดของคะแนนรวม แบบทดสอบรูปแบบปิรามิดทั้ง 2 ฉบับใช้ค่าความแตกต่างระหว่างค่าความยากของข้อสอบในชั้นที่อยู่ติดกันแบบแปรผัน นำแบบทดสอบรูปแบบปิรามิดแต่ละชุดไปทดสอบกับกลุ่มตัวอย่าง จำนวน 236 คน โดยจัดเป็นกลุ่มแปรผัน 2 กลุ่ม และนำแบบทดสอบมาตรฐานเดิมทั้ง 2 ชุด ไปทดสอบกับกลุ่มเปรียบเทียบ จำนวน 250 คน ผลปรากฏว่าแบบทดสอบมาตรฐานเดิมขนาดสั้นชุดหนึ่ง มีความสัมพันธ์กันสูงกับคะแนนรวมในการทดสอบ และสูงกว่าความสัมพันธ์กับแบบทดสอบรูปแบบปิรามิด แบบทดสอบรูปแบบปิรามิดชุดหนึ่งมีความสัมพันธ์ต่ำกว่ากับคะแนนรวมของแบบทดสอบและต่ำกว่าความสัมพันธ์กับแบบทดสอบมาตรฐานเดิม

ลอร์ด (Lord. 1971) ได้ใช้คอมพิวเตอร์จำลองสถานการณ์ เพื่อศึกษาประเมินคุณภาพการทดสอบเฉพาะบุคคลแบบเฟลิกซีเลเวล โดยใช้ข้อสอบเฟลิกซีเลเวล จำนวน 119 ข้อ ซึ่งผู้สอบแต่ละคนจะต้องตอบข้อสอบแบบเฟลิกซีเลเวล จำนวน 60 ข้อ ผลการศึกษาพบว่า เมื่อขนาดของช่วงห่างความยากของข้อสอบมากขึ้นจะทำให้ค่าอินฟอร์เมชันลดลงทุกระดับความสามารถ และจะมีประสิทธิภาพสูงสุดเมื่อมีค่าอยู่ระหว่าง 0.033 ถึง 0.067 แบบทดสอบจะไม่มีประสิทธิภาพเมื่อใช้ช่วงห่างความยากใกล้เคียงกับ 0.1 เมื่อเปรียบเทียบกับ การทดสอบแบบมาตรฐานเดิม การทดสอบแบบเฟลิกซีเลเวลจะให้ค่าอินฟอร์เมชันสูงกว่าการทดสอบแบบมาตรฐานเดิม เมื่อผู้สอบมีความสามารถสูงหรือต่ำ ส่วนผู้สอบที่มีความสามารถกลาง ๆ การทดสอบแบบมาตรฐานเดิมจะให้ค่าอินฟอร์เมชันสูงกว่าการทดสอบแบบเฟลิกซีเลเวล

ลาร์กินและไวส์ (Lakin and Weiss. 1974) ได้ทดสอบวัดความสามารถโดยใช้แบบทดสอบเฉพาะบุคคลรูปแบบปิรามิด และใช้เครื่องคอมพิวเตอร์ดำเนินการสอบ โดยใช้โครงสร้างข้อสอบแบบปิรามิด 3 ชุด และแบบทดสอบมาตรฐานเดิมที่มีความยากปานกลาง ดำเนินการสอบกับนักเรียนในระดับวิทยาลัยที่เรียนจิตวิทยา 2 กลุ่ม พบว่า แบบทดสอบปิรามิด 15 ชั้น ให้ค่าสหสัมพันธ์แบบทดสอบซ้ำค่อนข้างต่ำกว่าแบบทดสอบมาตรฐานเดิมจำนวน 40 ข้อ อย่างไรก็ตาม เมื่อพิจารณาด้านผลกระทบด้านการจำแล้ว แบบทดสอบรูปแบบปิรามิดให้การประมาณค่าความสามารถที่คงที่มากกว่าแบบทดสอบมาตรฐานเดิมที่มีความยาวเท่ากัน

แมคคินลีย์ และคนอื่นๆ (McKinly and Others. 1983 ; <http://ercir.syr.edu/plweb-cgi/fastweb>) ได้ศึกษาการประเมินผลกระบวนการทดสอบ Logistic Tailored แบบหนึ่งและสาม พารามิเตอร์เพื่อการใช้ในศูนย์ข้อสอบขนาดย่อย โดยมีวัตถุประสงค์เพื่อเปรียบเทียบการประเมินความสามารถซึ่งได้มาจากกระบวนการทดสอบ tailored โดยมีพื้นฐานของ การคำนวณแบบหนึ่ง พารามิเตอร์ (1PL) และ การคำนวณแบบสามพารามิเตอร์ (3PL) ขั้นตอนแรกของการศึกษาใช้ข้อมูลจริง ในขณะที่ขั้นตอนที่สองใช้ข้อมูลทำขึ้น ในขั้นตอนแรกนั้นข้อมูลจากผู้เข้ารับการทดสอบ 3,000 คน จากการทดสอบย่อยในวิชาคณิตศาสตร์ ACT 40 ข้อ 2,000 คน แรกนั้นถูกใช้เพื่อหาพารามิเตอร์ ประมาณการสำหรับทั้งสองรูปแบบ โดยการใช้ตัวประมาณการเหล่านี้ การทดสอบ tailored แบบ 1PL และ 3PL ถูกกระทำขึ้นโดยใช้ข้อมูลตอบสนองจาก 1,000 คนที่เหลือ กระบวนการทดสอบ tailor ทั้งสองนั้นมีการใช้การประเมินความสามารถแบบเหมือนกันมากที่สุด และ กระบวนการการเลือกข้อสอบที่ใช้ข้อมูลมากที่สุด ชุดประเมินความสามารถสองชุดจะได้รับการเปรียบเทียบ ในขั้นตอนที่สอง ข้อมูลตอบสนอง 3000 กรณีถูกสร้างขึ้นโดยใช้ ตัวพารามิเตอร์ 3PL จากขั้นตอนแรกเป็นพารามิเตอร์จริง ความสามารถที่แท้จริงจะถูกเลือกมาจากค่ากระจายปกติมาตรฐาน 2000 กรณีแรกถูกใช้สำหรับการปรับข้อสอบแบบ 1PL และ 3PL และอีก 1000 ข้อที่เหลือถูกใช้ในการสร้างแบบทดสอบ tailored แบบ 1PL และ 3PL ชุดตัวประมาณด้านความสามารถสองชุดถูกเปรียบเทียบซึ่งกันและกัน และ เปรียบเทียบกับพารามิเตอร์ความสามารถที่แท้จริง ผลจากทั้งสองขั้นตอนในการศึกษาแสดงให้เห็นว่า การทดสอบ tailored แบบ 1PL และ 3PL ให้ผลประเมินความสามารถที่สัมพันธ์ร่วมกันอย่างสูง และ ไม่มีข้อได้เปรียบในด้านการประเมินความสามารถของการใช้รูปแบบหนึ่งต่ออีกรูปแบบหนึ่ง เนื่องจากว่าวิธีการประมาณค่าด้วย 1PL นั้นมีค่าใช้จ่ายถูกกว่า จึงแนะนำให้ใช้รูปแบบนี้สำหรับการประมาณค่าพารามิเตอร์ของข้อสอบกับการทดสอบแบบ tailored

เอสเต้ (Estes. 1985 : <http://ercir.syr.edu/plweb-cgi/fastweb>) ได้ศึกษาตัวอย่างของ ธนาคารข้อสอบที่สนับสนุนต่อการพัฒนาแบบทดสอบแบบท้องถิ่น : สองกรณีศึกษาพร้อมปฏิบัติการ โดยข้อมูลต่าง ๆ ถูกรวบรวมโดย Assessment Development and Use Project ซึ่งจัดตั้งโดย Northwest Regional Educational Laboratory (NWREL) เพื่อที่จะสนับสนุนความต้องการในการ พัฒนาแบบทดสอบจากหน่วยงานรัฐและตัวแทนภาคท้องถิ่น ลักษณะธนาคารข้อสอบเฉพาะจะ รายงานเฉพาะกรณีศึกษาสองกรณี กรณีแรกคือ Wisconsin Item Bank ซึ่งพัฒนาโดย Wisconsin State Department of Public Instruction ซึ่งเป็นโรงเรียนที่ใช้การทดสอบความพร้อมเป็นฐาน อยู่ใน วิสคอนซิน โดยจะมีการอธิบายรายละเอียดเกี่ยวกับการออกแบบ ความเป็นมา ปัญหา และการ ปฏิบัติการของธนาคารข้อสอบ และกรณีที่สองอธิบายถึงธนาคารข้อสอบที่พัฒนาโดยเขตเมืองอื่นๆ สิ่ง ที่กล่าวถึงก็คือภูมิหลัง และ วัตถุประสงค์ในการพัฒนา, การใช้ในปัจจุบัน, ปัญหาและเนื้อหาในการ พัฒนา, การจัดการ และการใช้ และหนทางที่จะใช้เทคโนโลยีร่วมกัน ในส่วนสุดท้ายพิจารณาถึง เนื้อหาที่เกี่ยวข้องว่าธนาคารข้อสอบจะทำงานได้ดีอย่างไร และสามารถตอบสนองศักยภาพของตนเอง ได้ดีแค่ไหน

วิสน์ไวส์กี (ตริงใจ พูลผลอำนาจ. 2534 ; อ้างอิงจาก Wisniewski. 1986) ได้ตรวจสอบความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบ จากการประยุกต์ใช้ราชต์โมเดลกับการทดสอบ เอ ที (Adaptive Testing) แบบบี เอส เอ็ม (Binary Search Method : BSM) การวิเคราะห์ข้อมูล ใช้การวิเคราะห์ถดถอยพหุคูณ (Multiple Regression) โดยใช้คะแนนจากการทดสอบบี เอส เอ็ม เป็นตัวทำนายคะแนนการทดสอบแบบมาตรฐานเดิมที่ใช้กระดาษ-ดินสอ (paper Pencil Testing) ผลการศึกษาพบว่า คะแนนจากการทดสอบบี เอส เอ็ม สามารถทำนายคะแนนจากการทดสอบแบบมาตรฐานเดิมที่ใช้กระดาษดินสอได้ดี นอกจากนั้นการทดสอบแบบบี เอส เอ็ม ยังสามารถที่จะควบคุมเวลาในการตอบเป็นรายข้อ และสามารถควบคุมสภาพแวดล้อมที่เป็นตัวแปรแทรกซ้อนได้ดีกว่าการทดสอบมาตรฐานเดิม

แบล็คมอร์ (ตริงใจ พูลผลอำนาจ. 2534 ; อ้างอิงจาก Blackmore. 1987) ได้ศึกษาเปรียบเทียบการดำเนินการสอบ 3 วิธี คือ ใช้คอมพิวเตอร์ดำเนินการสอบ (Computerized) ใช้คอมพิวเตอร์ปรับความยากให้เหมาะสมกับระดับความสามารถของผู้สอบ (Computerized Adaptive Testing) และการดำเนินการสอบแบบมาตรฐานเดิม คือ ใช้กระดาษกับดินสอ (Paper and Pencil) ได้ทำการศึกษากับนักเรียนเกรดสิบสองที่กำลังจะจบการศึกษาในปี ค.ศ. 1986 ของโรงเรียนประจำตำบลการ์เดนโกรฟ (Gardeen Grove High School) ในรัฐคาลิฟอร์เนีย จำนวน 24 คน โดยแบ่งออกเป็น 3 กลุ่มๆ ละ 8 คน เพื่อดำเนินการทดสอบ คือ ความถนัดทางการเรียน จำนวน 6 ฉบับย่อย คือ การให้เหตุผลทางภาษา (Verbal Reasoning) ความสามารถทางตัวเลข (Numerical Ability) การให้เหตุผลเชิงนามธรรม (Abstract Reasoning) การให้เหตุผลเชิงกล (Mechanical Reasoning) มิติสัมพันธ์ (Space relations) และการใช้ภาษา (Language Usage) ผลจากการศึกษาพบว่า คะแนนจากข้อสอบย่อยทุกฉบับจากการดำเนินการสอบทั้ง 3 วิธี ไม่แตกต่างกัน ส่วนทัศนคติและความคิดเห็น ต่าง ๆ เกี่ยวกับการดำเนินการสอบทั้ง 3 วิธี จากการใช้แบบทดสอบตามและสัมภาษณ์ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ลี (ตริงใจ พูลผลอำนาจ. 2534 ; อ้างอิงจาก Lee. 1987) ได้พัฒนาการทดสอบแบบเอ็มเอเอที (Microcomputerized Adaptive Achievement Test : MAAT) โดยใช้ยุทธวิธี 2 แบบ คือ แบบเอ็มพีอาร์ที ของวอลท์ (Wald's Sequential Probability Ratio Test : PETA) และยุทธวิธีของลอว์ด แบบเฟล็กซ์เลเวล (Logd's Flexilevel Test) โดยใช้ข้อสอบที่มีอยู่ในคลังข้อสอบพีอีทีเอ (The Pitt Educational Testing Aids : PETA) ได้ทดลองศึกษากับนักศึกษาระดับบัณฑิตศึกษา จำนวน 3 ห้องเรียน ในรายวิชาการวิจัยการศึกษาเบื้องต้น ผลการศึกษาพบว่า

1. มีความเป็นไปได้สูงในอนาคตที่จะนำเอาการทดสอบแบบเอ็มเอเอที (MAAT) มาใช้ในการทดสอบผลสัมฤทธิ์ทางการเรียนในระดับบัณฑิตศึกษา ผู้ดำเนินการสอบไม่จำเป็นต้องมีความรู้ทางด้านจิตวิทยาหรือคอมพิวเตอร์มากนักก็สามารถดำเนินการสอบได้ เพียงแต่ดำเนินการตามขั้นตอนต่าง ๆ ตามคู่มือครูที่เขียนไว้อย่างละเอียด

2. คะแนนจากการทดสอบแบบเอ็มเอเอที (MAAT) มีความสัมพันธ์ทางบวกกับคะแนนผลการสอบกลางเทอม

3. นักศึกษามีทัศนคติที่ดีต่อการทดสอบแบบเอ็มเอเอที (MAAT) ในทางบวก

โฮ (ตริงใจ พูลผลอำนาจ. 2534 ; อ้างอิงจาก Ho. 1989) ได้ใช้ไมโครคอมพิวเตอร์ศึกษาเปรียบเทียบยุทธวิธีการใช้ แอดาปทีฟ เทสติ้ง (Adaptive Testing) 3 วิธี คือยุทธวิธีของเบส์ (Bayesian) โมเดลเบย์ (Model Bayesian) และความเป็นไปได้สูงสุด (Maximum Likelihood) ยุทธวิธีทั้ง 3 นี้ ได้เปรียบเทียบโดยใช้ขนาดของคลังข้อสอบ (Bank Sizes) 3 ขนาด คือ ขนาด 86 ข้อ 71 ข้อ และ 56 ข้อ ชนิดของคลังข้อสอบ (Bank Type) 3 แบบ คือ แบบที่มีข้อสอบแบบสุ่ม แบบที่มีเฉพาะข้อสอบที่มีค่าอำนาจจำแนกสูง และแบบที่มีเฉพาะข้อสอบที่ง่าย ๆ ส่วนระดับความสามารถของผู้สอบ มี 3 ระดับ คือ สูง ปานกลาง ต่ำ การวิเคราะห์ข้อมูลโดยใช้การวิเคราะห์ความแปรปรวน 3 ทาง (Three ways Anova with repeated reasures) ผลการศึกษาพบว่า

1. ยุทธวิธีของโมเดลเบย์ จะมีประสิทธิภาพ (Efficient) สูงสุด ยุทธวิธีของเบย์ จะให้ค่าความเที่ยงสูงสุด และยุทธวิธีความเป็นไปได้สูงสุดจะให้ค่าไม่คงที่ (Inconsistent) ภายใต้การทดสอบในสถานการณ์ต่าง ๆ
2. การเลือกขนาดคลังข้อสอบดูเหมือนว่าจะส่งผลน้อยมาก ยกเว้นในกลุ่มผู้สอบที่มีความสามารถต่ำเมื่อใช้ข้อสอบเริ่มต้นที่มีความยากสูงกว่าความสามารถที่แท้จริงของผู้สอบ
3. ถ้าใช้ข้อสอบเริ่มต้นที่มีความยากเท่ากับหรือน้อยกว่าความสามารถที่แท้จริงของผู้สอบแล้ว ความถูกต้องในการประมาณค่าจะมีมากขึ้น
4. มีปฏิสัมพันธ์ (Interaction) ระหว่างชนิดของคลังข้อสอบกับยุทธวิธีของ Adaptive Testing และระหว่างชนิดของคลังข้อสอบกับระดับความสามารถของผู้สอบ
5. ความแตกต่างระหว่างคุณลักษณะทางสถิติของการเลือกข้อสอบกับไม่เลือกข้อสอบมีความสัมพันธ์กันต่ำ

ดอดด์ และคนอื่น (Dodd and Others. 1993 : 61-77) ได้ทำการศึกษาตัวแปรที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์รูปแบบ Partial Credit Model จาก 3 ตัวแปร ได้แก่ เกณฑ์การยุติการทดสอบ จำนวนข้อสอบของการทดสอบหรือแบบทดสอบ และการกระจายค่าความยากของข้อสอบ ผลการศึกษาพบว่า เกณฑ์การยุติการทดสอบตามวิธีค่าความคลาดเคลื่อนมาตรฐานแบบมีเงื่อนไขที่หลากหลายเป็นเกณฑ์ยุติการทดสอบที่ดีกว่าใช้เกณฑ์การยุติการทดสอบด้วยวิธีสารสนเทศน้อยที่สุด ด้านจำนวนข้อสอบการทดสอบแบบปรับเหมาะนั้นควรมีจำนวนข้อเท่ากับ 30 ข้อ จะทำให้แบบทดสอบมีค่าความยากง่ายปานกลางเหมาะสมกับความสามารถของผู้สอบ และความยากง่ายของข้อสอบควรให้กระจายครอบคลุมกับความสามารถของผู้สอบในทุกระดับความสามารถ

เบอร์เกอร์ (Berger. 1994 : 141-153) ได้ทำการศึกษาแนวทางทั่ว ๆ ไปเพื่อทำการออกแบบขั้นตอนของการทดสอบด้วยแบบทดสอบฟอร์มคงที่ (Fixed-Form) และการทดสอบแบบปรับเหมาะด้วยวิธี Testlets เนื่องจากการคัดเลือกข้อสอบจากธนาคารข้อสอบของแบบทดสอบฟอร์มคงที่

และการทดสอบแบบปรับเหมาะด้วยวิธี Testlets ก่อนข้างมีปัญหาในการคัดเลือกข้อสอบให้เหมาะสมกับความสามารถของผู้สอบได้ ผลการศึกษาได้แนวทางในการออกแบบขั้นตอนในการคัดเลือกข้อสอบสำหรับการทดสอบฟอร์มคงที่และการทดสอบปรับเหมาะแบบ testlets ด้วยการนำโปรแกรมทางคณิตศาสตร์เข้ามาช่วยในการคัดเลือกข้อสอบให้กับแบบทดสอบทั้งสอง ด้วยวิธีที่คล้ายกับการทดสอบแบบปรับเหมาะ (Optimal) วิธีการนี้ยังสามารถนำไปประยุกต์ใช้ได้ง่ายในสถานการณ์ต่าง ๆ และการประมาณค่าความสามารถของผู้สอบจากการคัดเลือกข้อสอบด้วยวิธีดังกล่าวจะให้ความเที่ยงตรงสูงอีกด้วย

เฮทเทอร์ และ ซีเกล (Hetter and Segall. 1994 : 197-204) ทำการศึกษาเปรียบเทียบ ✓
วิธีการคัดเลือกข้อสอบสำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ โดยทำการคัดเลือกแบ่งกลุ่มข้อสอบออกเป็น 3 กลุ่ม กล่าวคือกลุ่มที่ 1 ใช้คอมพิวเตอร์ในการจัดกลุ่มของ ข้อสอบ กลุ่มที่ 2 ใช้ค่าพารามิเตอร์ของข้อสอบจากการทดสอบด้วยการเขียนตอบลงกระดาษคำตอบ และกลุ่มที่ 3 ใช้ค่าพารามิเตอร์จากการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะตามทฤษฎีการตอบสนองข้อสอบ การศึกษาเปรียบเทียบครั้งนี้เพื่อวิเคราะห์รูปแบบโครงสร้างของข้อสอบ (Structural Model) ผลปรากฏว่ารูปแบบโครงสร้างของข้อสอบจากการจัดกลุ่มของข้อสอบโดยใช้ค่าพารามิเตอร์ของข้อสอบจากการทดสอบด้วยการเขียนตอบลงบนกระดาษคำตอบและการใช้ค่าพารามิเตอร์จากการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบปรับเหมาะตามทฤษฎีการตอบสนองข้อสอบไม่แตกต่างกัน และความเชื่อมั่นจากการทดสอบของทั้ง 2 กลุ่มก็ไม่แตกต่างกัน แต่การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบตามทฤษฎีการตอบสนองข้อสอบจะใช้จำนวนข้อสอบน้อยกว่า

ไบซอท และคนอื่น ๆ (Bizot and Others. 1994 : [http://ercir.syr.edu/plweb-
cgi/fastweb](http://ercir.syr.edu/plweb/cgi/fastweb)) ได้ทำการศึกษาผลกระทบจากการใช้รูปแบบการทดสอบและค่าพารามิเตอร์จากการ ✓
ทดสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) โดยมีจุดมุ่งหมายเพื่อประเมินโมเดลตามทฤษฎีการตอบสนองข้อสอบระหว่างโมเดลแบบสองพารามิเตอร์กับโมเดลแบบสามพารามิเตอร์ เพื่อทำการประมาณค่าความสามารถเริ่มต้นการทดสอบและเกณฑ์การยุติการสอบ โดยใช้ข้อสอบคำศัพท์จากบทสนทนาจำนวน 80 ข้อ กับกลุ่มตัวอย่าง 3 กลุ่ม คือ กลุ่มนักเรียนชั้นมัธยมศึกษาตอนปลาย กลุ่มนักเรียนชั้นมัธยมศึกษาตอนต้น และกลุ่มของนักเรียนชั้นมัธยมศึกษาตอนต้นและตอนปลายจากกลุ่มแรกและกลุ่มที่สองเพื่อทำการสอบซ้ำ วิธีการประมาณค่าความสามารถเริ่มแรกนั้นใช้ 2 วิธีคือ การประมาณค่าความสามารถแบบสุ่มและการประมาณค่าความสามารถที่ได้จากการคำนวณด้วยโมเดลแบบสองพารามิเตอร์กับโมเดลแบบสามพารามิเตอร์ ทำการเก็บข้อมูลกับกลุ่มตัวอย่างทั้งหมด 2,697 คน ผลการศึกษาปรากฏว่า รูปแบบการทดสอบและค่าพารามิเตอร์จากการทดสอบตามทฤษฎีการตอบสนองข้อสอบแตกต่างกัน โดยโมเดลตามทฤษฎีการตอบสนองข้อสอบแบบสามพารามิเตอร์มีความเหมาะสมที่สุดทั้งในแง่เชิงทฤษฎีและเชิงปฏิบัติการ

ซวิก และวินเกอร์สกาย (Zwick and Wingersky. 1994 : 121-140) ได้ศึกษาวิธีการต่าง ๆ สำหรับการประเมินข้อสอบที่ทำหน้าที่ต่างกันในการทดสอบแบบปรับเหมาะกับความสามารถ

ของผู้สอบด้วยคอมพิวเตอร์ ใช้วิธี Mantel-Haenszel ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ ผู้เข้าสอบแต่ละคนจะได้รับข้อสอบ 25 ข้อ จากแบบทดสอบที่มีข้อสอบ 75 ข้อ โดยใช้รูปแบบของ ทฤษฎีการตอบสนองข้อสอบแบบ three-parameter logistic และคะแนนของผู้เข้าสอบจะจับคู่ระหว่าง คะแนนจริงที่คาดหวังตามการตอบสนองข้อสอบ กับพารามิเตอร์ข้อสอบที่ประมาณค่าได้ ผลการวิจัย พบว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ สามารถประเมินการทำหน้าที่ต่างกันของข้อสอบได้อย่างมีประสิทธิภาพ

ดอดด์ และคูก (Dodd and Koch. 1995 : 5 - 22) ทำการศึกษาการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์กับข้อสอบแบบให้คะแนนหลายค่า (Polytomous) และทำการศึกษาค้นคว้าขั้นตอนและวิธีใช้การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ รูปแบบ Polytomous ในหลาย ๆ รูปแบบตามโมเดลทฤษฎีการตอบสนองข้อสอบ เพื่อทำการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แบบให้คะแนนหลายค่าไป ใช้กับการทดสอบในวิชาการตลาด จิตวิทยา และวิชาการศึกษา ผลการศึกษายังไม่ได้ข้อสรุปว่า รูปแบบการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แบบให้คะแนนหลายค่ารูปแบบใดเหมาะสมอย่างไร ซึ่งดอดด์ ได้เสนอแนะไว้ว่าควรให้มีการวิจัยการใช้รูปแบบ Polytomous CAT ต่อไป

ชางและยีน (Chang and Ying. 1996 : 213 - 229) ทำการศึกษาวิธีการคัดเลือก ข้อสอบของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ โดยใช้ สารสนเทศแบบ Global Information กับการคัดเลือกข้อสอบโดยวิธีสารสนเทศสูงสุดของ Fisher ที่ระดับความสามารถซึ่งถูกประมาณค่าก่อนหน้าของผู้สอบ เนื่องจากการคัดเลือกข้อสอบด้วยวิธี สารสนเทศมากที่สุดของ Fisher อาจมีประสิทธิภาพน้อยลงกว่าที่ควรจะเป็น เมื่อค่าความสามารถที่ ประมาณได้ ($\hat{\theta}$) ไม่ใกล้เคียงกับค่าความสามารถจริง (θ) โดยเฉพาะอย่างยิ่งเมื่อการประมาณค่า ความสามารถเริ่มแรกด้วยข้อสอบเพียงไม่กี่ข้อ ค่าความสามารถที่ได้จากการประมาณอาจแตกต่างกับ ค่าความสามารถจริงอยู่มาก ผลการศึกษาเปรียบเทียบวิธีการคัดเลือกข้อสอบ ปรากฏว่าวิธีการ คัดเลือกข้อสอบของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ โดย วิธีใช้สารสนเทศแบบ Global Information สามารถลดความลำเอียงของการทดสอบลง และสามารถลด ความคลาดเคลื่อนมาตรฐานของการวัดลง

หลุย (Luecht. 1996 : 389 - 404) ได้ศึกษาการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ในมิติที่หลากหลายในการทดสอบเพื่อขอประกาศนียบัตร หรือใบรับรอง ซึ่งใช้ทฤษฎีการตอบสนองข้อสอบในมิติที่หลากหลาย (MIRT) ซึ่งซีเกล (Segall) เป็นผู้สร้างขึ้นเพื่อนำมาประยุกต์ใช้ในการออกใบรับรองหรือประกาศนียบัตร มีเนื้อหาผสมผสานและ ซับซ้อน แต่แบบทดสอบจะต้องประกอบด้วยข้อสอบที่มีความสมดุลของเนื้อหา ข้อสอบจะแบ่งเป็น คะแนนย่อย โดยใช้แนวคิดของการปรับเหมาะตามวิธี MIRT เพื่อที่จะปรับปรุงความเชื่อมั่นของคะแนน ย่อย กรอบแนวคิดโดยสรุปเป็นการลงสรุปทางสถิติทั้งแบบ univariate และ multivariate ซึ่งต้องมี

ระบบในการจัดการเนื้อหาและรูปแบบของโครงสร้างข้อสอบแบบปรับเหมาะให้มีความสมดุลย์ ผลการวิจัยปรากฏว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ที่ใช้ ทฤษฎีการตอบสนองข้อสอบในมิติที่หลากหลาย (MIRT) นั้นให้ค่าความแปรปรวนของความเชื่อมั่นสูง มีประโยชน์ต่อการวัดทางจิตวิทยา

วังและวิสเฟิล (Wang and Vispoel. 1998 : 109 - 135) ได้ศึกษาลักษณะของวิธีการ ประมาณค่าความสามารถในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ โดยใช้วิธีการประมาณค่าพารามิเตอร์ 4 วิธีคือ Maximum likelihood estimation (MLE), Three bayesian approaches, วิธีของ Owen, Expected a posteriori (EAP) และ Maximum a posteriori ตามทฤษฎีการประมาณค่าความสามารถของผู้สอบและการวิจัยเชิงประจักษ์ก่อนหน้านี้ ซึ่งให้เห็นความแตกต่างอย่างชัดเจนระหว่างวิธี MLE และวิธี Bayesian ซึ่ง MLE จะมีความลำเอียงต่ำกว่า ส่วน ความคลาดเคลื่อนมาตรฐานสูงกว่า มีความสอดคล้องต่ำกว่าและมีการจัดสอบที่มีประสิทธิภาพต่ำกว่า ความคลาดเคลื่อนมาตรฐานจริงในขณะที่ความคลาดเคลื่อนมาตรฐานสำหรับวิธี Bayesian ขึ้นอยู่กับ การกระจายของค่าส่วนเบี่ยงเบนมาตรฐาน (SD) จากการประมาณค่า ความคลาดเคลื่อนมาตรฐานจริง ที่เกิดภายหลัง ส่วนวิธี EAP ให้ผลดีที่สุด การใช้เกณฑ์การเริ่มต้นโดยจัดผู้เข้าสอบเป็น 3 กลุ่ม ตาม ระดับความสามารถจะช่วยลดความลำเอียงลง ในกรณีที่ใช้วิธีการของ Bayesian แต่มีผลต่อวิธีของ MLE เพียงเล็กน้อย จากความลำเอียงของผลการวิจัยเหล่านี้ จึงเป็นข้อเสนอแนะสำหรับการเลือก วิธีการประมาณค่าความสามารถของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วย คอมพิวเตอร์ให้เหมาะสมในบริบทของการตัดสินใจที่แตกต่างกัน

เรเวิลต้า และพอนโซดา (Reveulta and Ponsoda. 1998 : 311- 327) ได้เปรียบเทียบ วิธีควบคุมการนำเสนอข้อสอบในการสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ 2 วิธีคือ วิธี Progressive Method และวิธี Restricted Maximum Information Method วิธีทั้งสองได้ ถูกเปรียบเทียบกับวิธีการคัดเลือกข้อสอบวิธีอื่น ๆ อีก 6 วิธี (1. Maximum Information 2. One-Parameter 3. McBride and Martin 4. Randomesques 5. Sympson and Hetter และ 6. Random Item Selection) ที่เกี่ยวกับตัวแปรความถูกต้องของแบบทดสอบและการนำเสนอข้อสอบ ผลการวิจัย แสดงให้เห็นว่าวิธีการ Restricted Maximum Method มีประโยชน์ในการลดอัตราการนำเสนอข้อสอบ มากที่สุด และวิธี Progressive Method ลดจำนวนข้อสอบที่ไม่เคยใช้ลง ทั้ง 2 วิธี จะทำให้ข้อสอบมี ความถูกต้องมาก ดังนั้นวิธีการแบบรวมทั้ง Progressive – Restricted method อาจมีประโยชน์ใน การควบคุมการนำเสนอข้อสอบ โดยไม่ได้ลดความถูกต้องของข้อสอบลง

สต็อกกิงและสวานสัน (Stocking and Swanson.1998: 271- 279) ได้พัฒนาออกแบบ วิชาการข้อสอบที่ใช้กับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ วิธีการ optimal design ถูกประยุกต์ใช้ในการออกแบบวิชาการข้อสอบของการทดสอบแบบปรับเหมาะ ดังกล่าว เพื่อที่จะหาวิธีการที่จะลดการนำเสนอข้อสอบ เพิ่มความปลอดภัยของข้อสอบ และวิธีการที่จะ เพิ่มประสิทธิภาพของการคัดเลือกข้อสอบ ผลปรากฏว่า การออกแบบสร้างวิชาการข้อสอบด้วยวิธี

Optimal design สามารถเพิ่มความปลอดภัยของข้อสอบ การคัดเลือกและเพิ่มประสิทธิภาพในการคัดเลือกข้อสอบมากขึ้น

ไวส์ และคนอื่น ๆ (Wise and Others. 1999 : 185 - 198) ทำการศึกษาเกี่ยวกับการทบทวนคำตอบของผู้สอบซึ่งอาจทำให้ผู้สอบได้คะแนนเพิ่มขึ้นจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ซึ่งได้เสนอยุทธศาสตร์ที่ให้ผู้สอบสามารถเปลี่ยนความยากของข้อสอบให้เหมาะกับความสามารถของตนเองได้หรือให้ผู้เข้าสอบสามารถตรวจทานคำตอบของตนเองได้ ผลการวิจัยพบว่า ข้อสอบที่ยากเกินไปไม่สามารถจำแนกความสามารถของผู้สอบที่มีความสามารถพิเศษได้ ดังนั้นการใช้ยุทธศาสตร์ KINGSBURY ที่ทำให้ผู้สอบสามารถเปลี่ยนข้อสอบที่มีความยากให้เหมาะสมกับความสามารถของตนเอง จะทำให้ผู้สอบทำคะแนนจากการทดสอบได้มากขึ้น

พอนโซดา และคนอื่น ๆ (Ponsoda and Others. 1999 : 167 - 184) ได้ศึกษาเปรียบเทียบความยากง่ายของการทดสอบแบบปรับเหมาะด้วยตนเอง (SATs) และการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (CAT) ผลการวิจัยพบว่า ไม่มีความแตกต่างกันระหว่างแบบทดสอบทั้ง 4 ลักษณะ (SAT ที่ง่าย, SAT ที่ยาก, CAT ที่ง่าย และ CAT ที่ยาก) สำหรับค่าความสามารถที่ได้จากการประมาณและความวิตกกังวลในการทดสอบมีความแตกต่างกันอย่างมีนัยสำคัญ กล่าวคือ ค่าความสามารถที่ได้จากการประมาณด้วยการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จะมีความถูกต้องสูงกว่าค่าความสามารถที่ได้จากการประมาณจากแบบทดสอบแบบปรับเหมาะด้วยตนเอง และความวิตกกังวลจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จะน้อยกว่าด้วย

เชน และ ชาง (Chen and Chang. 2000 : 241 - 255) ศึกษาเปรียบเทียบหลักเกณฑ์การคัดเลือกข้อสอบในขั้นตอนเริ่มต้นของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ โดยใช้หลักเกณฑ์การคัดเลือกข้อสอบ 5 วิธี ได้แก่ Fisher information (FI), Fisher interval information (FII), Fisher information with a posterior distribution (FIP), Kullback-Leibler information (KL) และ Kullback-Leibler information with a posterior distribution (KLP) ทั้งในด้านประสิทธิภาพและความถูกต้องของการประมาณค่าความสามารถของผู้สอบใน (θ) ผลการวิจัยพบว่า ขั้นตอนแรกวิธี FII, FIP และ KLP ดีกว่าวิธี FI เพียงเล็กน้อยที่ค่าความสามารถ ($\theta = -3$ และ -2) สำหรับแบบทดสอบที่มีจำนวนข้อมากกว่า 10 ข้อ ให้ผลไม่แตกต่างกัน

เซงและเลียว (Cheng and Liou. 2000 : 257 - 265) ได้ทำการศึกษาการประมาณค่าระดับความสามารถในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ผลการศึกษาพบว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์นั้นจะประมาณค่าความสามารถของผู้สอบได้ถูกต้องสมเหตุสมผล จากจำนวนข้อสอบเพียงเล็กน้อย การนำการทดสอบแบบปรับเหมาะไปใช้อย่างมีประสิทธิภาพขึ้นอยู่กับ 1) ความถูกต้องของวิธีการทางสถิติที่ใช้สำหรับประมาณค่าความสามารถของผู้สอบ และ 2) ประสิทธิภาพของเกณฑ์ในการคัดเลือกข้อสอบ วิธีการประมาณค่าความสามารถของผู้สอบเหมาะสมสำหรับ CAT นอกจากนี้ยังตรวจสอบความถูกต้อง

ของขั้นตอน CAT ที่แตกต่างกัน ผลการวิจัยแสดงให้เห็นว่า การประมาณค่าความสามารถของผู้เข้าสอบถูกต้องจะต้องไม่มีความลำเอียงในการคัดเลือกข้อสอบเริ่มต้นในการทดสอบแบบปรับเหมาะ และ การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จะมีประสิทธิภาพมากขึ้นเมื่อสอบด้วยข้อสอบ 10 ข้อหรือมากกว่า

เอเดเจนและสตรัคแมน (Eggen and Straetmans, 2000 : 713 - 734) ได้ศึกษาการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เพื่อจำแนกผู้เข้าสอบออกเป็น 3 ประเภท วัตถุประสงค์เพื่อสำรวจความเป็นไปได้ในการใช้การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ในสถานการณ์ที่ผู้เข้าสอบถูกจำแนกออกเป็น 3 กลุ่ม ด้วยการทดสอบแบบ 2 ขั้นตอน กล่าวคือ ขั้นตอนที่ 1 เป็นการทดสอบทางสถิติ ขั้นตอนที่ 2 เป็นการประมาณค่าทางสถิติ สำหรับวิธีการคัดเลือกข้อสอบนั้น พิจารณาจากค่าสารสนเทศสูงสุดและวิธีการควบคุมการนำเสนอ ข้อสอบ ผลการวิจัยพบว่า จำนวนเฉลี่ยของข้อสอบในแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ (CAT) มีจำนวนลดลงอย่างน้อย 22% เมื่อเปรียบเทียบกับแบบทดสอบการจัดอันดับที่ใช้กระดาษเขียนตอบ นอกจากนี้ การทดสอบทางสถิติเป็นทางเลือกที่ดีจะประมาณค่าทางสถิติ และการกำหนดยุทธศาสตร์ในการคัดเลือกข้อสอบที่มีสารสนเทศสูงสุด

การวิจัยและพัฒนาในเรื่องการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ได้พัฒนาก้าวหน้ามากขึ้น และมีการนำไปใช้ในวงการศึกษาย่างกว้างขวางในปัจจุบัน ประโยชน์ที่ได้จะเกิดแก่ครู นักเรียนและผู้เชี่ยวชาญด้านการวัดผล สิ่งสำคัญที่สุดอีกประการหนึ่งคือ ลักษณะเด่นของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ สามารถนำไปใช้ในหลักสูตรที่มีการออกแบบโปรแกรมให้นักเรียนเรียนรู้ด้วยตนเอง ผลของโปรแกรมการเรียนรู้ด้วยตนเองอีกประการหนึ่งคือ นักเรียนจะต้องทดสอบบ่อยๆ มากกว่าการเรียนรู้ในโปรแกรมแบบเดิม ซึ่งจะต้องทำการทดสอบในเวลาที่แตกต่างกัน ดังนั้นการทดสอบแบบมาตรฐานเดิมจึงทำให้เกิดปัญหาในเรื่องความน่าเชื่อถือของผลการสอบและปัญหาการเปรียบเทียบผลการทดสอบจากการทดสอบในเวลาที่แตกต่างกัน ดังนั้นการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์สามารถแก้ปัญหาต่างๆ ดังกล่าวนี้ได้ จึงทำให้ผู้วิจัยสนใจศึกษาค้นคว้าเพื่อพัฒนาระบบการทดสอบและศึกษาประสิทธิภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ในครั้งนี้

บทที่ 3

วิธีดำเนินการวิจัย

การศึกษาวิจัยในครั้งนี้ผู้วิจัยได้ดำเนินการศึกษาวิจัยแบ่งออกเป็น 2 ระยะ

1. การวิจัยระยะที่ 1 เพื่อพัฒนาเครื่องมือ และหาประสิทธิภาพของเครื่องมือประกอบด้วย
 - พัฒนาข้อสอบและคลังข้อสอบ
 - พัฒนาโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะ
 - ศึกษาประสิทธิภาพของโปรแกรมคอมพิวเตอร์สำหรับการทดสอบ
2. การวิจัยระยะที่ 2 เพื่อศึกษาผลของการวิจัย
 - ศึกษาผลของตัวแปรอิสระที่คาดว่าจะมีอิทธิพลต่อตัวแปรตาม

โดยผู้วิจัยได้ดำเนินการวิจัยตามระยะดังกล่าวดังต่อไปนี้

ประชากรและกลุ่มตัวอย่าง

ประชากรและกลุ่มตัวอย่างสำหรับการพัฒนาข้อสอบและคลังข้อสอบ

ประชากรสำหรับการพัฒนาข้อสอบและคลังข้อสอบ ในครั้งนี้ ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 1 ปีการศึกษา 2546 โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาปทุมธานี เขตที่ 1 และ 2 จำนวนทั้งหมด 23 โรงเรียน เป็นโรงเรียนขนาดใหญ่ 6 โรงเรียน ขนาดกลาง 8 โรงเรียน และขนาดเล็ก 9 โรงเรียน จำนวนนักเรียนทั้งสิ้น 6,714 คน

กลุ่มตัวอย่างที่ใช้ในการพัฒนาข้อสอบและคลังข้อสอบ ได้มาโดยวิธีการสุ่ม ซึ่งมีขั้นตอนในการสุ่มดังนี้

ขั้นที่ 1 สุ่มโรงเรียนจำนวน 12 โรงเรียน จากโรงเรียนมัธยมศึกษา ในโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษาปทุมธานี เขตที่ 1 และ 2 โดยการสุ่มแบบแบ่งชั้น (Stratified random sampling) ซึ่งใช้ขนาดของโรงเรียนเป็นชั้นของการสุ่ม ได้โรงเรียนขนาดใหญ่ 3 โรงเรียน ขนาดกลาง 4 โรงเรียน และขนาดเล็ก 5 โรงเรียน จำนวนห้องเรียนทั้งสิ้น 88 ห้อง

ขั้นที่ 2 สุ่มห้องเรียน จากโรงเรียนที่ได้จากการสุ่มในขั้นตอนแรกโดยวิธีการสุ่มอย่างง่าย (Simple random sampling) โดยการสุ่มจากโรงเรียนขนาดใหญ่โรงละ 3 ห้อง ได้ห้องเรียน 9 ห้อง โรงเรียนขนาดกลางโรงละ 2 ห้อง ได้ห้องเรียน 8 ห้อง และโรงเรียนขนาดเล็กโรงละ 1 ห้อง ได้ห้องเรียน 5 ห้อง รวมห้องเรียน 22 ห้อง นักเรียนจำนวน 597 คน ดังแสดงในตาราง 4

ตาราง 4 จำนวนนักเรียนกลุ่มตัวอย่าง จำแนกเป็นรายโรงเรียน ขนาดของโรงเรียน และห้องเรียน

ที่	โรงเรียน	ขนาด โรงเรียน	จำนวน ห้อง	กลุ่ม ตัวอย่าง
1	ธรรมศาสตร์คลองหลวงวิทยาคม	ใหญ่	3	90
2	ชัยบุรี	ใหญ่	3	85
3	ชัยรักษ์	ใหญ่	3	77
4	มัธยมวัดหัตถสารเกษตร	กลาง	2	60
5	ลำลูกกา	กลาง	2	55
6	สายปัญญารังสิต	กลาง	2	53
7	ปทุมธานี “นันทมนี่บำรุง”	กลาง	2	52
8	ชัยสิทธิ์वास “พัฒนสายบำรุง”	เล็ก	1	28
9	มัธยมสังคีตวิทยา กรุงเทพมหานคร	เล็ก	1	25
10	สามโคก	เล็ก	1	25
11	สิงห์ปทุมราชภัฏวิทยา	เล็ก	1	24
12	วราชธานีรัตนมาตุวิทยา	เล็ก	1	23
	รวม	3	22	597

จากตาราง 4 โรงเรียนที่ใช้เป็นกลุ่มตัวอย่างมีจำนวน 12 โรงเรียน จากขนาดของโรงเรียน 3 ขนาด จำนวนห้องเรียน 22 ห้องเรียน และจำนวนนักเรียนกลุ่มตัวอย่างทั้งสิ้น 597 คน เพื่อใช้เป็นกลุ่มตัวอย่างในการทดสอบเพื่อหาคุณภาพข้อสอบเป็นรายข้อ เพื่อหาค่าพารามิเตอร์ของข้อสอบโดยใช้โมเดลโลจิสติกแบบ 3 พารามิเตอร์ หาคุณภาพของข้อสอบโดยตรวจสอบความเป็นมิติเดียวด้วยการวิเคราะห์องค์ประกอบ หาความเที่ยงตรงเชิงโครงสร้าง และหาความเชื่อมั่นของแบบทดสอบที่ผู้วิจัยสร้างเพิ่มเติมขึ้น

ประชากรและกลุ่มตัวอย่างสำหรับการพัฒนาโปรแกรม ศึกษาประสิทธิภาพของโปรแกรมการทดสอบ และใช้ในการวิจัย

ประชากรสำหรับการพัฒนาโปรแกรม ศึกษาประสิทธิภาพของโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ และใช้ในการวิจัย ได้แก่ นักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2546 โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษา ปทุมธานี เขตที่ 1 และ 2 จาก 176 โรงเรียน เป็นโรงเรียนขนาดใหญ่ 24 โรงเรียน โรงเรียนขนาดกลาง 64 โรงเรียน และโรงเรียนขนาดเล็ก 88 โรงเรียน จำนวนนักเรียนทั้งสิ้น 7,933 คน

กลุ่มตัวอย่างสำหรับการพัฒนาโปรแกรม ศึกษาประสิทธิภาพของโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ และใช้ในการวิจัย ได้แก่ นักเรียนชั้น

ประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2546 โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษา ปทุมธานี เขตที่ 1 และ 2 จาก 21 โรงเรียน จำนวนนักเรียน 968 คน ได้มาโดยวิธีการสุ่ม ซึ่งมีขั้นตอนในการสุ่มดังนี้

ขั้นตอนแรก สุ่มโรงเรียน 21 โรงเรียน จากโรงเรียนระดับประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2546 ของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษา ปทุมธานี เขตที่ 1 และ 2 โดยการสุ่มแบบแบ่งชั้น (stratified random sampling) ซึ่งใช้ขนาดของโรงเรียนเป็นชั้นของการสุ่ม ผลของการสุ่มได้โรงเรียนขนาดใหญ่ 4 โรงเรียน มีห้องเรียน 18 ห้อง โรงเรียนขนาดกลาง 6 โรงเรียน มีห้องเรียน 18 ห้องเรียน และโรงเรียนขนาดเล็ก 11 โรงเรียน มีห้องเรียน 17 ห้อง รวมโรงเรียนทั้งหมด 21 โรงเรียนมีห้องเรียนทั้งหมด 53 ห้อง

ขั้นตอนที่ 2 สุ่มห้องเรียนจากโรงเรียนที่ได้จากการสุ่มในขั้นตอนแรก โดยการสุ่มอย่างง่าย (simple random sampling) จากโรงเรียนขนาดใหญ่ โรงเรียนละ 3 ห้อง ได้ห้องเรียน 12 ห้อง จำนวนนักเรียน 335 คน จากโรงเรียนขนาดกลาง โรงเรียนละ 2 ห้อง ได้ห้องเรียน 12 ห้อง จำนวนนักเรียน 330 คน และได้จากโรงเรียนขนาดเล็ก โรงเรียนละ 1 ห้อง ได้ห้องเรียน 11 ห้อง จำนวนนักเรียน 303 คน รวมนักเรียนที่เป็นกลุ่มตัวอย่างทั้งสิ้น จำนวน 968 คน

ขั้นตอนที่ 3 สุ่มโรงเรียนจากขั้นตอนที่ 1 และห้องเรียนจากขั้นตอนที่ 2 ออกเป็น 3 กลุ่ม ได้แก่

กลุ่มที่ 1 สำหรับทดสอบเพื่อกำหนดเวลาให้กับข้อสอบในแต่ละข้อ โดยใช้กลุ่มตัวอย่างจากโรงเรียนขนาดใหญ่ 1 โรงเรียน ห้องเรียน 3 ห้อง ขนาดกลาง 1 โรงเรียน ห้องเรียน 2 ห้อง และขนาดเล็ก 3 โรงเรียน ห้องเรียน 3 ห้อง รวมโรงเรียน 5 โรงเรียน ห้องเรียนทั้งสิ้น 8 ห้อง จำนวนนักเรียน 235 คน ใช้เป็นกลุ่มตัวอย่างในการทดสอบเพื่อกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อของโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

กลุ่มที่ 2 สำหรับหาประสิทธิภาพของโปรแกรม โดยใช้กลุ่มตัวอย่างจากโรงเรียนขนาดใหญ่ 1 โรงเรียน ห้องเรียน 3 ห้อง และขนาดเล็ก 2 โรงเรียน ห้องเรียน 2 ห้อง รวมโรงเรียน 3 โรงเรียน ห้องเรียน 5 ห้อง จำนวนนักเรียน 135 คน ใช้เป็นกลุ่มตัวอย่างในการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ และทดสอบด้วยแบบสอบวัดความพึงพอใจในการสอบ เพื่อศึกษาประสิทธิภาพของโปรแกรม

กลุ่มที่ 3 สำหรับการท้าวิจัย โดยใช้กลุ่มตัวอย่างจากโรงเรียนขนาดใหญ่ 2 โรงเรียน ห้องเรียน 6 ห้อง โรงเรียนขนาดกลาง 6 โรงเรียน ห้องเรียน 12 ห้อง และโรงเรียนขนาดเล็ก 5 โรงเรียน ห้องเรียน 5 ห้อง รวมห้องเรียน 23 ห้องเรียน จำนวนนักเรียน 598 คน ใช้เป็นกลุ่มตัวอย่างในการวิจัยเพื่อศึกษาผลของตัวแปรอิสระที่มีต่อตัวแปรตาม ดังแสดงไว้ในตาราง 5

ตาราง 5 จำนวนนักเรียนกลุ่มตัวอย่าง จำแนกเป็นรายโรงเรียน ขนาดของโรงเรียน และห้องเรียน

ที่	โรงเรียน	ขนาด โรงเรียน	จำนวน ห้อง	กลุ่ม ตัวอย่าง	พัฒนาโปรแกรม		ทำวิจัย
					กลุ่มที่ 1	กลุ่มที่ 2	กลุ่มที่ 3
1	วัดชินวราราม	ใหญ่	3	85	✓		
2	วัดคลองหนึ่ง(แก้วนิมิตร)	ใหญ่	3	85			✓
3	วัดแสงสวรรค์	ใหญ่	3	85			✓
4	ชุมชนวัดบัวแก้วเกษร	ใหญ่	3	80		✓	
5	วัดบางเดื่อ	กลาง	2	60	✓		
6	ไทยรัฐวิทยา 69 (คลองหลวง)	กลาง	2	60			✓
7	วัดนาบุญ	กลาง	2	60			✓
8	ชุมชนวัดหน้าไม้	กลาง	2	50			✓
9	วัดมุจลินดาราม	กลาง	2	50			✓
10	วัดสระบัว	กลาง	2	50			✓
11	ทองพูนอุทิศ	เล็ก	1	30	✓		
12	บางขวดอนุสรณ์	เล็ก	1	30	✓	✓	
13	สังข์อำวิทยา	เล็ก	1	30	✓		
14	คลองสอง	เล็ก	1	30			
15	วัดดาวเรือง	เล็ก	1	38			
16	วัดเวฬุวัน	เล็ก	1	28			✓
17	วัดไพรีฟ้า	เล็ก	1	26			✓
18	วัดสุทธาวาส	เล็ก	1	26			✓
19	วัดบัวสุวรรณประดิษฐ์	เล็ก	1	25			✓
20	วัดสังฆรักษ์บำรุง	เล็ก	1	25		✓	
21	วัดบ่อเงิน	เล็ก	1	25			✓
	รวม	21	35	968	235	135	598

จากตาราง 5 โรงเรียนที่ใช้เป็นกลุ่มตัวอย่างมีจำนวน 21 โรงเรียน จากขนาดของโรงเรียน 3 ขนาด จำนวนห้องเรียน 35 ห้องเรียน และจำนวนนักเรียนกลุ่มตัวอย่างทั้งสิ้น 968 คน ใช้เป็นกลุ่มตัวอย่างในการทดสอบเพื่อกำหนดเวลาที่ใช้ในการตอบข้อสอบแต่ละข้อ จำนวนนักเรียนกลุ่มตัวอย่าง 235 คน ใช้เป็นกลุ่มตัวอย่างในการทดสอบเพื่อศึกษาประสิทธิภาพของโปรแกรม จำนวนนักเรียนกลุ่มตัวอย่าง 135 คน และใช้เป็นกลุ่มตัวอย่างในการวิจัยเพื่อศึกษาผลของตัวแปรอิสระที่มีต่อตัวแปรตาม จำนวนนักเรียนกลุ่มตัวอย่าง 598 คน

การดำเนินการวิจัยระยะที่ 1 : การพัฒนาเครื่องมือ

การพัฒนาข้อสอบและคลังข้อสอบ พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะ และศึกษาประสิทธิภาพของโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

การพัฒนาเครื่องมือที่ใช้ในการวิจัย

1. การพัฒนาข้อสอบและคลังข้อสอบ

ข้อสอบที่ใช้สำหรับสร้างคลังข้อสอบ ได้แก่ ข้อสอบวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 วัดความเข้าใจเรื่องเศษส่วน แบบเลือกตอบ 4 ตัวเลือก ที่ผู้วิจัยสร้างขึ้นจำนวน 44 ข้อ และข้อสอบของรังสรรค์ มณีเล็ก (2540) จำนวน 200 ข้อ เนื่องจากข้อสอบดังกล่าวมีกระบวนการสร้างอย่างเป็นระบบ มีคุณภาพที่น่าเชื่อถือ มีค่าพารามิเตอร์ของข้อสอบโดยเฉลี่ยค่าใกล้เคียงกัน กล่าวคือค่าอำนาจจำแนกประมาณ 1.0 ซึ่งอยู่ในเกณฑ์ดี ค่าความยากประมาณ 0.8 ซึ่งอยู่ในเกณฑ์ปานกลางค่อนข้างไปทางยาก และค่าการเดามีค่าประมาณ 0.2 ซึ่งอยู่ในเกณฑ์ดี ส่วนผลการตรวจสอบความเป็นมิติเดียวของข้อสอบวัดความเข้าใจเรื่องเศษส่วน โดยการวิเคราะห์องค์ประกอบ ปรากฏว่า ค่าไอเกน (Eigen value) ของตัวประกอบที่ 1 สูงกว่าค่าไอเกนของตัวประกอบอื่น ๆ ที่เหลือ และตัวประกอบที่เหลือดังกล่าวมีค่าไอเกนใกล้เคียงกัน และข้อสอบที่ผู้วิจัยสร้างขึ้นนั้น ผู้วิจัยได้สร้างตามฟิสิกส์การสร้างข้อสอบของรังสรรค์ มณีเล็ก ได้ข้อสอบที่ครอบคลุมฟิสิกส์ดังกล่าว และผู้เชี่ยวชาญด้านการสอนคณิตศาสตร์และผู้เชี่ยวชาญด้านการวัดผล พิจารณาแล้วเห็นว่าข้อสอบที่ผู้วิจัยสร้างเพิ่มเติมขึ้นสามารถวัดได้ตามโครงสร้างฟิสิกส์เดิม ดังนั้นจึงอนุมานได้ว่า ข้อสอบวัดความเข้าใจเรื่องเศษส่วนชั้นประถมศึกษาปีที่ 6 ที่ผู้วิจัยสร้างขึ้น และของรังสรรค์ มณีเล็ก มีความเป็นมิติเดียวตรงตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

กระบวนการสร้างแบบทดสอบคณิตศาสตร์ วัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 กระทำอย่างเป็นระบบ ตามขั้นตอนและทฤษฎีการสร้างแบบทดสอบครบถ้วน ผลการวิเคราะห์ข้อมูล แสดงถึงคุณภาพของแบบทดสอบอยู่ในเกณฑ์ที่มีคุณภาพน่าเชื่อถือ มีข้อสอบในคลังเดิมจำนวน 200 ข้อ และข้อสอบที่ผู้วิจัยสร้างขึ้นตามขั้นตอนและกระบวนการสร้างข้อสอบของรังสรรค์ มณีเล็ก อีกจำนวน 44 ข้อ รวมเป็นข้อสอบ 244 ข้อ ที่มีค่าความยากอยู่ในเกณฑ์ตามช่วงของความยากตั้งแต่ -3.0 ถึง 3.0 ซึ่งมีวิธีการสร้างข้อสอบ พอสรุปได้ดังนี้

1) ศึกษาฟิสิกส์ (facet) ของคำถามและตัวเลือก เพื่อใช้เป็นกรอบในการสร้างข้อสอบวัดความเข้าใจเรื่องเศษส่วนของรังสรรค์ มณีเล็ก

2) สร้างข้อสอบเพื่อวัดความเข้าใจเรื่องเศษส่วน จำนวน 54 ข้อ ให้ครอบคลุมฟิสิกส์ (facet) ของคำถามและตัวเลือก

3) นำข้อสอบไปให้ผู้เชี่ยวชาญด้านหลักสูตรการสอน การวัดและประเมินผล การศึกษาด้านคณิตศาสตร์ตรวจสอบความเที่ยงตรงตามเนื้อหาและความเที่ยงตรงตามโครงสร้าง โดยใช้เทคนิคของโรวินลลี และแฮมเบิลตัน (Rovinelli and Hambleton) กำหนดเกณฑ์ในการพิจารณา 3

เกณฑ์ คือ ตรง ไม่ตรงและไม่แน่ใจว่าตรงตามเนื้อหา และโครงสร้างที่ใช้วัดหรือไม่ ซึ่งวัดเนื้อหาเรื่อง เศษส่วน และวัดโครงสร้างความเข้าใจ

4) คัดเลือกข้อสอบที่เป็นไปตามเกณฑ์ในข้อ 3 และปรับปรุงข้อสอบให้เป็นไปตามเกณฑ์ตามคำแนะนำของผู้เชี่ยวชาญ

5) นำข้อสอบไปทดลองสอบกับกลุ่มตัวอย่างเพื่อวิเคราะห์หาคุณภาพรายข้อโดยใช้ ทฤษฎีการตอบสนองข้อสอบ วิเคราะห์ข้อสอบโดยใช้โมเดลโลจิสติกแบบ 3 พารามิเตอร์ คือค่าอำนาจ จำแนก (a) ค่าความยาก (b) และค่าการเดา (c) กับกลุ่มตัวอย่างจำนวน 597 คน

6) คัดเลือกข้อสอบได้ข้อสอบจำนวน 44 ข้อ จากข้อสอบทั้งหมด 54 ข้อที่มีคุณภาพ ค่าพารามิเตอร์เป็นไปตามเกณฑ์ กล่าวคือ ค่าอำนาจจำแนกตั้งแต่ 0.8 ขึ้นไป ค่าความยากอยู่ในช่วง -3.0 ถึง 3.0 และค่าการเดาน้อยกว่าหรือเท่ากับ 0.3 ซึ่งค่าพารามิเตอร์ดังกล่าวนี้ จะถูกนำไปเก็บไว้ในคลังข้อสอบรวมกับข้อสอบของรังสรรค์ มณีเล็ก เพื่อใช้ในการประมาณค่าความสามารถของผู้สอบ และใช้เป็นดัชนีสำหรับคัดเลือกข้อสอบจากคลังข้อสอบออกมาให้ผู้สอบตอบบนหน้าจอคอมพิวเตอร์ ดังนั้น จึงมี ข้อสอบในคลังเพื่อใช้ในการทำวิจัยในครั้งนี้ รวมทั้งสิ้นจำนวน 244 ข้อ

2. การพัฒนาโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะ

โปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ที่ใช้ในการค้นคว้าวิจัยในครั้งนี้ ผู้วิจัยได้พัฒนาโปรแกรมตามลำดับขั้นดังนี้

2.1 การศึกษาข้อมูลเพื่อพัฒนาโปรแกรมการทดสอบ

ขั้นตอนการดำเนินการศึกษาวิจัยเพื่อเก็บข้อมูล และวิเคราะห์ข้อมูลเพื่อกำหนดคุณลักษณะของข้อสอบตามตัวแปรที่ศึกษา

จากการศึกษาค่าพารามิเตอร์ความยากของข้อสอบวัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 ของรังสรรค์ มณีเล็ก จำนวน 200 ข้อ และข้อสอบของผู้วิจัยสร้างเพิ่มเติมขึ้น จำนวน 44 ข้อ รวมข้อสอบทั้งสิ้น จำนวน 244 ข้อ ซึ่งเรียงลำดับข้อตามความยากง่ายของข้อสอบ จากข้อง่ายไปหาข้อยาก เพื่อใช้เป็นแนวทางในการพิจารณากำหนดการจัดชุดข้อสอบเพื่อนำเสนอให้ผู้สอบตอบในแต่ละชั้น และใช้เป็นข้อมูลพื้นฐานในการศึกษาเพื่อกำหนดเวลาให้ตอบข้อสอบในแต่ละข้อ ซึ่งผู้วิจัยได้ทำตามลำดับขั้นดังนี้

2.1.1 การจัดชุดข้อสอบให้ตอบในแต่ละชั้น

ผู้วิจัยได้พิจารณาแนวทางในการกำหนดการจัดชุดข้อสอบเพื่อนำเสนอให้ผู้สอบตอบในแต่ละชั้น ตามลำดับขั้นดังนี้

(1) การจัดชุดข้อสอบให้ตอบชั้นละ 1 ข้อ เป็นการนำเสนอข้อสอบให้ผู้สอบตอบตามธรรมดาโดยทั่วไป กล่าวคือหลังจากทราบผลการประมาณค่าความสามารถของผู้สอบที่ตอบข้อสอบข้อที่ผ่านมา คอมพิวเตอร์จะนำเสนอข้อสอบข้อต่อไปอีก 1 ข้อ ตามผลการตอบก่อนหน้านั้น เมื่อคอมพิวเตอร์คำนวณประมาณค่าความสามารถของผู้สอบแล้ว คอมพิวเตอร์จะคัดเลือกข้อสอบข้อต่อไป โดยดูค่าฟังก์ชันสารสนเทศของข้อสอบที่สูงที่สุด 1 ข้อ ณ ความสามารถของผู้สอบในขณะนั้น นำเสนอเป็นข้อต่อไป

(2) การจัดชุดข้อสอบให้ตอบขั้นละ 2 ข้อ เป็นการจัดชุดข้อสอบเสนอแก่ผู้เข้าสอบครั้งละ 2 ข้อ แล้วจึงทำการประมาณค่าความสามารถของผู้สอบทั้งนี้เพื่อให้เกิดความมั่นใจในการเลือกเส้นทางของข้อสอบว่า การที่จะเลือกข้อสอบชุดต่อไปนั้น จะเพิ่มค่าความยากมากขึ้น หรือจะลดค่าความยากลง หรือยังคงจัดชุดของข้อสอบที่มีค่าความยากเท่าเดิมให้ผู้สอบอีกครั้ง ข้อสอบทั้ง 2 ข้อที่นำเสนอแต่ละครั้งนั้นจะต้องมีค่าสารสนเทศของข้อสอบที่สูงที่สุดเท่ากัน หรือข้อที่มีค่าสารสนเทศของข้อสอบที่สูงที่สุดใกล้เคียงกันมากที่สุด

(3) การจัดชุดข้อสอบให้ตอบขั้นละ 3 ข้อ เป็นการจัดชุดข้อสอบเสนอแก่ผู้เข้าสอบ ครั้งละ 3 ข้อ แล้วจึงทำการประมาณค่าความสามารถของผู้สอบทั้งนี้เพื่อเพิ่มความมั่นใจมากยิ่งขึ้นในการเลือกเส้นทางของข้อสอบว่าจะเลือกชุดของข้อสอบต่อไปนั้น จะเพิ่มค่าความยากมากขึ้น หรือลดค่าความยากลง หรือยังคงจัดชุดของข้อสอบที่มีค่าความยากเท่าเดิมอีกครั้ง ข้อสอบทั้งสามข้อที่นำเสนอแต่ละครั้งนั้นจะต้องมีค่าสารสนเทศของข้อสอบที่สูงที่สุดเท่ากันทุกข้อหรือมีค่าสูงที่สุดใกล้เคียงกันมากที่สุด

2.1.2 การกำหนดเวลาให้ตอบในแต่ละข้อ

การวิจัยในครั้งนี้ผู้วิจัยได้กำหนดการใช้เวลาในการตอบในแต่ละข้อเป็น 3 ลักษณะด้วยกัน ในแต่ละลักษณะมีการกำหนดเวลาที่ใช้ในการตอบดังนี้

(1) ไม่กำหนดเวลาที่ใช้ในการตอบในแต่ละข้อ เป็นการสอบที่เปิดโอกาสให้ผู้สอบใช้เวลาอย่างเต็มที่ในการตอบข้อสอบในแต่ละข้อ ซึ่งเป็นลักษณะการสอบแบบปกติที่ใช้กันโดยทั่วไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

(2) การกำหนดเวลาให้ตอบในแต่ละข้อเฉลี่ยเท่ากันทุกข้อ กล่าวคือข้อสอบทุก ๆ ข้อในคลังข้อสอบจะกำหนดเวลาที่ใช้ในการตอบข้อสอบเท่ากันทุก ๆ ข้อ เวลาเฉลี่ยที่กำหนดนี้ได้จากการนำข้อสอบทั้งหมดไปทำการทดสอบกับกลุ่มตัวอย่างแล้วนำเวลาที่ใช้ในการสอบมาหาค่าเฉลี่ยเพื่อกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อ

(3) การกำหนดเวลาให้ตอบในแต่ละข้อตามน้ำหนักค่าความยากของข้อสอบ กล่าวคือถ้าเป็นข้อสอบยากจะต้องใช้เวลาในการตอบมากกว่าข้อสอบง่าย ซึ่งมีขั้นตอนในการกำหนดเวลา ดังนี้

ขั้นที่ 1 นำข้อสอบมาจัดแบ่งเป็นชุดตามระดับความยากของข้อสอบซึ่งแบ่งตามช่วงค่าความยาก ช่วงละ .50 โดยเริ่มแบ่งจากค่าความยากที่ 0.00 ไปยังค่าความยาก -3.00 และ 3.00 ได้ชุดของข้อสอบจำนวน 12 ชุด แสดงได้ดังตาราง 6

ตาราง 6 การแบ่งชุดข้อสอบตามช่วงระดับความยาก

ชุดที่	ช่วงค่าความยาก	ข้อสอบข้อที่	จำนวนข้อ
1	-3.00 ถึง -2.51		0
2	-2.50 ถึง -2.01	71	1
3	-2.00 ถึง -1.51	201, 213, 147, 32, 65	5
4	-1.50 ถึง -1.01	195, 146, 114, 39, 38, 96, 120	7
5	-1.00 ถึง -.051	116, 134, 26, 81, 137, 108, 104, 55, 42, 176, 222, 84, 103, 163, 154, 174, 19, 49, 182, 100, 101	21
6	-0.50 ถึง -0.01	232, 87,180, 233, 54, 126, 183, 56, 212, 77, 68, 226, 123, 109, 76, 15, 229, 82, 205, 184, 75, 72, 208, 78, 170, 196, 106, 155, 166, 11, 62, 159, 50, 221, 46, 14, 53	37
7	0.00 ถึง 0.50	35, 122, 202, 79, 164, 227, 138, 151, 230, 125, 156, 175, 99, 216, 192, 217, 215, 190, 8, 210, 211,7, 27, 51, 209, 206, 124, 60, 88, 191, 74, 207, 139, 169, 228, 73, 63, 237, 57, 218, 43, 160, 135, 225, 167	45
8	0.51 ถึง 1.00	105, 1, 148, 83, 113, 58, 214, 133, 13, 91, 112, 130, 136, 98, 2, 220, 194, 18, 127, 95, 85, 234, 165, 188, 223, 40, 23, 86, 204, 70, 224, 25, 61, 179, 173, 24, 110, 181, 29, 219	40
9	1.01 ถึง 1.50	153, 131, 22, 162, 150, 193, 197, 92, 118, 93, 6, 186, 132, 231, 144, 89, 203, 129, 17, 36, 111, 52, 34, 59, 47, 236, 94, 152, 168, 115, 5, 30, 240, 189, 121, 243, 171, 157, 149, 9, 178, 67	42
10	1.51 ถึง 2.00	239, 200, 185, 21, 41, 187, 172, 45, 141, 199, 20, 66, 142, 143, 44, 28, 140, 37, 242, 244, 80, 48, 145, 12	24
11	2.01 ถึง 2.50	161, 10, 117, 107, 33, 198, 177, 238, 4, 3, 64, 119, 235, 241, 90, 69, 102, 128,	18
12	2.51 ถึง 3.00	97, 31, 158, 16	4
รวมทั้งสิ้น			244

จากตาราง 6 เป็นข้อสอบวัดความเข้าใจเรื่องเศษส่วนจำนวนทั้งสิ้น 244 ข้อและจัดแบ่งข้อสอบออกเป็น 12 ชุด คือชุดที่ 1 เป็นข้อสอบที่ง่ายมากแต่ไม่มีข้อสอบในช่วงนี้ ชุดที่ 2 เป็นข้อสอบง่ายจำนวน 1 ข้อ ชุดที่ 3 เป็นข้อสอบที่ยากกว่าชุดที่ 1 และ 2 จำนวน 5 ข้อ ส่วนข้อสอบชุดต่อไปจะยากขึ้นเรื่อย ๆ คือ ชุดที่ 4 จำนวน 7 ข้อ ชุดที่ 5 จำนวน 21 ข้อ ชุดที่ 6 จำนวน 37 ข้อ ชุดที่ 7

จำนวน 45 ข้อ ชุดที่ 8 จำนวน 40 ข้อ ชุดที่ 9 จำนวน 42 ข้อ ชุดที่ 10 จะเป็นข้อสอบยากกว่าทุกชุดที่กล่าวมา จำนวน 24 ข้อ ชุดที่ 11 เป็นข้อสอบที่ยากมาก จำนวน 18 ข้อ และชุดที่ 12 เป็นข้อสอบยากที่สุด จำนวน 4 ข้อ

ขั้นที่ 2 นำข้อสอบ 8 ชุด คือชุดที่มีจำนวนข้อสอบมากพอ ยกเว้นชุดที่ 1, 2, 3 และ 12 ที่มีข้อสอบน้อยมาก จากขั้นที่ 1 มาเก็บไว้ในคลังข้อสอบซึ่งเป็นคลังข้อสอบที่มีโปรแกรมคอมพิวเตอร์ ช่วยในการดำเนินการสอบ รวม 8 โปรแกรมย่อย ซึ่งในแต่ละโปรแกรมจะบรรจุข้อสอบแต่ละชุดจัดเป็นคลังย่อยเพื่อเก็บข้อสอบไว้ จำนวนข้อสอบในแต่ละคลังย่อย มีจำนวนข้อสอบ 20 ข้อ ยกเว้นชุดที่ 1 มี 7 ข้อและชุดที่ 8 มี 18 ข้อ โดยการคัดเลือกข้อสอบให้เป็นตัวแทนในแต่ละช่วง ทั้งนี้ในการกำหนดให้มีจำนวนข้อสอบ 20 ข้อ เพราะจากผลการใช้คลังข้อสอบที่บรรจุข้อสอบชุดนี้ที่ผ่านมา ผู้สอบจะใช้ข้อสอบโดยเฉลี่ย 20 ข้อ และเพื่อไม่ให้เกิดความเมื่อยล้าจากการทำข้อสอบที่ยาวเกินไปอันอาจมีผลต่อการใช้เวลาที่คลาดเคลื่อนไป และการกำหนดเวลาในการตอบในแต่ละชุดจะเป็นไปตามเกณฑ์เดียวกัน เมื่อข้อสอบเท่ากันทุกชุดดังกล่าว ข้อสอบแต่ละชุดที่นำไปทดลองสอบเพื่อกำหนดเวลา แสดงไว้ใน ตาราง 7

ตาราง 7 จำนวนชุดข้อสอบที่นำไปทดสอบเพื่อกำหนดเวลาที่ใช้ในการสอบ

ชุดที่	ช่วงค่าความยาก	ข้อสอบข้อที่	จำนวนข้อ
1	-1.50 ถึง -1.01	195, 146, 114, 39, 38, 96, 120	7
2	-1.00 ถึง -0.51	116, 134, 26, 81, 137, 108, 104, 55, 42, 176, 84, 103, 163, 154, 174, 19, 49, 182, 100, 101	20
3	-0.50 ถึง -0.01	232, 180, 54, 183, 212, 77, 109, 76, 15, 82, 184, 75, 72, 78, 106, 155, 159, 221, 14, 53	20
4	0.00 ถึง 0.50	35, 202, 164, 138, 230, 156, 99, 215, 211, 27, 124, 60, 191, 74, 139, 228, 63, 57, 160, 167	20
5	0.51 ถึง 1.00	105, 1, 83, 58, 133, 91, 130, 98, 220, 18, 95, 234, 188, 40, 86, 70, 25, 179, 24, 219	20
6	1.01 ถึง 1.50	153, 22, 150, 197, 118, 6, 132, 144, 203, 17, 111, 34, 47, 94, 168, 5, 240, 121, 171, 67	20
7	1.51 ถึง 2.00	239, 185, 41, 172, 141, 199, 20, 66, 142, 143, 44, 28, 140, 37, 242, 244, 80, 48, 145, 12	20
8	2.01 ถึง 2.50	161, 10, 117, 107, 33, 198, 177, 238, 4, 3, 64, 119, 235, 241, 90, 69, 102, 128,	18
รวมทั้งสิ้น			145

ขั้นที่ 3 นำข้อสอบ 8 ชุดไปสอบกับนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 8 ห้องเรียน นักเรียน 235 คน โดยวิธีการจัดผู้เข้าสอบอย่างมีระบบ คือ นักเรียนในแต่ละห้องจะต้องทดสอบด้วยแบบทดสอบทั้ง 8 ชุด กล่าวคือ นักเรียนคนที่ 1 ถึงคนที่ 8 จะทำการทดสอบกับข้อสอบชุดที่ 1 ถึง 8 ตามลำดับ นักเรียนคนที่ 9 ถึง 16 จะทำการทดสอบกับข้อสอบชุดที่ 1 ถึง 8 ตามลำดับเช่นกัน และทำการสอบกับผู้เข้าสอบตามลำดับข้างต้นจนครบทั้ง 8 ห้อง

ขั้นที่ 4 วิเคราะห์และหาค่าสถิติข้อสอบแต่ละชุดโดยวิธีหาค่าเฉลี่ยของเวลาที่ใช้ในการตอบแต่เพื่อไม่ให้การกำหนดเวลาเป็นการทดสอบแบบ Speed Test ผู้วิจัยจึงเพิ่มเวลาให้กับข้อสอบโดยการนำค่าเฉลี่ยของเวลาที่ทำแบบทดสอบทั้งฉบับรวมกับค่าความเบี่ยงเบนมาตรฐานก่อน จึงหาค่าเฉลี่ยเป็นรายข้อ

ขั้นที่ 5 นำค่าสถิติของข้อสอบที่ค้นพบดังกล่าวข้างต้น ไปใช้กำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อ ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ โดยกำหนดเวลาดังนี้

(1) ไม่กำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อ

(2) กำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อ เฉลี่ยเท่ากันทุกข้อ

ตามค่าเฉลี่ยของเวลาที่ได้จากการทดลองสอบตามกระบวนการข้างต้น

(3) กำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อของแต่ละชุด เฉลี่ยตามน้ำหนักค่าความยากของข้อสอบในชุดนั้น ๆ ซึ่งผู้วิจัยจัดแบ่งข้อสอบตามระดับความยากออกเป็น 12 ชุด และกำหนดเวลาในการตอบข้อสอบในชุดเดียวกันให้เวลาเท่ากัน ตามค่าเฉลี่ยเวลาของข้อสอบแต่ละชุดที่ค้นพบ ส่วนชุดที่ 1, 2, 3 และ 12 ในชุดดังกล่าวที่ไม่ได้นำไปทดลองสอบเนื่องจากมีจำนวนข้อสอบน้อยมากในชุดดังกล่าวนั้น จึงทำการกำหนดเวลาในการตอบโดยดูแนวโน้มการลดลงของเวลาเมื่อข้อสอบง่าย และแนวโน้มการเพิ่มขึ้นเมื่อข้อสอบยากขึ้น เพื่อกำหนดเวลาให้กับข้อสอบ

2.1.3 การสร้างรูปแบบของข้อสอบที่มีสื่อประสมประกอบการทดสอบ

การวิจัยในครั้งนี้ ผู้วิจัยได้กำหนดรูปแบบของข้อสอบเป็น 2 รูปแบบ คือ รูปแบบของข้อสอบที่ไม่มีสื่อประสมประกอบการทดสอบ และรูปแบบของข้อสอบที่มีสื่อประสมประกอบการทดสอบ ซึ่งรูปแบบของข้อสอบดังกล่าวมีลักษณะแตกต่างกันดังนี้

1) รูปแบบของข้อสอบที่ไม่มีสื่อประสมประกอบการทดสอบเป็นรูปแบบของการนำเสนอข้อสอบบนหน้าจอคอมพิวเตอร์ที่ประกอบด้วยตัวข้อสอบ ตัวเลือก และฉากหลังบนหน้าจอคอมพิวเตอร์ที่มีสีพื้นเพียงสีเดียวเหมือนกันทุกข้อ ตลอดช่วงระยะเวลาการทดสอบ

2) รูปแบบของข้อสอบที่มีสื่อประสมประกอบการทดสอบ เป็นรูปแบบของการนำเสนอข้อสอบบนหน้าจอคอมพิวเตอร์ที่ประกอบไปด้วย ตัวข้อสอบ ตัวเลือก ภาพกราฟฟิก ภาพวีว สีสั่น และเสียงเพลงบรรเลงประกอบการทดสอบขณะดำเนินการสอบ ข้อสอบแต่ละข้อจะมีรูปแบบของสื่อประสมแตกต่างกัน ขึ้นอยู่กับคุณลักษณะหรือองค์ประกอบของข้อสอบแต่ละข้อนั้นด้วย

สำหรับการสร้างรูปแบบของข้อสอบที่มีสื่อประสมประกอบการทดสอบในครั้งนี้ ผู้วิจัยได้คัดเลือกสื่อประสมบางส่วนที่ใช้กับแบบทดสอบวัดผลการเรียนหลังการใช้คอมพิวเตอร์มัลติมีเดียแบบฝึกของ ศิริพงษ์ ภูพันนา (2546) ที่ได้พัฒนาบทเรียนคอมพิวเตอร์มัลติมีเดียแบบฝึก และได้

พัฒนาคอมพิวเตอร์มัลติมีเดียสำหรับการทดสอบหลังการใช้แบบฝึกขั้นพร้อม ๆ กันนี้ด้วย เนื่องจาก การวิจัยดังกล่าวได้พัฒนาสื่อประสมโดยผ่านการพิจารณาและประเมินจากผู้เชี่ยวชาญด้านเทคโนโลยี หลายท่าน ผู้วิจัยจึงเชื่อว่าสื่อประสมดังกล่าวมีคุณภาพและสามารถนำมาพัฒนาใช้กับการทดสอบแบบ ปรับเหมาะของผู้วิจัยได้ ดังนั้นการวิจัยในครั้งนี้ ผู้วิจัยจึงเลือกสื่อประสมบางส่วนจากผลการวิจัย ดังกล่าวมาพัฒนาสร้างรูปแบบข้อสอบแบบปรับเหมาะที่มีสื่อประสมประกอบการทดสอบในครั้งนี้

2.2 กระบวนการพัฒนาโปรแกรมการทดสอบ

การพัฒนาโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะ ผู้วิจัย ดำเนินการดังนี้

ทำการศึกษาผลการวิจัย ของรังสรรค์ มณีเล็ก (2540) ที่ค้นพบผลของตัวแปร ที่ส่งผลต่อคุณภาพของการทดสอบแบบปรับเหมาะ กล่าวคือ ถ้าเริ่มต้นสอบด้วยข้อสอบที่มีความยาก ง่ายปานกลางจะทำให้ใช้จำนวนข้อสอบน้อยกว่าการเริ่มต้นสอบจากข้อยากหรือข้อง่าย วิธีคัดเลือก ข้อสอบ ข้อต่อไปในคลังข้อสอบควรใช้ค่าสารสนเทศของข้อสอบที่สูงที่สุดจะทำให้ใช้ข้อสอบน้อยลง และเหมาะกับคลังข้อสอบขนาดใหญ่ การประมาณค่าความสามารถของผู้สอบด้วยวิธีเบส์ให้ค่าความ เทียงตรงเชิงสภาพสูงกว่า และเกณฑ์การยุติการสอบ ควรใช้เกณฑ์ค่าความคลาดเคลื่อนในการ ประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.3 และจากการพัฒนาโปรแกรมการทดสอบ แบบปรับเหมาะกับความสามารถของผู้สอบของสายชล ออบทม (2540) ซึ่งได้พัฒนาโปรแกรมการ ทดสอบแบบปรับเหมาะอย่างมีขั้นตอนเป็นระบบ โปรแกรมสามารถรองรับการทำงานใน ระบบปฏิบัติการวินโดวส์ 95 และผลจากการประเมินการใช้โปรแกรมแสดงถึงคุณภาพของโปรแกรมอยู่ ในเกณฑ์ที่มีคุณภาพ และมีประสิทธิภาพ กล่าวคือ โปรแกรมการทดสอบสามารถรองรับข้อสอบแบบ เลือกตอบที่มีตัวเลือก 3 ถึง 5 ตัวเลือก รองรับข้อสอบที่ผ่านการวิเคราะห์ค่าพารามิเตอร์แบบ 3 พารามิเตอร์ ตามทฤษฎีการตอบสนองข้อสอบ (IRT) และมีกลุ่มของคำสั่งต่าง ๆ ของโปรแกรมเพื่อใช้ ในการทำงานได้อย่างสะดวก มีความทันสมัย ผู้สอบมีความสนใจในการทดสอบมากยิ่งขึ้น ดังนั้นผู้วิจัย จึงเลือกข้อค้นพบของรังสรรค์ มณีเล็ก และเลือกแนวทางการพัฒนาโปรแกรมการทดสอบของสายชล ออบทม มาพัฒนาเป็นโปรแกรมการทดสอบที่ใช้ในการศึกษาค้นคว้าของผู้วิจัย ซึ่งผู้วิจัยได้พัฒนา โปรแกรมและสร้างข้อกำหนดคุณลักษณะของข้อสอบขึ้นใหม่ที่คาดว่าจะส่งผลต่อคุณภาพของการ ทดสอบแบบปรับเหมาะยิ่งขึ้น คือ การจัดชุดข้อสอบให้ตอบในแต่ละชั้น ต่างกัน 3 รูปแบบ การใช้เวลา ในการตอบในแต่ละข้อต่างกัน 3 วิธี และการสร้างรูปแบบของข้อสอบที่มีสื่อประสมกับไม่มีสื่อประสม 2 รูปแบบ จะส่งผลต่อเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของ แบบทดสอบอย่างน้อยแค่ไหนอย่างไร จึงได้รูปแบบของโปรแกรมการทดสอบแบบปรับเหมาะกับความ สามารถของผู้สอบด้วยคอมพิวเตอร์ จำนวนทั้งสิ้น 18 โปรแกรมที่เกิดจากการจัดหมู่ของตัวแปร ดังกล่าว ดังแสดงรายละเอียดของโปรแกรมในตาราง 8

ตาราง 8 ตารางแสดงการกำหนดคุณลักษณะของข้อสอบตามตัวแปรที่ศึกษา

โปรแกรมย่อย	จำนวนข้อสอบ ที่ใช้สอบในแต่ละชั้น	เวลาที่ใช้ในการตอบแต่ละข้อ	รูปแบบของข้อสอบ
1	1	ไม่จำกัดเวลา	ไม่มีสื่อประสม
2	1	ไม่จำกัดเวลา	มีสื่อประสม
3	1	จำกัดเวลาเฉลี่ยเท่ากันทุกข้อ	ไม่มีสื่อประสม
4	1	จำกัดเวลาเฉลี่ยเท่ากันทุกข้อ	มีสื่อประสม
5	1	จำกัดเวลาตามน้ำหนักค่าความยาก	ไม่มีสื่อประสม
6	1	จำกัดเวลาตามน้ำหนักค่าความยาก	มีสื่อประสม
7	2	ไม่จำกัดเวลา	ไม่มีสื่อประสม
8	2	ไม่จำกัดเวลา	มีสื่อประสม
9	2	จำกัดเวลาเฉลี่ยเท่ากันทุกข้อ	ไม่มีสื่อประสม
10	2	จำกัดเวลาเฉลี่ยเท่ากันทุกข้อ	มีสื่อประสม
11	2	จำกัดเวลาตามน้ำหนักค่าความยาก	ไม่มีสื่อประสม
12	2	จำกัดเวลาตามน้ำหนักค่าความยาก	มีสื่อประสม
13	3	ไม่จำกัดเวลา	ไม่มีสื่อประสม
14	3	ไม่จำกัดเวลา	มีสื่อประสม
15	3	จำกัดเวลาเฉลี่ยเท่ากันทุกข้อ	ไม่มีสื่อประสม
16	3	จำกัดเวลาเฉลี่ยเท่ากันทุกข้อ	มีสื่อประสม
17	3	จำกัดเวลาตามน้ำหนักค่าความยาก	ไม่มีสื่อประสม
18	3	จำกัดเวลาตามน้ำหนักค่าความยาก	มีสื่อประสม

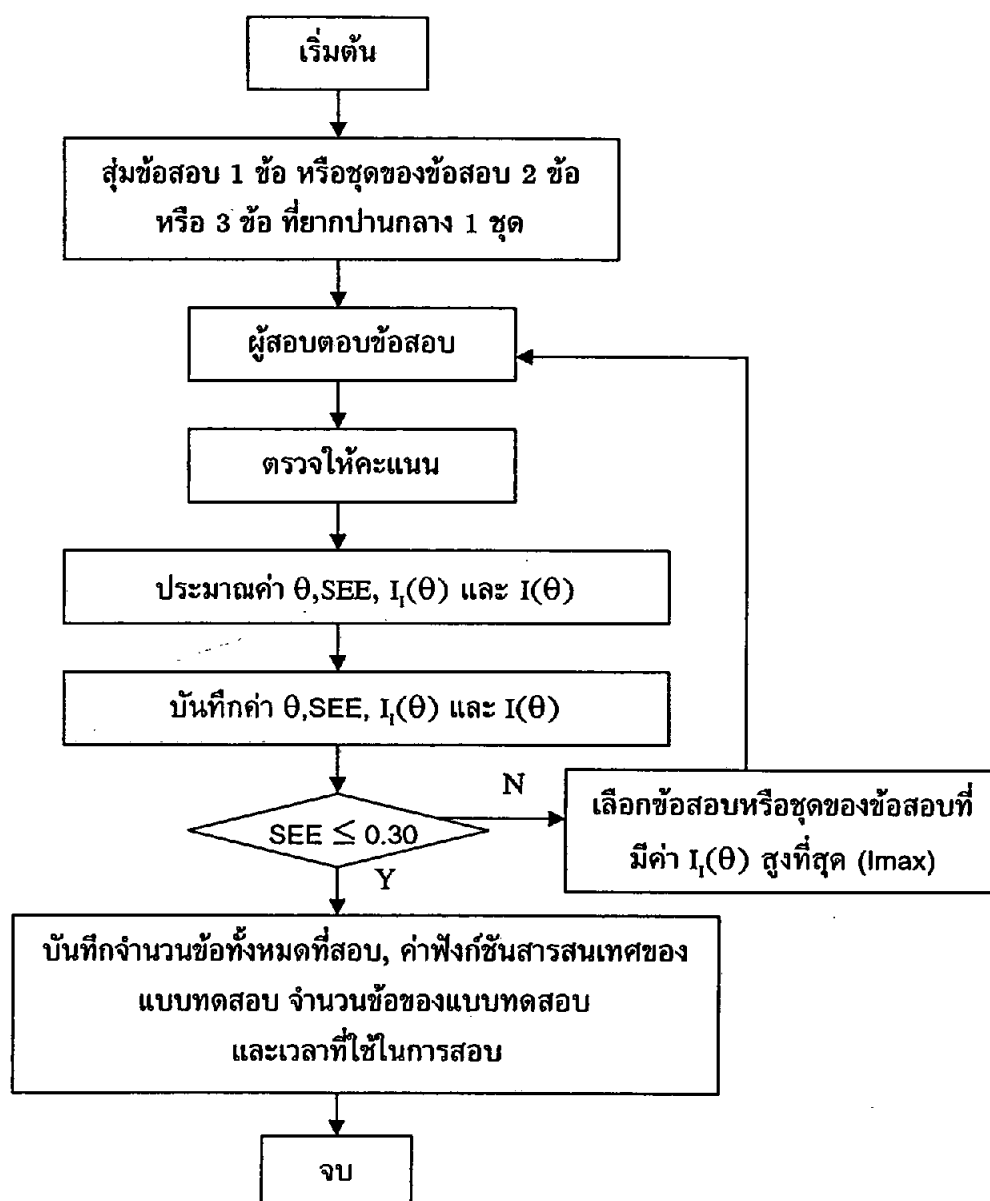
ตาราง 8 ตารางแสดงการกำหนดคุณลักษณะของข้อสอบตามตัวแปรที่ศึกษา ได้โปรแกรมจากการกำหนดคุณลักษณะตัวแปร 3 ตัว ได้แก่ การกำหนดจำนวนข้อสอบให้ตอบในแต่ละชั้น ได้ 3 แบบ การจำกัดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อ 3 แบบ และการจัดรูปแบบของข้อสอบ 2 รูปแบบ รวมได้โปรแกรมย่อยจำนวนทั้งสิ้น 18 โปรแกรม

การพัฒนาโปรแกรมการทดสอบ

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ทั้ง 18 โปรแกรมย่อยตามตารางข้างต้นนี้ ผู้วิจัยได้พัฒนาโปรแกรมคอมพิวเตอร์ดังกล่าวตามขั้นตอนดังนี้

1. เขียนผังโปรแกรมการทำงานของคอมพิวเตอร์ในการดำเนินการสอบ การวิเคราะห์ผล การสอบ และการบันทึกผลการสอบ ซึ่งมีผังโปรแกรมจำนวน 18 โปรแกรม ที่ได้จากการจัดหมู่ของตัวแปรอิสระ 3 ตัวคือ การจัดชุดข้อสอบให้ตอบในแต่ละชั้น ซึ่งมี 3 รูปแบบ ประกอบด้วย จัดชุดข้อสอบให้ตอบชั้นละ 1 ข้อ จัดชุดข้อสอบให้ตอบชั้นละ 2 ข้อ และจัดชุดข้อสอบให้ตอบชั้นละ 3 ข้อ การ

จำกัดเวลาที่ใช้ในการตอบแต่ละข้อมี 3 วิธี ประกอบด้วย ไม่จำกัดเวลาที่ใช้ในการตอบแต่ละข้อ จำกัดเวลาที่ใช้ในการตอบแต่ละข้อเฉลี่ยเท่ากันทุกข้อ และจำกัดเวลาที่ใช้ในการตอบแต่ละข้อตามน้ำหนักค่าความยากของข้อสอบและการสร้างรูปแบบของข้อสอบ 2 รูปแบบ ประกอบด้วยรูปแบบของข้อสอบที่ไม่มีสื่อประสมและรูปแบบของข้อสอบที่มีสื่อประสมประกอบการสอบ ดังตัวอย่างผังโปรแกรมตามภาพประกอบ 15



ภาพประกอบ 15 แสดงผังการทำงานของโปรแกรมคอมพิวเตอร์ในการดำเนินการสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

2. ผู้วิจัยทำการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ขึ้นให้โปรแกรมการทดสอบสามารถทำงานในระบบปฏิบัติการวินโดวส์ 98 (Windows

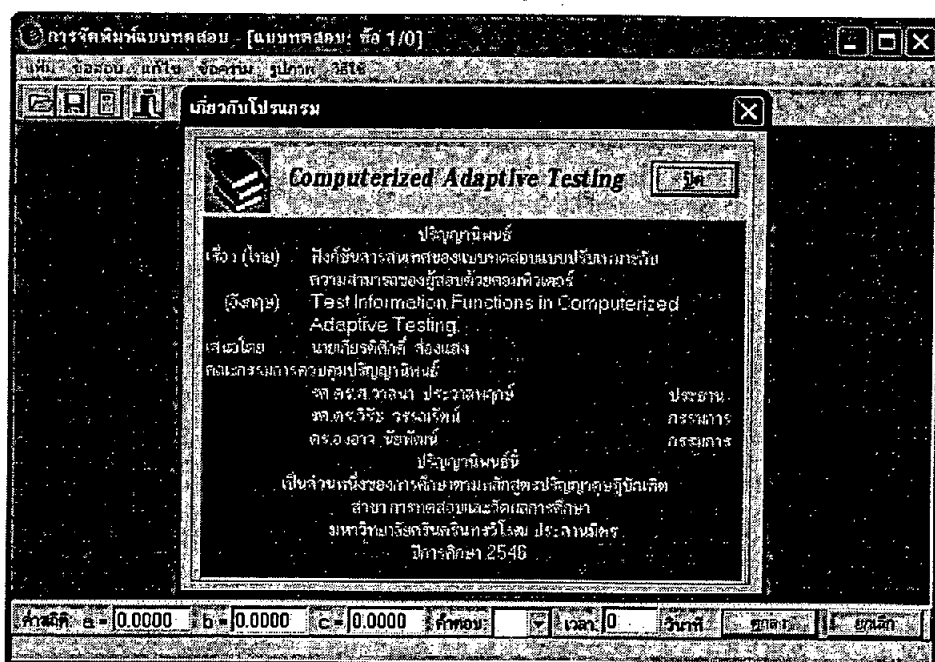
98) ขึ้นไป เนื่องจากเป็นระบบปฏิบัติการที่ใช้กันแพร่หลายในเครื่องคอมพิวเตอร์ในปัจจุบัน และเลือกใช้โปรแกรมบอร์แลนเดลฟาย (BORLAND DELPHI) ในการเขียนโปรแกรมการทดสอบแบบปรับเหมาะ เพราะเป็นภาษาที่สามารถเขียนภาพกราฟฟิคได้ เขียนสูตรสมการฟังก์ชันทางคณิตศาสตร์ได้ และเป็นภาษาที่มีโครงสร้างชัดเจนง่ายต่อการใช้งาน และสามารถใช้กับเครื่องไมโครคอมพิวเตอร์ได้

3. คัดเลือกข้อสอบข้อแรกเพื่อทำการเริ่มต้นสอบด้วยข้อสอบยากง่ายปานกลาง การคัดเลือกข้อสอบหรือชุดของข้อสอบถัดไปด้วยการใช้ค่าสารสนเทศของข้อสอบหรือชุดของข้อสอบที่สูงที่สุด การประมาณค่าความสามารถของผู้สอบใช้วิธีของเบส์ เกณฑ์การยุติการสอบใช้เกณฑ์ค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบต่ำกว่าหรือเท่ากับ 0.3 และเลือกใช้รูปแบบการทดสอบหลายขั้นตอนแบบทางแยกแปรผัน

4. เขียนโปรแกรมคอมพิวเตอร์ตามแผนผังของโปรแกรมตามที่ได้กำหนดไว้ ประกอบด้วยโปรแกรมสำหรับการพิมพ์เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบ และโปรแกรมการทดสอบแบบปรับเหมาะ รวมทั้งหมด 18 โปรแกรม ซึ่งเป็นโปรแกรมการทดสอบแบบปรับเหมาะตามโครงสร้างคุณลักษณะของตัวแปรที่ผู้วิจัยได้ศึกษาค้นคว้าไว้ในการวิจัยระยะที่ 1 คือตัวแปรการจัดชุดข้อสอบให้ตอบในแต่ละชั้น การกำหนดเวลาที่ใช้ในการตอบแต่ละข้อ และรูปแบบของข้อสอบที่ไม่มีสื่อประสมและมีสื่อประสมประกอบการทดสอบ ดังภาพตัวอย่างของโปรแกรม ต่อไปนี้

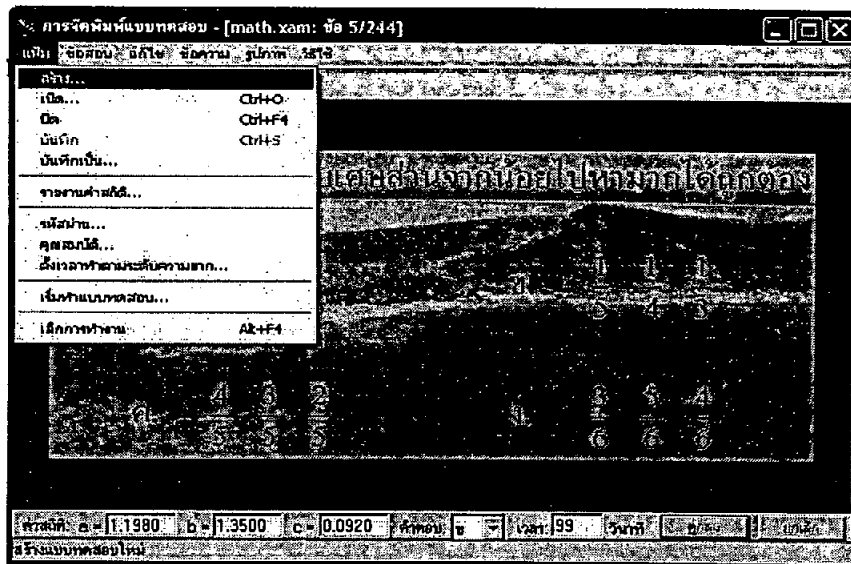
4.1 ภาพตัวอย่างของโปรแกรมสำหรับการพิมพ์เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไข

ข้อสอบ



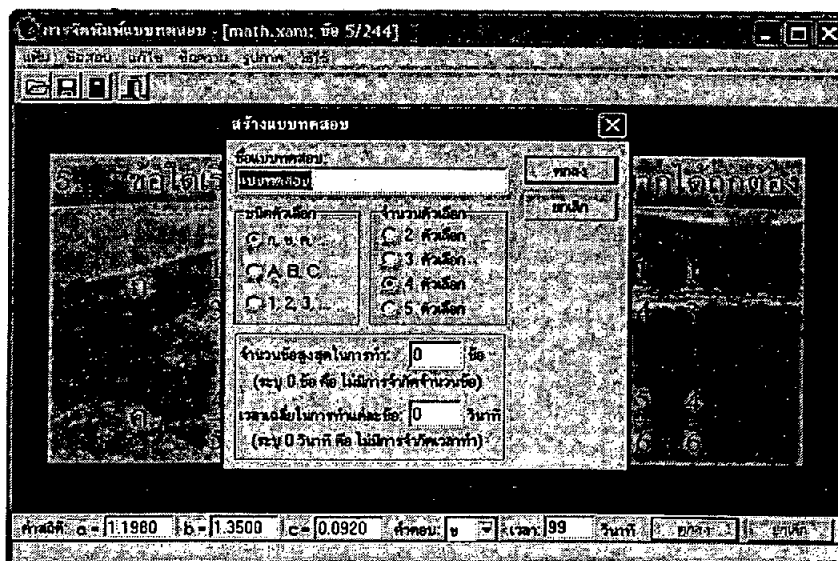
ภาพประกอบ 16 โลโก้หน้าแรกแสดงโปรแกรมจัดพิมพ์แบบทดสอบ

เมื่อคลิกปุ่มปิด หรือกด Enter บนแป้นพิมพ์ โปรแกรมจะแสดงหน้าจอถัดไปให้ป้อนรหัสผ่านเพื่อเข้าสู่โปรแกรมการจัดพิมพ์เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบ และเมื่อผู้ใช้โปรแกรมป้อนรหัสผ่านได้ถูกต้อง ก็จะสามารถเข้าสู่โปรแกรมการจัดพิมพ์เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบได้ ดังภาพประกอบ 17



ภาพประกอบ 17 หน้าจอโปรแกรมคำสั่งเพิ่ม - สร้าง

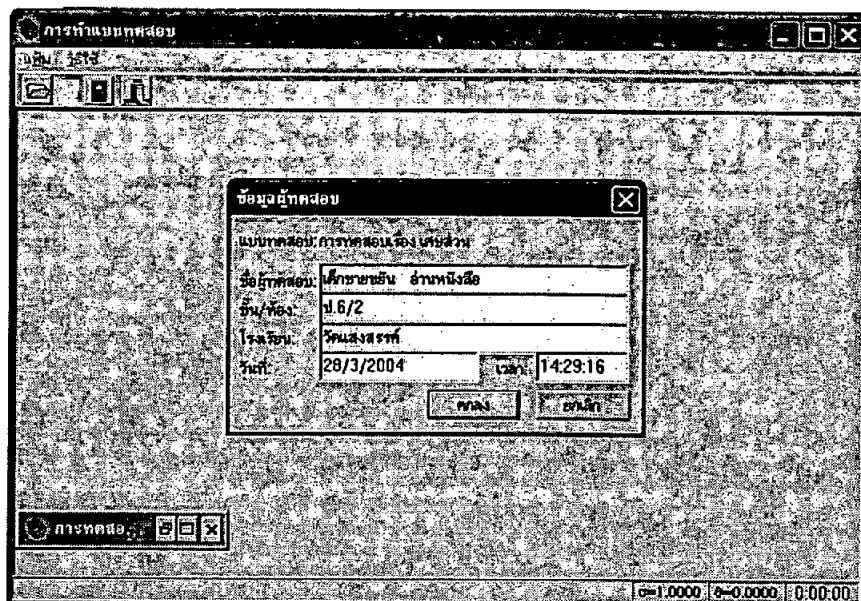
เมื่อเลือกคำสั่งเพิ่ม และคลิกที่คำสั่งสร้าง โปรแกรมจะยอมให้ผู้สร้างแบบทดสอบกำหนด ชื่อแบบทดสอบ ชนิดของตัวเลือก จำนวนตัวเลือก จำนวนข้อสูงสุดของแบบทดสอบ เวลาเฉลี่ยในการตอบข้อสอบแต่ละข้อ ดังภาพประกอบที่ 18



ภาพประกอบ 18 หน้าจอโปรแกรมสร้างแบบทดสอบ

สำหรับคู่มือการใช้โปรแกรมการพิมพ์เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบ โดยคำสั่งอื่นๆ ที่ใช้กับโปรแกรม ผู้วิจัยได้เสนอรายละเอียดไว้ใน ภาคผนวก ข ส่วนแรก

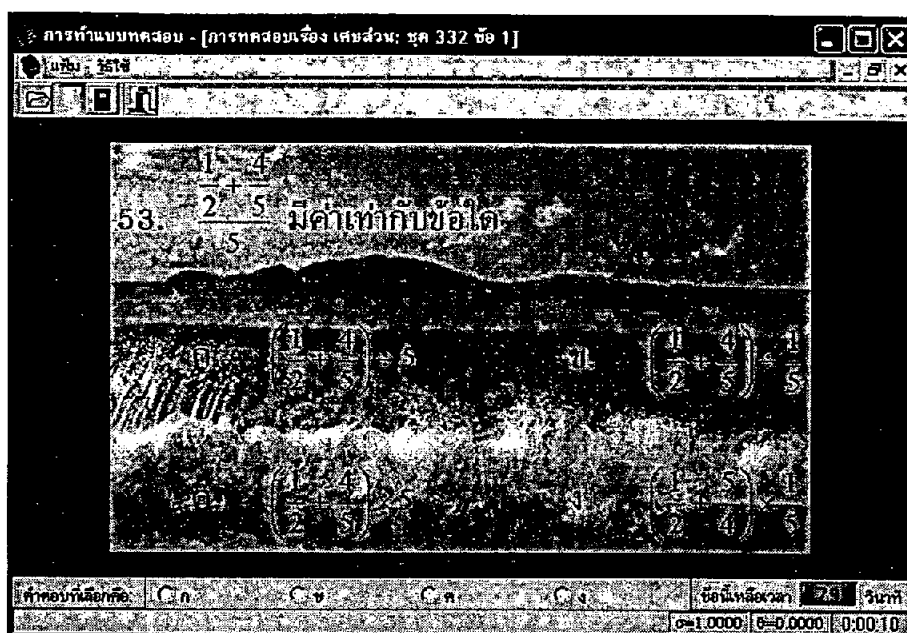
4.2 ภาพตัวอย่างของโปรแกรมสำหรับการทดสอบแบบปรับเหมาะ



ภาพประกอบ 19 หน้าจอโปรแกรมสำหรับการทำแบบทดสอบ

หน้าจอแรกของโปรแกรมสำหรับการทำแบบทดสอบ จะเป็นหน้าจอสำหรับกรอกข้อมูลส่วนตัวของผู้ทดสอบ ซึ่งประกอบด้วย ชื่อผู้สอบ ชั้น/ห้อง โรงเรียน วันที่ และเวลาเริ่มสอบ เมื่อผู้สอบป้อนข้อมูลส่วนตัวครบสมบูรณ์ แล้วคลิกเมาส์ที่ปุ่มตกลง โปรแกรมจะเริ่มให้ทำแบบทดสอบด้วยข้อสอบข้อแรกที่มีความยากง่ายปานกลาง

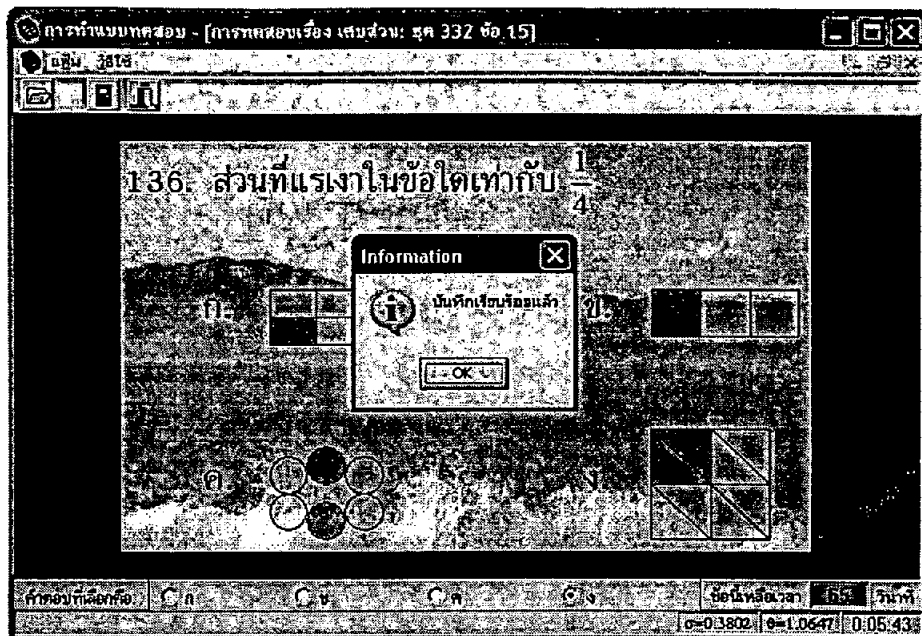
หน้าจอโปรแกรมการทำแบบทดสอบ จะบอกชื่อและชุดรูปแบบของแบบทดสอบ จำนวนข้อสอบที่กำลังทำขณะนั้น ซึ่งขณะที่ปรากฏภาพตัวอย่างของโปรแกรมถัดไปนี้ เป็นการทดสอบวัดความเข้าใจเรื่องเศษส่วน ในชุดที่ 332 หมายความว่าจัดชุดข้อสอบให้ตอบชั้นละ 3 ข้อ กำหนดเวลาให้ตอบตามน้ำหนักความยากของข้อสอบ และเป็นข้อสอบที่มีสื่อประสมประกอบการทดสอบ ดังภาพประกอบ 20



ภาพประกอบ 20 หน้าจอโปรแกรมการทำแบบทดสอบ

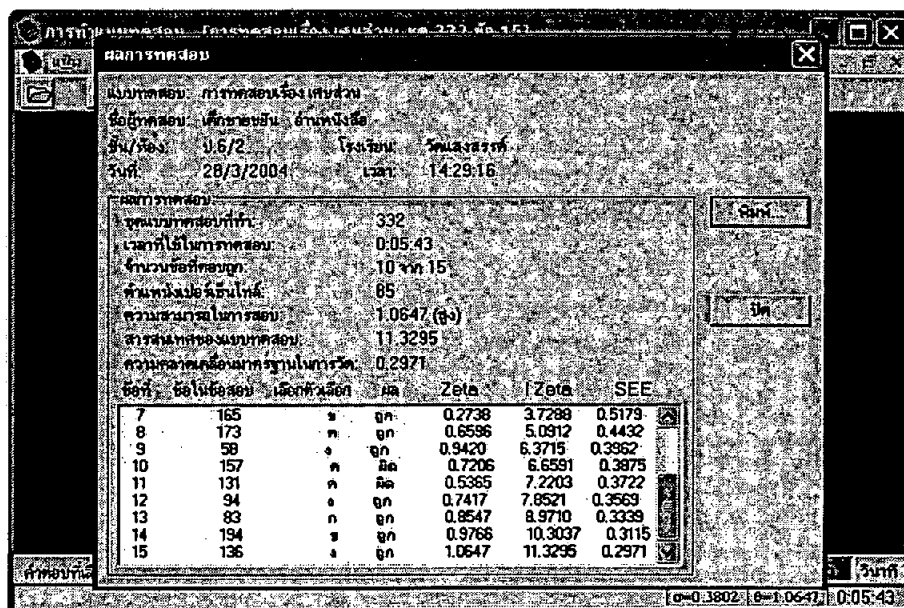
จากภาพประกอบ 20 ข้อสอบที่ปรากฏขณะนี้เป็นข้อสอบข้อที่ 1 ให้เลือกตอบตามข้อ ก ข ค หรือ ง ที่ปรากฏในขอบหน้าต่างล่างของหน้าจอ และแบบทดสอบชุดนี้จะมีช่องบอกเวลาอยู่ขวามือสุด ซึ่งขณะนี้เหลือเวลาสำหรับตอบข้อที่ 1 เท่ากับ 79 วินาที ซึ่งยังมีเวลามากพอในการคิดหา คำตอบช่องบอกเวลานี้จึงแสดงด้วยสีเขียว ถ้าเหลือเวลาสำหรับการทำข้อสอบน้อยกว่า 30 วินาที ช่องบอกเวลานี้จะแสดงด้วยสีเหลือง และถ้าเหลือเวลาสำหรับการทำข้อสอบน้อยกว่า 10 วินาที ช่องบอกเวลานี้จะแสดงด้วยสีแดง ซึ่งเป็นการเตือนบอกเวลาช่วงสุดท้าย การทดสอบจะดำเนินการไปจนกระทั่งผลการทดสอบเป็นไปตามเกณฑ์ยุติการสอบ การทดสอบก็จะสิ้นสุดลง

ดังภาพประกอบ 21



ภาพประกอบ 21 หน้าจอโปรแกรมเมื่อบันทึกผลการสอบเรียบร้อยแล้ว

เมื่อทำการบันทึกผลการทดสอบเสร็จสิ้นลง และใช้เมาส์คลิกปุ่ม OK โปรแกรมจะประมวลผลการทดสอบ และรายงานผลการสอบทันที ดังภาพประกอบ 22



ภาพประกอบ 22 หน้าจอโปรแกรม แสดงผลการทดสอบ

ผลการทดสอบที่ปรากฏ ประกอบด้วย ชื่อแบบทดสอบ ชื่อ-สกุลผู้สอบ ชั้นเรียน โรงเรียน วันที่ และเวลาเริ่มต้นสอบ ส่วนผลการทดสอบประกอบด้วย ชุดของแบบทดสอบที่ทำ เวลาที่ใช้ในการสอบ จำนวนข้อสอบที่ตอบถูกและจำนวนข้อที่ทำทั้งหมด ตำแหน่งเปอร์เซ็นต์ ค่าความสามารถของผู้สอบ ค่าฟังก์ชันสารสนเทศของแบบทดสอบ ค่าความคลาดเคลื่อนในการประมาณ

ค่าความสามารถ ลำดับข้อที่ท่า ลำดับข้อสอบในคลัง ตัวเลือกที่เลือกตอบ ตรวจคำตอบ ค่าความสามารถ ค่าฟังก์ชันสารสนเทศของข้อสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ข้อที่ท่าเสร็จขณะนั้น ๆ

สำหรับคู่มือการใช้โปรแกรมการทดสอบ โดยคำสั่งอื่นๆ ที่ใช้กับโปรแกรมการทดสอบนั้น ผู้วิจัยได้เสนอรายละเอียดไว้ใน ภาคผนวก ข ส่วนท้ายและในทุกโปรแกรมการทดสอบนั้น จะต้องมีการคำนวณค่าต่าง ๆ ได้แก่

4.3 ค่าความน่าจะเป็นในการตอบข้อสอบได้ถูกต้อง

การคำนวณค่าความน่าจะเป็นในการตอบข้อสอบได้ถูกต้อง ผู้วิจัยเลือกวิธีการคำนวณตามโมเดลโลจิสติกแบบ 3 พารามิเตอร์ ซึ่งมีสูตรในการคำนวณดังนี้ (Hambleton and Swaminathan, 1985)

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

เมื่อ $P_i(\theta)$ แทน ความน่าจะเป็นที่ผู้สอบซึ่งมีความสามารถเท่ากับ θ ตอบข้อสอบข้อที่ i ได้ถูกต้อง

θ แทน ค่าความสามารถของผู้สอบ

a_i แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ i

b_i แทน ค่าความยากง่ายของข้อสอบข้อที่ i

c_i แทน ค่าความน่าจะเป็นของผู้สอบที่มีความสามารถต่ำมาก

แต่สามารถตอบข้อสอบข้อที่ i ได้ถูกต้อง ซึ่งก็คือค่าการเดา

D แทน ค่าคงที่มีค่าเท่ากับ 1.7

e แทน ค่าคงที่มีค่าเท่ากับ 2.71828

4.4 ค่าความสามารถของผู้สอบ ด้วยวิธีของเบส์ (Bayesian) โดยเลือกใช้วิธีของโอเวน (Owen) ซึ่งมีสูตรที่ใช้ในการคำนวณดังนี้

ถ้าตอบถูกต้อง

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

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

ถ้าตอบผิด

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

$$\sigma_{m+1}^2 = \sigma_m^2 \left[1 - \left[\frac{\Theta(D)}{1 + \frac{1}{a^2 \sigma_m^2}} \right] \left[\frac{\Theta(D)}{\Phi(D)} - D \right] / [\Phi(D)] \right]$$

- เมื่อ
- $\Theta(D)$ แทน ค่าออร์ดิเนต (ordinate) ของโค้งปกติที่จุด D
 - $\Phi(D)$ แทน พื้นที่ใต้โค้งปกติ จากค่า D ต่ำสุด จนถึงจุด D ที่มีค่า เท่ากับ $\frac{b - \theta_m}{\sqrt{\frac{1}{a^2} + \sigma_m^2}}$ และ $A = c + (1 - c)\Phi(-D)$
 - 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 ค่าความแปรปรวนในการประมาณค่าความสามารถของผู้สอบ เมื่อตอบข้อสอบข้อที่ $m + 1$
 - a แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ $m + 1$
 - b แทน ค่าความยากง่ายของข้อสอบข้อที่ $m + 1$
 - c แทน ค่าการเดาของข้อสอบข้อที่ $m + 1$

4.5 ค่าฟังก์ชันสารสนเทศของข้อสอบและค่าฟังก์ชันสารสนเทศของแบบทดสอบ ซึ่งคำนวณได้จากสูตรดังนี้

$$I_i(\theta) = \frac{[P_i'(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

- เมื่อ $I_i(\theta)$ หมายถึง สารสนเทศของข้อสอบ i ที่ระดับความสามารถ θ
- $P_i'(\theta)$ หมายถึง ค่าอนุพันธ์ของโอกาสในการตอบข้อสอบข้อที่ i ถูก
 - $P_i(\theta)$ หมายถึง โอกาสในการตอบข้อสอบข้อที่ i ถูก ที่ระดับความสามารถ θ
 - $Q_i(\theta)$ หมายถึง โอกาสในการตอบข้อสอบข้อที่ i ผิด ที่ระดับความสามารถ θ

$$I(\theta) = \sum_{i=1}^n \frac{[P_i'(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

- เมื่อ $I(\theta)$ หมายถึง สารสนเทศของแบบทดสอบฉบับ I ที่ระดับความสามารถ θ

4.6 ความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบ คำนวณได้จากสูตร

$$SE(\theta) = \frac{1}{\sqrt{I(\theta)}}$$

เมื่อ $I(\theta)$ หมายถึง สารสนเทศของแบบทดสอบฉบับ I ที่ระดับความสามารถ θ

5. ทดลองใช้โปรแกรม

5.1 ทดลองใช้โปรแกรมเบื้องต้นโดยทดสอบกับนักเรียนโรงเรียนวัดบัวสุวรรณ ประดิษฐ์ อำเภอลาดหลุมแก้ว จังหวัดปทุมธานี โปรแกรมละ 5 คน จำนวน 18 โปรแกรม นักเรียนทั้งสิ้น 90 คน

5.2 ทดลองใช้โปรแกรมโดยผู้เชี่ยวชาญ นำโปรแกรมคอมพิวเตอร์สำหรับการดำเนินการสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ไปให้ผู้เชี่ยวชาญด้านการใช้โปรแกรมการทดสอบจำนวน 3 ท่าน ตรวจสอบประสิทธิภาพของโปรแกรมในด้านการทำงานว่าเป็นไปตามเงื่อนไข กฎเกณฑ์ตามตัวแปรที่ผู้วิจัยกำหนดไว้

6. วิเคราะห์ผลการตรวจสอบประสิทธิภาพของโปรแกรมโดยผู้เชี่ยวชาญ พร้อมทั้งขอคำแนะนำในส่วนที่โปรแกรมบกพร่องเพื่อเป็นข้อมูลในการปรับปรุงแก้ไขโปรแกรม

7. ปรับปรุงแก้ไขโปรแกรมคอมพิวเตอร์ในส่วนที่มีข้อบกพร่องและในส่วนที่โปรแกรมไม่สามารถทำงานเป็นไปตามเงื่อนไขที่กำหนดหรือปรับปรุงโปรแกรมให้มีประสิทธิภาพสูงขึ้น

8. นำโปรแกรมไปตรวจสอบประสิทธิภาพในการใช้งานจริง

8.1 ตรวจสอบประสิทธิภาพการทำงานของโปรแกรมโดยผู้สอน หรือผู้ดำเนินการสอบโดยขอความอนุเคราะห์จากครูผู้สอน และศึกษานิเทศก์ ของสำนักงานเขตพื้นที่การศึกษาปทุมธานีเขตที่ 1 จำนวน 5 ท่าน เป็นผู้ทดลองใช้ และตรวจสอบประสิทธิภาพการทำงานของโปรแกรมด้านความสะดวกรวดเร็ว ความถูกต้องแม่นยำ และความทนทานต่อความผิดพลาดขณะใช้โปรแกรม

8.2 ตรวจสอบประสิทธิภาพด้านความพึงพอใจของผู้สอบโดยการทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่างจำนวน 135 คน เพื่อวิเคราะห์หาประสิทธิภาพของโปรแกรม ด้านความพึงพอใจในการสอบ

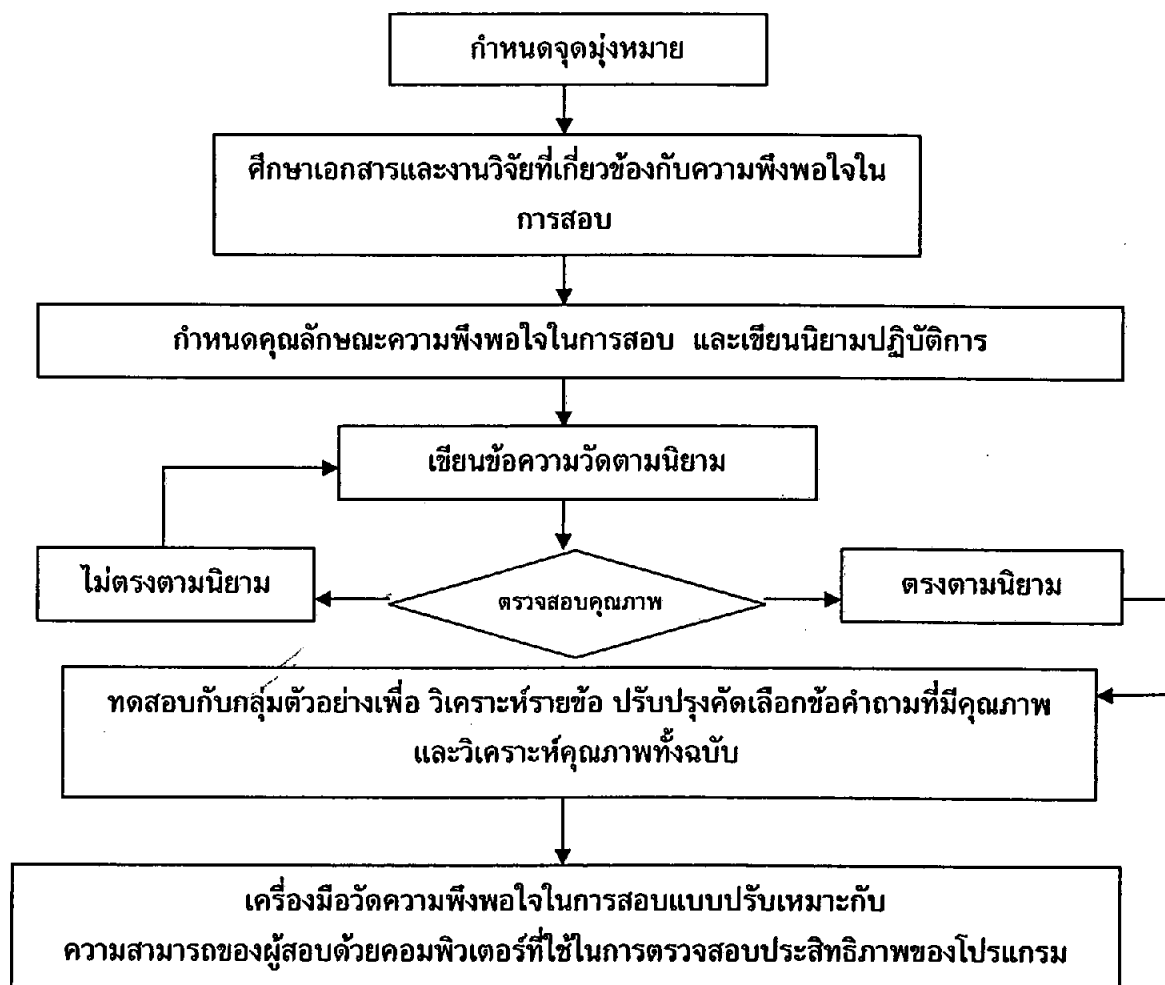
9. ปรับปรุงและแก้ไขโปรแกรมให้สมบูรณ์ยิ่งขึ้น และนำโปรแกรมไปทดสอบกับกลุ่มตัวอย่างเพื่อศึกษาวิจัยผลของตัวแปรอิสระในการวิจัยในระยะที่ 2 ต่อไป

3. การพัฒนาเครื่องมือเพื่อหาประสิทธิภาพของโปรแกรมการทดสอบ

เครื่องมือประเมินประสิทธิภาพของโปรแกรมการทดสอบ ผู้วิจัยใช้แบบสอบวัดความพึงพอใจในการสอบด้วยการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ซึ่งเป็นแบบสอบวัดที่เป็นข้อความวัดคุณลักษณะความพึงพอใจที่เกิดจากการได้รับการตอบสนองความต้องการในปัจจุบันต่าง ๆ ที่ก่อให้เกิดแรงจูงใจหรือกระตุ้นให้ผู้สอบเกิดความผ่อนคลายไม่วิตกกังวลและเต็มใจที่จะทำการสอบ ปัจจัยที่ก่อให้เกิดความพึงพอใจในการสอบนี้ ประกอบด้วย รูปแบบการนำเสนอ

ข้อสอบ เวลาที่ใช้ในการตอบข้อสอบแต่ละข้อ กระบวนการและบรรยากาศในการดำเนินการสอบ และการใช้สื่อประกอบการทดสอบ โดยมีทั้งข้อความทางบวก และข้อความทางลบ และกำหนดน้ำหนักความพึงพอใจในการทดสอบไว้ 5 ระดับ คือพึงพอใจมาก พึงพอใจ เฉยๆ ไม่พึงพอใจ และไม่พึงพอใจอย่างมาก

วิธีการดำเนินการสร้างแบบวัดความพึงพอใจในการสอบ ผู้วิจัยได้ดำเนินการตามลำดับขั้นตอนดังแสดงในภาพประกอบ 23 ดังนี้



ภาพประกอบ 23 แสดงลำดับขั้นตอนการสร้างเครื่องมือวัดความพึงพอใจในการสอบ

การสร้างเครื่องมือวัดความพึงพอใจในการสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ สรุปเป็นขั้นตอนได้ดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างเครื่องมือวัด เพื่อสร้างแบบทดสอบวัดความพึงพอใจในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ที่มีคุณภาพด้านความเที่ยงตรงเชิงโครงสร้าง มีอำนาจจำแนก และมีความเชื่อมั่น

2. ศึกษา नियาม ทฤษฎี เอกสาร เพื่อทำการวิเคราะห์คุณลักษณะความพึงพอใจในการสอบของผู้สอบ
3. กำหนดคุณลักษณะความพึงพอใจในการทดสอบจากการวิเคราะห์ตามข้อที่ 2 และเขียน नियามปฏิบัติการ ตามคุณลักษณะความพึงพอใจในการทดสอบดังกล่าว
4. เขียนข้อความวัดคุณลักษณะความพึงพอใจในการทดสอบตาม नियามเชิงปฏิบัติการ โดยมีทั้งข้อความทางบวก และข้อความทางลบ และกำหนดน้ำหนักความพึงพอใจในการทดสอบไว้ 5 ระดับ คือ พึงพอใจมาก พึงพอใจ เฉยๆ ไม่พึงพอใจ ไม่พึงพอใจอย่างมาก
5. ตรวจสอบคุณภาพของข้อคำถามวัดความพึงพอใจในการทดสอบด้านความเที่ยงตรงเชิงพิณิจ(Face Validity) โดยให้ผู้เชี่ยวชาญจำนวน 5 ท่าน เป็นผู้พิจารณาว่าข้อความนั้น ๆ วัดได้ตรงคุณลักษณะตาม नियามเชิงปฏิบัติการหรือไม่
6. ทดสอบกับกลุ่มตัวอย่าง โดยการนำข้อความวัดที่สร้างขึ้นไปทดสอบกับกลุ่มตัวอย่างพร้อม ๆ กับการทดสอบ เพื่อกำหนดเวลาในการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะ กับนักเรียนชั้นประถมศึกษาปีที่ 6 ในสังกัดสำนักงานเขตพื้นที่การศึกษาปทุมธานี จำนวน 235 คน
7. วิเคราะห์ข้อคำถามรายข้อ (Item Analysis) เพื่อหาค่าอำนาจจำแนกรายข้อโดยการทดสอบค่าที (T-test) และวิเคราะห์คุณภาพทั้งฉบับ
8. นำแบบสอบวัดความพึงพอใจในการทดสอบที่มีคุณภาพไปใช้ในการเก็บรวบรวมข้อมูล เพื่อตรวจสอบประสิทธิภาพของโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ กับนักเรียนชั้นประถมศึกษาปีที่ 6 ในสังกัดสำนักงานเขตพื้นที่การศึกษาปทุมธานี จำนวน 135 คน

**ตัวอย่างแบบสอบวัดความพึงพอใจในการสอบด้วยการทดสอบแบบปรับเหมาะด้วย
คอมพิวเตอร์
คำชี้แจง**

ขอให้ท่านอ่านและพิจารณาข้อความต่อไปนี้ ว่าท่านมีความพึงพอใจต่อการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์มากน้อยเพียงใด โดยทำเครื่องหมาย ✓ ลงในช่องแสดงระดับความพึงพอใจดังนี้

- 5 หมายถึง มีความพึงพอใจมาก
- 4 หมายถึง มีความพึงพอใจ
- 3 หมายถึง มีความพึงพอใจปานกลาง
- 2 หมายถึง ไม่ค่อยพึงพอใจ
- 1 หมายถึง ไม่พึงพอใจอย่างยิ่ง

ข้อความ	ระดับความพึงพอใจ				
	5	4	3	2	1
1. ผู้ดำเนินการสอบชี้แจงวิธีการสอบชัดเจน เข้าใจง่าย	✓				
2. บรรยากาศภายในห้องสอบ					✓
3. ระยะเวลาที่ให้ทำข้อสอบแต่ละข้อ			✓		

จากข้อ 1. ถ้าท่านทำเครื่องหมาย ✓ ลงในช่อง 5 หมายความว่า ท่านมีความพึงพอใจมากในตัวผู้ดำเนินการสอบที่สามารถชี้แจงวิธีการสอบที่ชัดเจน และเข้าใจได้ง่าย

จากข้อ 2. ถ้าท่านทำเครื่องหมาย ✓ ลงในช่อง 1 หมายความว่า ท่านไม่พึงพอใจอย่างยิ่งในบรรยากาศภายในห้องสอบ

จากข้อ 3. ถ้าท่านทำเครื่องหมาย ✓ ลงในช่อง 3 หมายความว่า ท่านมีความพึงพอใจปานกลางในเรื่องของระยะเวลาที่ท่านคิดและตอบข้อสอบในแต่ละข้อ

การเก็บรวบรวมข้อมูล : การพัฒนาเครื่องมือ

การเก็บรวบรวมข้อมูลเพื่อพัฒนาข้อสอบและคลังข้อสอบ พัฒนาโปรแกรมการทดสอบ และศึกษาประสิทธิภาพของโปรแกรมการทดสอบผู้วิจัย ได้ดำเนินการเก็บรวบรวมข้อมูลดังนี้

1. ขอนหนังสือจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ไปขอความอนุเคราะห์เก็บรวบรวมข้อมูลจากสำนักงานเขตพื้นที่การศึกษามุทมาณี เพื่อขอความร่วมมือในการทำวิจัย โดยขอเก็บรวบรวมข้อมูลกับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ในโรงเรียนที่ได้รับการสุ่มเป็นกลุ่มตัวอย่าง
2. นำหนังสือจากสำนักงานเขตพื้นที่การศึกษามุทมาณี ไปประสานงานกับโรงเรียนที่ได้รับการสุ่มเป็นกลุ่มตัวอย่าง เพื่อแจ้งวัตถุประสงค์ของการทำวิจัย พร้อมทั้งขอใช้ห้องคอมพิวเตอร์สำหรับการทดสอบ นัดหมายวัน เวลา เพื่อทำการทดสอบกับนักเรียนกลุ่มตัวอย่าง

3. ทำการทดสอบเพื่อพัฒนาข้อสอบและคลังข้อสอบกับกลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 1 จำนวนกลุ่มตัวอย่าง 597 คน ด้วยแบบทดสอบวัดความเข้าใจเรื่องเศษส่วนที่ผู้วิจัยสร้างเพิ่มเติมขึ้นจำนวน 54 ข้อ เพื่อศึกษาคุณภาพรายข้อด้านค่าอำนาจจำแนก ค่าความยากง่าย ค่าการเดา ด้วยทฤษฎีการตอบสนองข้อสอบ โดยใช้โมเดลโลจิสติกแบบ 3 พารามิเตอร์ในการวิเคราะห์ทำการวิเคราะห์คุณภาพของแบบทดสอบด้านความเป็นมิติเดียวโดยการวิเคราะห์องค์ประกอบ

4. ทำการทดสอบเพื่อพัฒนาโปรแกรมกับกลุ่มตัวอย่างนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวนกลุ่มตัวอย่าง 235 คน ด้วยโปรแกรมการทดสอบแบบปรับเหมาะ จำนวน 8 โปรแกรมย่อย เพื่อใช้เป็น ข้อมูลไปกำหนดเวลาที่ใช้ในการสอบแต่ละข้อ ซึ่งวิธีการจัดนักเรียนเข้าห้องสอบนั้นได้จัดนักเรียนแบบเรียงสลับเป็นระบบกล่าวคือ คนที่ 1 – 8 จะได้รับการทดสอบกับโปรแกรมคอมพิวเตอร์ข้อสอบชุดที่ 1 – 8 ตามลำดับ นักเรียนคนที่ 9 – 16 จะได้รับการทดสอบกับโปรแกรมคอมพิวเตอร์ข้อสอบชุดที่ 1 – 8 ตามลำดับเช่นเดียวกัน จัดนักเรียนเข้าสอบเช่นเดียวกันนี้กับทุกห้องจนครบทั้ง 5 ห้อง และทดสอบด้วยแบบสอบวัดความพึงพอใจในการสอบ เพื่อหาคุณภาพของแบบสอบวัด

5. ทำการทดสอบเพื่อหาประสิทธิภาพของโปรแกรมกับกลุ่มตัวอย่างนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวนกลุ่มตัวอย่าง 135 คน ด้วยโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ที่ทำการกำหนดคุณลักษณะของข้อสอบตามตัวแปรที่ผู้วิจัยศึกษาเรียบร้อยแล้วและเป็นโปรแกรมการทดสอบที่ผ่านการพิจารณาประสิทธิภาพจากผู้เชี่ยวชาญด้านโปรแกรมการทดสอบแล้วจำนวน 18 โปรแกรมย่อย ซึ่งวิธีการจัดนักเรียนเข้าห้องสอบนั้นกระทำเช่นเดียวกับการจัดนักเรียนเข้าสอบในข้อที่ 4 และทำการทดสอบด้วยแบบสอบวัดความพึงพอใจในการสอบกับผู้เข้าสอบทุกคน หลังจากการทดสอบด้วยโปรแกรมการทดสอบเสร็จสิ้นลง เพื่อหาประสิทธิภาพของโปรแกรมการทดสอบด้านความพึงพอใจในการสอบ

6. การบันทึกผลการสอบ ทำการบันทึกผลด้วยกระดาษคำตอบสำหรับการทดสอบเพื่อพัฒนาข้อสอบและคลังข้อสอบกับกลุ่มตัวอย่างกลุ่มที่ 1 ส่วนการทดสอบเพื่อพัฒนาโปรแกรมตามข้อที่ 4 กับกลุ่มตัวอย่างกลุ่มที่ 2 บันทึกผลการสอบด้วยอุปกรณ์บันทึกข้อมูล (Handy Drive) จากเครื่องคอมพิวเตอร์ และทำการบันทึกผลด้วยกระดาษคำตอบสำหรับกับกลุ่มตัวอย่างที่ทำการทดสอบเพื่อวัดความพึงพอใจในการสอบ

การวิเคราะห์ข้อมูลและสถิติที่ใช้ในการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลจากการดำเนินการวิจัยระยะที่ 1 เพื่อหาคุณภาพของข้อสอบรายข้อที่ผู้วิจัยสร้างขึ้น การกำหนดเวลาที่ใช้ในการตอบข้อสอบแต่ละข้อ การหาคุณภาพของแบบสอบวัดความพึงพอใจในการสอบ และหาประสิทธิภาพของโปรแกรมการทดสอบ ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลดังนี้

1. วิเคราะห์ข้อมูลคุณภาพข้อสอบที่ผู้วิจัยสร้างขึ้น

1.1 วิเคราะห์คุณภาพของข้อสอบเป็นรายข้อ เพื่อหาค่าพารามิเตอร์ของข้อสอบ คือ ค่าอำนาจจำแนก ค่าความยากง่าย และค่าการเดาตามทฤษฎีการตอบสนองข้อสอบ โดยใช้โมเดลโลจิสติกแบบ 3 พารามิเตอร์ การหาค่าพารามิเตอร์ของข้อสอบดังกล่าวนี้วิเคราะห์ด้วยโปรแกรม Bilog

1.2 ตรวจสอบความเป็นมิติเดียวของข้อสอบ ด้วยการวิเคราะห์องค์ประกอบ (factor analysis) และพิจารณาค่าไอเกนของตัวประกอบที่ได้จากการวิเคราะห์ ถ้าหากตัวประกอบที่ 1 มีค่าไอเกนสูงกว่าค่าไอเกนของตัวแปรประกอบที่ 2 ตั้งแต่ 3 เท่าขึ้นไปและค่าไอเกนของตัวประกอบ 2 มีค่าใกล้เคียงกับค่าไอเกนของตัวประกอบที่เหลือ หรือสามารถอธิบายความแปรปรวนร่วมได้ร้อยละ 70 ขึ้นไปแสดงว่าข้อสอบดังกล่าวมีความเป็นมิติเดียว

1.3 คัดเลือกข้อสอบที่มีคุณภาพตามเกณฑ์ ซึ่งเป็นข้อสอบที่มีค่าอำนาจจำแนก ตั้งแต่ 0.8 ขึ้นไป ค่าความยากง่าย อยู่ในช่วง -3.00 ถึง 3.00 ค่าการเดาน้อยกว่าหรือเท่ากับ 0.3 ได้ข้อสอบใหม่จำนวน 44 ข้อ นำเข้าไปเก็บรวมไว้ในคลังข้อสอบของรังสรรค์ มณีเล็ก ซึ่งมีข้อสอบอยู่ในคลังแล้วจำนวน 200 ข้อ รวมเป็นข้อสอบในคลังทั้งสิ้น 244 ข้อ

2. การวิเคราะห์ข้อมูลเพื่อกำหนดเวลาที่ใช้ในการตอบข้อสอบแต่ละข้อ โดยการหาค่าเฉลี่ยของข้อสอบทั้งฉบับ และหาค่าเฉลี่ยของข้อสอบแต่ละชุดที่แบ่งตามช่วงค่าความยาก โดยใช้สูตร

$$\bar{X} = \frac{\sum X}{nN}$$

เมื่อ \bar{X} คือ ค่าเฉลี่ยของเวลาที่ใช้ในการตอบแต่ละข้อ

$\sum X$ คือ ผลรวมของเวลาที่ใช้ในการสอบแต่ละข้อ

n คือ จำนวนข้อสอบทั้งฉบับหรือจำนวนข้อสอบในแต่ละชุด

N คือ จำนวนผู้เข้าสอบทั้งหมดหรือผู้เข้าสอบในแต่ละชุด

3. การวิเคราะห์ข้อมูลเพื่อหาคุณภาพของแบบสอบวัดความพึงพอใจในการสอบ

3.1 หาค่าเฉลี่ยของคะแนนสอบ (\bar{X}) และค่าความเบี่ยงเบนมาตรฐาน (SD)

3.2 การวิเคราะห์ข้อสอบเป็นรายข้อ (Item Analysis) เพื่อหาค่าอำนาจจำแนกโดยใช้

t-test

3.3 ความเที่ยงตรง (Validity) ความเที่ยงตรงเชิงโครงสร้าง (Construct Validity) หาโดยวิธีวิเคราะห์ความเที่ยงตรงเชิงโครงสร้างแบบ Item-Total Correlation โดยซึ่งใช้สูตรสัมประสิทธิ์สหสัมพันธ์อย่างง่าย

3.4 ค่าความเชื่อมั่น (Reliability) คำนวณโดยใช้สูตรสัมประสิทธิ์อัลฟา (α – Coefficient)

4. วิเคราะห์ข้อมูลเพื่อหาประสิทธิภาพของโปรแกรมการทดสอบ

4.1 พิจารณาความเห็นของผู้เชี่ยวชาญและบัณฑิตก่าแนะนำเพื่อทำการปรับปรุงแก้ไขในส่วนที่บกพร่องของโปรแกรม

4.2 พิจารณาความเห็นของผู้ใช้โปรแกรมได้แก่ ครูผู้สอน หรือผู้ดำเนินการสอบและบัณฑิตก่าแนะนำเพื่อทำการปรับปรุงแก้ไขในส่วนที่บกพร่องของโปรแกรม

4.3 หาค่าเฉลี่ย และทดสอบความแตกต่างค่าเฉลี่ยของคะแนนความพึงพอใจในการสอบของนักเรียนกลุ่มตัวอย่าง

การดำเนินการวิจัยระยะที่ 2 เพื่อศึกษาผลของตัวแปรอิสระ ที่มีต่อตัวแปรตาม

การวิจัยในระยะที่ 2 เป็นการศึกษามูลของตัวแปรอิสระ 3 ตัวคือ จำนวนข้อสอบที่นำเสนอให้ตอบในแต่ละชั้น เวลาที่ใช้ในการตอบแต่ละข้อ และรูปแบบของข้อสอบที่ไม่มีและมีสื่อประสม ที่มีต่อเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ซึ่งดำเนินการดังนี้

เครื่องมือที่ใช้ในการวิจัย

1. แบบทดสอบ

เครื่องมือที่ใช้สำหรับเก็บรวบรวมข้อมูล ได้แก่ แบบทดสอบวิชาคณิตศาสตร์วัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 ของรังสรรค์ มณีเล็ก (2540) และข้อสอบวัดในเรื่องเดียวกันที่ ผู้วิจัยสร้างเพิ่มเติมขึ้น ซึ่งมีลักษณะเป็นแบบทดสอบวัดความเข้าใจเรื่อง เศษส่วน รูปแบบของข้อสอบเป็นแบบปรนัยชนิดเลือกตอบ 4 ตัวเลือก ให้คะแนนแบบ 0, 1 จำนวนทั้งสิ้น 244 ข้อ ค่าความยากง่ายของข้อสอบอยู่ในช่วง -3.0 ถึง 3.0 ค่าอำนาจจำแนกเริ่มที่ 0.4 ขึ้นไป ค่าการเดาต่ำกว่าหรือเท่ากับ 0.3 ข้อสอบทั้งหมดจะถูกเก็บไว้ในในคลังข้อสอบ ซึ่งใช้ระบบคอมพิวเตอร์ช่วยในการจัดเก็บข้อสอบ และช่วยจัดชุดแบบทดสอบเสนอให้แก่ผู้เข้าสอบแต่ละคน โดยผู้เข้าสอบจะต้องทำการสอบบนหน้าจอคอมพิวเตอร์ ส่วนการตรวจให้คะแนนและรายงานผลการสอบจะกระทำและรายงานผลบนหน้าจอคอมพิวเตอร์ทันทีหลังจากยุติการสอบ

2. โปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะ

โปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบที่ใช้ในการค้นคว้าวิจัยในครั้งนี้ เป็นโปรแกรมการสอบที่ผู้วิจัยได้พัฒนาขึ้น ซึ่งมีจำนวนทั้งสิ้น 18 โปรแกรม ซึ่งได้จากการจัดหมู่ตัวแปรอิสระ 3 ตัว คือการจัดชุดข้อสอบให้ตอบในแต่ละชั้น 3 รูปแบบ ได้แก่ การจัดชุดข้อสอบให้ตอบชั้นละ 1 ข้อ ชั้นละ 2 ข้อ และชั้นละ 3 ข้อ การกำหนดเวลาที่ใช้ในการตอบแต่ละข้อ 3 ลักษณะ ได้แก่ ไม่กำหนดเวลาที่ใช้ในการตอบแต่ละข้อ กำหนดเวลาที่ใช้ในการตอบแต่ละข้อเฉลี่ยเท่ากันทุกข้อ และกำหนดเวลาที่ใช้ในการตอบแต่ละข้อตามน้ำหนักค่าความยากของข้อสอบ และรูปแบบของ ข้อสอบ 2 รูปแบบ ได้แก่ รูปแบบของข้อสอบที่ไม่มีสื่อประสมประกอบการทดสอบ และรูปแบบของข้อสอบที่มีสื่อประสมประกอบการทดสอบ ซึ่งโปรแกรมย่อยทั้ง 18 โปรแกรมนี้มีเงื่อนไขแสดงได้ดังตาราง 9 และกระบวนการพัฒนาโปรแกรม ผู้วิจัยได้ดำเนินการพัฒนาโปรแกรมไว้ในการดำเนินการวิจัยระยะที่ 1

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลทำการวิจัยในระยะที่ 2 เป็นการเก็บรวบรวมข้อมูลเพื่อศึกษาผลของตัวแปรอิสระที่มีต่อตัวแปรตาม ซึ่งผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลดังนี้

1. ขอนหนังสือจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ไปขอความอนุเคราะห์ให้เก็บรวบรวมข้อมูลจากสำนักงานเขตพื้นที่การศึกษามุทมาศ เพื่อขอความร่วมมือในการทำวิจัย โดยขอเก็บรวบรวมข้อมูลกับนักเรียนชั้นประถมศึกษาปีที่ 6 ในโรงเรียนที่ได้รับการสุ่มเป็นกลุ่มตัวอย่าง

2. นำหนังสือจากสำนักงานเขตพื้นที่การศึกษามุทมาศ ไปประสานงานกับโรงเรียนที่ได้รับการสุ่มเป็นกลุ่มตัวอย่าง เพื่อแจ้งวัตถุประสงค์ของการทำวิจัย พร้อมทั้งขอใช้ห้องคอมพิวเตอร์สำหรับการทดสอบ นัดหมายวัน เวลา เพื่อทำการทดสอบกับนักเรียนกลุ่มตัวอย่าง

3. เตรียมโปรแกรมการทดสอบแบบปรับเหมาะที่จะนำไปทดสอบกับนักเรียนกลุ่มตัวอย่างทั้ง 18 โปรแกรม โดยลงโปรแกรมไว้ในเครื่องคอมพิวเตอร์ล่วงหน้าและใส่รหัสเพื่อป้องกันการเข้าสู่โปรแกรมได้ก่อนการดำเนินการสอบ

4. ทดสอบกับนักเรียนชั้นประถมศึกษาปีที่ 6 ที่เป็นกลุ่มตัวอย่างด้วยโปรแกรมการทดสอบแบบปรับเหมาะ เพื่อใช้เป็นข้อมูลในการศึกษาผลของตัวแปรอิสระที่มีต่อตัวแปรตาม ซึ่งวิธีการจัด นักเรียนเข้าห้องสอบนั้นได้จัดนักเรียนแบบเรียงสลับเป็นระบบกล่าวคือ นักเรียนคนที่ 1 – 18 จะได้รับการทดสอบกับโปรแกรมคอมพิวเตอร์ข้อสอบชุดที่ 1 – 18 ตามลำดับ นักเรียนคนที่ 19 – 36 จะได้รับการทดสอบกับโปรแกรมคอมพิวเตอร์ข้อสอบชุดที่ 1 – 18 ตามลำดับเช่นเดียวกัน ทำการจัดนักเรียนเข้าสอบเช่นเดียวกันนี้กับทุกห้องจนครบทั้ง 13 ห้อง

5. ทำการบันทึกผลการสอบของนักเรียนที่เป็นกลุ่มตัวอย่างทุกคนด้วยอุปกรณ์บันทึกข้อมูล (Handy Drive) จากเครื่องคอมพิวเตอร์ที่ใช้ทำการทดสอบ เพื่อนำข้อมูลผลการทดสอบของนักเรียน กลุ่มตัวอย่างไปวิเคราะห์ศึกษาผลของตัวแปรอิสระที่มีต่อตัวแปรตามต่อไป

การวิเคราะห์ข้อมูลและสถิติที่ใช้ในการวิเคราะห์ข้อมูล

ผู้วิจัยทำการวิเคราะห์ข้อมูลในการวิจัยระยะที่ 2 โดยโปรแกรมบอร์แลนเดลฟาย (BORLAND DELPHI) ในการคำนวณค่าฟังก์ชันสารสนเทศของข้อสอบและแบบทดสอบ ซึ่งมีขั้นตอนการวิเคราะห์ ดังนี้

1. คำนวณค่าฟังก์ชันสารสนเทศของข้อสอบทุกข้อที่ระดับความสามารถของผู้สอบ โดยใช้สูตรดังนี้ (Hambleton and Swaminathan. 1985 : 105 - 107)

$$I_i(\theta) = \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

เมื่อ $I_i(\theta)$ คือ ค่าฟังก์ชันสารสนเทศของข้อสอบข้อที่ i

$$P'_i(\theta) = \frac{Da_iQ_i(\theta)[P_i(\theta) - c_i]}{(1 - c_i)}$$

$$\text{และ } Q_i(\theta) = 1 - P_i(\theta)$$

2. การวิเคราะห์ข้อมูลเพื่อเปรียบเทียบเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ

2.1 คำนวณค่าฟังก์ชันสารสนเทศของแบบทดสอบทั้ง 18 ฉบับ โดยใช้สูตรดังนี้ (Hambleton and Swaminathan. 1985 : 104)

$$I(\theta) = \sum_{i=1}^n \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

เมื่อ $I(\theta)$ คือ ค่าฟังก์ชันสารสนเทศของแบบทดสอบ

2.2 การวิเคราะห์เพื่อเปรียบเทียบเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ

2.2.1 หาค่าเฉลี่ยของเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ จากสูตร

$$\bar{X} = \frac{\sum X}{N}$$

เมื่อ \bar{X} แทน ค่าเฉลี่ยเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ

$\sum X$ แทน ผลรวมเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ ของแต่ละคน

N แทน จำนวนผู้สอบ

2.2.2 วิเคราะห์ความแปรปรวน เพื่อเปรียบเทียบค่าเฉลี่ยของเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ จากสูตร (รังสรรค์ มณีเล็ก. อ้างอิงจาก Glass and Stanley. 1970 : 353)

$$F = \frac{MS_B}{MS_W}$$

- เมื่อ F แทน ค่าสถิติทดสอบเอฟ สำหรับทดสอบความแตกต่างของค่าเฉลี่ย
 MS_B แทน ค่ากำลังสองเฉลี่ยของข้อมูลระหว่างกลุ่ม
 (mean square between groups)
 MS_W แทน ค่ากำลังสองเฉลี่ยของข้อมูลภายในกลุ่ม
 (mean square within groups)

2.2.3 เปรียบเทียบความแตกต่างของค่าเฉลี่ยเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ เป็นรายคู่โดยใช้วิธีเอสของเซฟเฟ (Scheffe's S Method) จากสูตร (อุทุมพร ทองอุไทย. 2523 : 158)

$$S = \sqrt{(k-1)F_{\alpha, v_1, v_2} MS_e \left[\sum_{j=1}^k \frac{(c_j)^2}{n_j} \right]}$$

- เมื่อ F แทน ค่า F ที่ชั้นความเป็นอิสระ v_1 และ v_2 ที่ระดับนัยสำคัญ α
 k แทน จำนวนระดับการทดลอง
 c_j แทน สัมประสิทธิ์การเปรียบเทียบ
 n_j แทน จำนวนคะแนนในระดับการทดสอบ j
 MS_e แทน ค่ากำลังสองเฉลี่ยของข้อมูลภายในกลุ่ม

บทที่ 4

ผลการวิเคราะห์ข้อมูล

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

a	แทน	ค่าอำนาจจำแนกของข้อสอบ
b	แทน	ค่าความยากง่ายของข้อสอบ
c	แทน	ค่าการเดาของข้อสอบ
Sa	แทน	ค่าความเบี่ยงเบนมาตรฐานของค่าอำนาจจำแนกของข้อสอบ
Sb	แทน	ค่าความเบี่ยงเบนมาตรฐานของค่าความยากง่ายของข้อสอบ
Sc	แทน	ค่าความเบี่ยงเบนมาตรฐานของค่าการเดาของข้อสอบ
r_{tt}	แทน	ค่าความเชื่อมั่นของแบบทดสอบสอบ
\bar{X}	แทน	ค่าเฉลี่ยของคะแนนความพึงพอใจในการสอบ เวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ
SD	แทน	ค่าความเบี่ยงเบนมาตรฐานของคะแนนความพึงพอใจในการสอบ เวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ
SS	แทน	ผลบวกกำลังสองของเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ
MS	แทน	ความแปรปรวนของเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ
df	แทน	ช่วงชั้นความเป็นอิสระ
t	แทน	ค่าวิกฤตของ t
F	แทน	ค่าวิกฤตของ F
N	แทน	จำนวนผู้สอบ
n	แทน	จำนวนข้อสอบ

การนำเสนอผลการวิเคราะห์ข้อมูล

ในการนำเสนอผลการวิเคราะห์ข้อมูลจากการวิจัยครั้งนี้ เพื่อให้ผลการวิเคราะห์ข้อมูลเป็นไปอย่างมีระบบตามขั้นตอนและสอดคล้องกับวัตถุประสงค์ของการวิจัย ผู้วิจัยจึงขอเสนอผลการวิเคราะห์ข้อมูลแบ่งออกเป็น 2 ตอน ตอนแรกเป็นผลการวิเคราะห์ข้อมูลเพื่อพัฒนาเครื่องมือและตรวจสอบคุณภาพของเครื่องมือ ส่วนตอนที่สองเป็นผลการทดสอบเวลาที่ใช้ในการสอบ จำนวนข้อสอบของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ จำแนกตามตัวแปรอิสระที่ศึกษา 3 ตัวแปร ซึ่งผู้วิจัยขอเสนอผลการวิเคราะห์โดยเรียงลำดับ ตามหัวข้อต่อไปนี้

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลในขั้นตอนเตรียมการ

1. ผลการวิเคราะห์ข้อมูลเพื่อการพัฒนาข้อสอบและคลังข้อสอบ
 - ค่าพารามิเตอร์ของข้อสอบวัดความเข้าใจเรื่องเศษส่วนที่ผู้วิจัยสร้างเพิ่มเติมขึ้นและค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบ
2. ผลการวิเคราะห์ข้อมูลเพื่อพัฒนาและประเมินคุณภาพโปรแกรมการทดสอบ
 - ผลการวิเคราะห์ข้อมูลเพื่อกำหนดเวลาที่ใช้ในการตอบแต่ละข้อ
 - ผลการวิเคราะห์ข้อมูลเพื่อประเมินคุณภาพของโปรแกรมด้านความพึงพอใจในการสอบ

ตอนที่ 2 ผลการวิเคราะห์ข้อมูลในขั้นตอนหลัก

1. ผลการวิเคราะห์ค่าสถิติพื้นฐาน
2. ผลการวิเคราะห์ข้อมูลเพื่อทดสอบสมมติฐานการวิจัย

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลในขั้นตอนเตรียมการ

1. ค่าพารามิเตอร์ของข้อสอบที่ผู้วิจัยสร้างเพิ่มเติมขึ้น

ข้อสอบที่ผู้วิจัยสร้างเพิ่มเติมขึ้น เป็นข้อสอบวัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 สร้างขึ้นตามฟ้าเซ็ทการสร้างข้อสอบของรังสรรค์ มณีเล็ก จำนวน 54 ข้อ นำไปทดสอบกับกลุ่มตัวอย่าง 597 คน ซึ่งจากผลการทดสอบ ได้ค่าอำนาจจำแนก ค่าความยากง่าย ค่าการเดาของข้อสอบ และค่าความเบี่ยงเบนมาตรฐาน ดังตาราง 9

ตาราง 9 ค่าอำนาจจำแนก ค่าความยากง่าย ค่าการเดา และความเบี่ยงเบนมาตรฐานของข้อสอบวัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 จำนวน 54 ข้อ ปรากฏผล ดังนี้

ข้อที่	เก็บไว้ในคลังข้อที่	ค่าอำนาจจำแนก		ค่าความยากง่าย		ค่าการเดา	
		a	Sa	b	Sb	C	Sc
1	201	0.9030	0.1780	-1.9320	0.4260	0.2390	0.097
2	ตัดทิ้ง	0.8540	0.2880	1.5410	0.5890	0.3570	0.0430
3	202	0.6330	0.1160	0.0350	0.1280	0.2180	0.0810
4	203	1.2120	0.3150	1.2110	0.3980	0.1880	0.0330
5	204	0.8500	0.2490	0.8620	0.3260	0.3000	0.0590
6	205	0.9240	0.1410	-0.2410	0.1360	0.2470	0.0820
7	206	0.7590	0.1420	0.2860	0.1370	0.1860	0.0820
8	207	0.9640	0.1870	0.3640	0.1550	0.1940	0.0590
9	208	1.0560	0.1400	-0.1960	0.1080	0.1630	0.6400
10	209	1.3300	0.2540	0.2850	0.1500	0.1770	0.0490

ตาราง 9 (ต่อ)

ข้อที่	เก็บไว้ในคลัง ข้อที่	ค่าอำนาจจำแนก		ค่าความยากง่าย		ค่าการเดา	
		a	Sa	b	Sb	C	Sc
11	210	0.9600	0.1990	0.2280	0.1450	0.2000	0.0670
12	211	1.3860	0.2790	0.2380	0.1540	0.1830	0.0510
13	ตัดทิ้ง	1.2250	0.3870	0.8520	0.3780	0.4000	0.440
14	212	1.6190	0.2270	-0.3580	0.1200	0.1570	0.0610
15	213	0.4510	0.0900	-1.8010	0.3900	0.2610	0.1010
16	214	1.5900	0.3940	0.6020	0.2670	0.2690	0.0390
17	ตัดทิ้ง	0.8190	0.3080	3.1120	1.2460	0.1000	0.0190
18	215	1.0210	0.1980	0.1710	0.1450	0.2190	0.067
19	216	1.2080	0.3070	0.1550	0.1800	0.3000	0.0700
20	217	0.7580	0.1390	0.1650	0.123	0.1820	0.0700
21	218	1.5960	0.4160	0.4140	0.2330	0.2840	0.045
22	219	1.4670	0.3210	1.0050	0.3060	0.1360	0.0270
23	ตัดทิ้ง	-	-	-	-	-	-
24	ตัดทิ้ง	0.8630	0.3550	3.013	1.3930	0.2760	0.0240
25	ตัดทิ้ง	0.3270	0.0710	0.2940	0.1740	0.3350	0.0970
26	220	1.0760	0.2400	0.6730	0.2190	0.1780	0.047
27	221	0.9270	0.1120	-0.0490	0.0850	0.1110	0.0480
28	222	0.8980	0.1210	-0.6360	0.1410	0.2050	0.0830
29	223	0.7340	0.1550	0.7740	0.2150	0.1800	0.057
30	224	0.9600	0.2600	0.8750	0.3020	0.2320	0.0510
31	225	1.1120	0.2480	0.4860	0.1980	0.2430	0.0550
32	226	0.8110	0.1500	-0.3250	0.1460	0.2490	0.0910
33	227	0.5180	0.1160	0.0560	0.1550	0.2960	0.0950
34	228	0.9640	0.2540	0.3920	0.2090	0.3010	0.0700
35	229	1.0760	0.1930	-0.2700	0.1400	0.2400	0.0830
36	230	1.2430	0.2200	0.1120	0.1380	0.2130	0.0600
37	231	0.5324	0.4220	1.1630	0.4950	0.3000	0.0360
38	232	0.8160	0.1250	-0.4970	0.1440	0.2290	0.0880

ตาราง 9 (ต่อ)

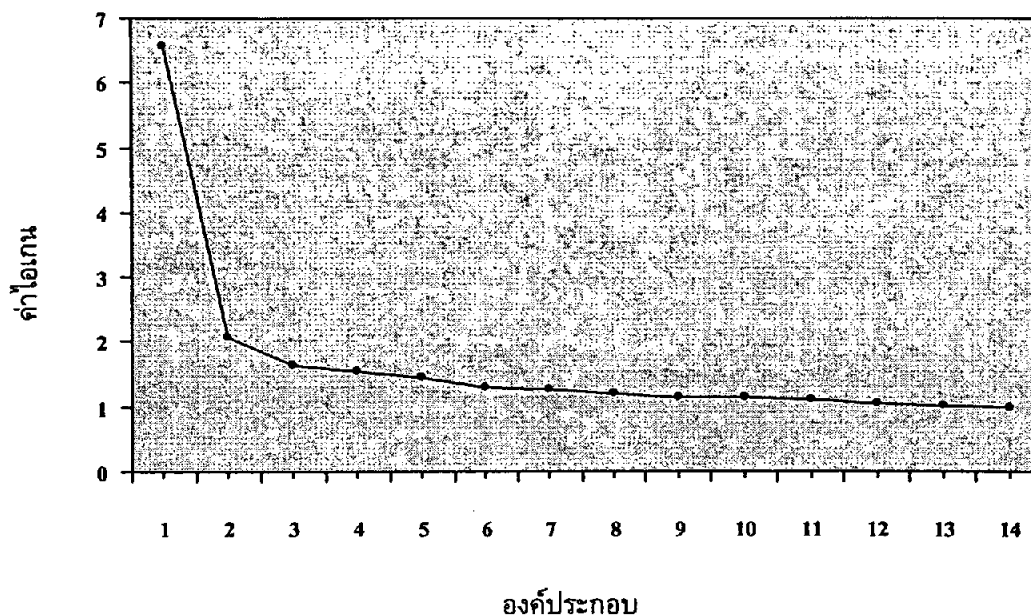
ข้อที่	เก็บไว้ในคลัง ข้อที่	ค่าอำนาจจำแนก		ค่าความยากง่าย		ค่าการเดา	
		a	Sa	b	Sb	C	Sc
39	233	1.3680	0.2100	-0.4470	0.1420	0.2240	0.0790
40	234	1.4060	0.2620	0.7500	0.2160	0.1400	0.030
41	235	1.0510	0.3820	2.2600	0.9620	0.2390	0.0260
42	ตัดทิ้ง	1.2910	0.3590	0.6630	0.2940	0.3450	0.0470
43	ตัดทิ้ง	-	-	-	-	-	-
44	236	1.3160	0.3130	1.2890	0.3950	0.1510	0.0270
45	237	1.2700	0.2250	0.405	0.1560	0.1760	0.0440
46	ตัดทิ้ง	0.5880	0.1630	0.9280	0.3230	0.3640	0.0730
47	238	1.2440	0.5220	2.1980	1.1520	0.2800	0.0250
48	239	0.9700	0.2970	1.5150	0.5320	0.2310	0.0370
49	240	1.3660	0.3150	1.3620	0.4120	0.1350	0.0250
50	ตัดทิ้ง	0.1840	0.0530	5.7020	1.6270	0.1560	0.0440
51	241	1.0260	0.3680	2.2880	0.9590	0.2480	0.0260
52	242	1.3570	0.4520	1.7060	0.7380	0.2160	0.0260
53	243	1.1640	0.3500	1.3890	0.5180	0.2570	0.0350
54	244	0.9710	0.3080	1.7110	0.6180	0.2160	0.0350
รวมเฉลี่ย	44 ข้อ	1.0350	0.3160	0.7080	1.2170	0.2310	0.0690

จากตาราง 9 พบว่าแบบทดสอบวัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 จำนวน 54 ข้อ มีค่าพารามิเตอร์ของข้อสอบ ดังนี้คือ ค่าอำนาจจำแนก มีค่าอยู่ระหว่าง 0.4510 ถึง 1.6190 ค่าความยากง่ายมีค่าอยู่ระหว่าง -1.9320 ถึง 2.2880 และค่าการเดามีค่าอยู่ระหว่าง 0.1110 ถึง 0.3000 และเมื่อพิจารณาค่าพารามิเตอร์ของข้อสอบที่มีคุณภาพตามเกณฑ์ คือค่าอำนาจจำแนกตั้งแต่ 0.80 ขึ้นไป ค่าความยากง่ายอยู่ในระหว่าง -3.0 ถึง 3.0 และค่าการเดาน้อยกว่าหรือเท่ากับ 0.3 แล้วพบว่าคงเหลือข้อสอบที่มีค่าพารามิเตอร์เป็นไปตามเกณฑ์ดังกล่าว จำนวน 44 ข้อ เมื่อพิจารณาค่าพารามิเตอร์ของข้อสอบทั้งหมด พบว่า ค่าเฉลี่ยของอำนาจจำแนกเท่ากับ 1.0350 ค่าเฉลี่ยของความยากง่ายเท่ากับ 0.7080 และค่าเฉลี่ยของการเดาเท่ากับ 0.2310 นำข้อสอบทั้ง 44 ข้อ ไปวิเคราะห์องค์ประกอบ เพื่อตรวจสอบความเป็นมิติเดียวของข้อสอบ หาความเที่ยงตรงเชิงโครงสร้าง และหาค่าความเชื่อมั่นของแบบทดสอบ ผลของการวิเคราะห์ดังกล่าวแสดงได้ด้วยตาราง และภาพประกอบดังต่อไปนี้

ตาราง 10 ค่าไอเกน (eigen value) ที่ได้จากการวิเคราะห์องค์ประกอบของแบบทดสอบวัดความเข้าใจเรื่องเศษส่วน จำนวน 14 องค์ประกอบ ที่มีค่าไอเกน มากกว่า 1

องค์ประกอบ	ค่าไอเกน
1	6.578
2	2.063
3	1.650
4	1.550
5	1.454
6	1.296
7	1.258
8	1.193
9	1.154
10	1.142
11	1.115
12	1.047
13	1.014
14	1.006

เมื่อนำค่าไอเกนของแบบทดสอบวัดความเข้าใจเรื่องเศษส่วนไปเขียนกราฟจะได้กราฟตามภาพประกอบ 24 ดังนี้



ภาพประกอบ 24 ค่าไอเกน (eigen value) ที่ได้จากการวิเคราะห์องค์ประกอบของแบบทดสอบวัดความเข้าใจเรื่องเศษส่วน จำนวน 14 องค์ประกอบที่มีค่าไอเกนมากกว่า 1

จากตาราง 10 และภาพประกอบ 24 พบว่าค่าไอเกนของตัวประกอบที่ 1 สูงกว่าค่าไอเกนของตัวประกอบอื่น ๆ ที่เหลือมากกว่า 3 เท่า และตัวประกอบที่เหลือมีค่าไอเกนใกล้เคียงกัน ซึ่งลักษณะดังกล่าวนี้พออนุมานได้ว่า ข้อสอบที่ได้มีความเป็นมิติเดียว ตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

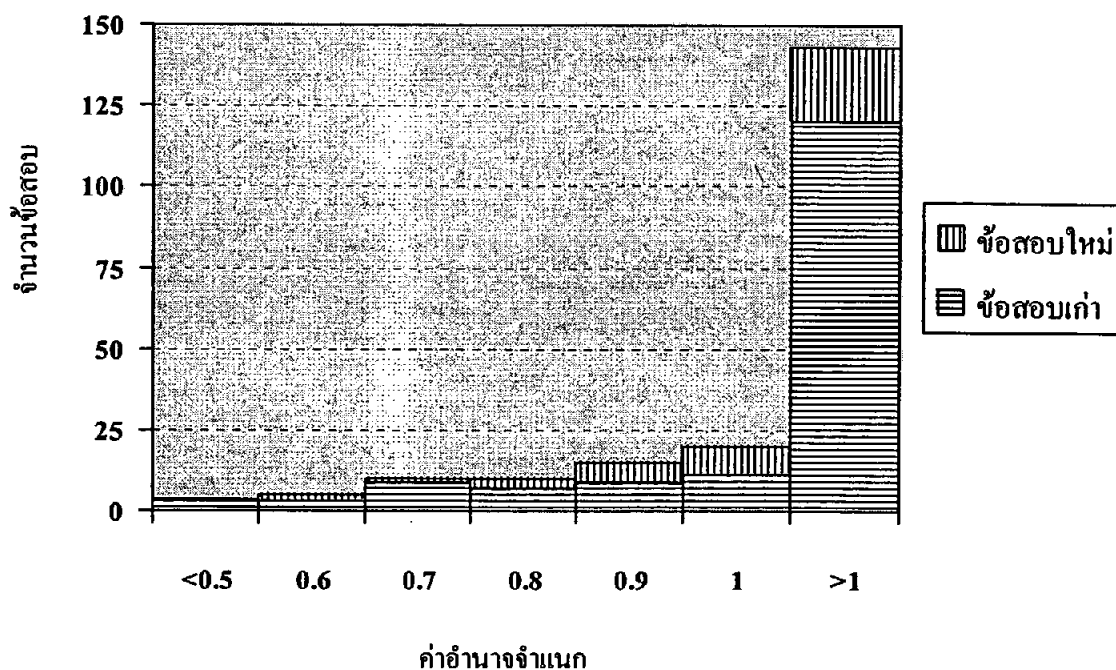
เมื่อนำผลของการตอบข้อสอบแต่ละข้อไปหาค่าสัมประสิทธิ์สหสัมพันธ์รายข้อกับผลการตอบของแบบทดสอบทั้งฉบับ เพื่อตรวจสอบความเที่ยงตรงเชิงโครงสร้าง (Construct Validity) ของแบบทดสอบ ผลการวิเคราะห์ดังกล่าวแสดงได้ดัง ตาราง 11 และค่าความเชื่อมั่นของแบบทดสอบปรากฏผล ต่อท้ายตาราง ดังนี้

ตาราง 11 ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนรายข้อ จำนวน 44 ข้อ กับคะแนนของแบบทดสอบวัด
ความเข้าใจเรื่องเศษส่วนทั้งฉบับ

ข้อสอบ ข้อที่	ค่าสหสัมพันธ์รายข้อ	ข้อสอบ ข้อที่	ค่าสหสัมพันธ์รายข้อ
1	.2109	29	.2807
3	.2765	30	.2790
4	.2575	31	.3769
5	.2053	32	.3518
6	.3792	33	.2086
7	.2899	34	.3321
8	.3519	35	.3979
9	.4382	36	.4343
10	.4623	37	.2182
11	.3796	38	.3210
12	.4538	39	.4626
14	.5469	40	.4082
15	.2167	41	.2047
16	.3752	44	.2820
18	.3797	45	.4314
19	.3802	47	.2308
20	.3313	48	.2853
21	.3923	49	.2603
22	.3620	51	.2434
26	.3802	52	.2449
27	.4241	53	.2176
28	.3583	54	.2807
$r_{tt} = .8505$			

จากตาราง 11 พบว่าค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนรายข้อของแบบทดสอบ ทั้ง 44 ข้อ กับคะแนนของแบบทดสอบทั้งฉบับ มีค่าสัมประสิทธิ์สหสัมพันธ์ทางบวกทุกข้ออยู่ระหว่าง 0.2047 ถึง 0.5464 และค่าความเชื่อมั่นของแบบทดสอบเท่ากับ 0.8505 จากลักษณะของความสัมพันธ์ดังกล่าวนี้พออนุมานได้ว่า ข้อสอบทุกข้อวัดโครงสร้างเดียวกันและมีความเชื่อมั่นในการวัดค่อนข้างสูง

เมื่อผู้วิจัยนำข้อสอบที่ผู้วิจัยสร้างเพิ่มเติมขึ้น และมีคุณภาพตามเกณฑ์ จำนวน 44 ข้อ เข้าไปรวมกับข้อสอบที่อยู่ในคลังเดิมของรังสรรค์ มณีเล็ก จำนวน 200 ข้อ จึงรวมเป็นข้อสอบในคลังทั้งสิ้น 244 ข้อ เพื่อให้เห็นจำนวนข้อของข้อสอบในคลังข้อสอบในแต่ละช่วงของค่าอำนาจจำแนก ค่าความยากง่าย และค่าการเดาของข้อสอบที่อยู่ในคลังเดิม และข้อสอบที่นำเข้าไปเก็บไว้ในคลังใหม่ ผู้วิจัยจึงได้นำเสนอตามภาพประกอบ 25 - 27

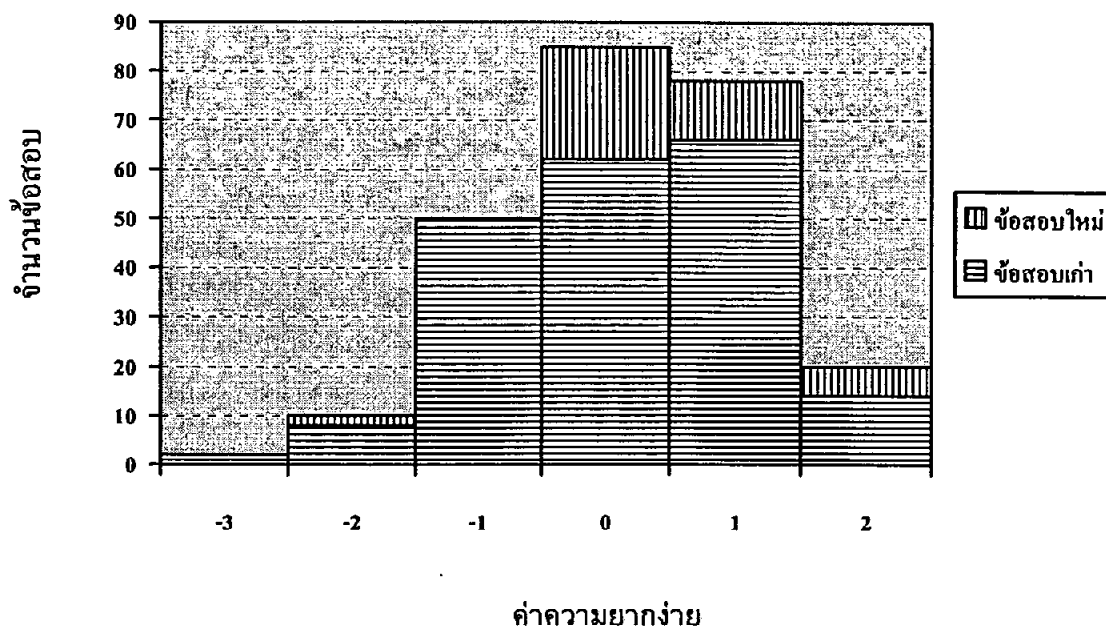


ภาพประกอบ 25 จำนวนข้อสอบวัดความเข้าใจ เรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 ในคลังข้อสอบเดิม 200 ข้อ และข้อสอบใหม่ 44 ข้อ จำแนกตามช่วงของค่าอำนาจจำแนก

จากภาพประกอบ 24 จะเห็นว่าข้อสอบส่วนใหญ่ในคลังข้อสอบมีอำนาจจำแนกมากกว่า 1.00 ประมาณ 143 ข้อ เป็นข้อสอบเก่า 120 ข้อ ข้อสอบใหม่ 23 ข้อ ถ้าพิจารณาโดยใช้เกณฑ์ค่าอำนาจจำแนกไม่ต่ำกว่า .80 แล้ว ข้อสอบที่มีค่าอำนาจจำแนก ตั้งแต่ .80 ขึ้นไปมีประมาณ 191 ข้อ คิดเป็นร้อยละ 78.28 ของข้อสอบในคลัง ซึ่งเป็นข้อสอบเก่า 150 ข้อ ข้อสอบใหม่ 41 ข้อ

ส่วนค่าอำนาจจำแนกของแบบทดสอบเก่าเท่ากับ 1.00 ค่าอำนาจจำแนกของแบบทดสอบใหม่เท่ากับ 1.035 และค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบทั้งสองฉบับเท่ากับ 1.006

เมื่อจำแนกข้อสอบในคลังข้อสอบออกตามค่าความยากง่ายแล้ว สามารถจำแนกได้
ดังภาพประกอบ 26

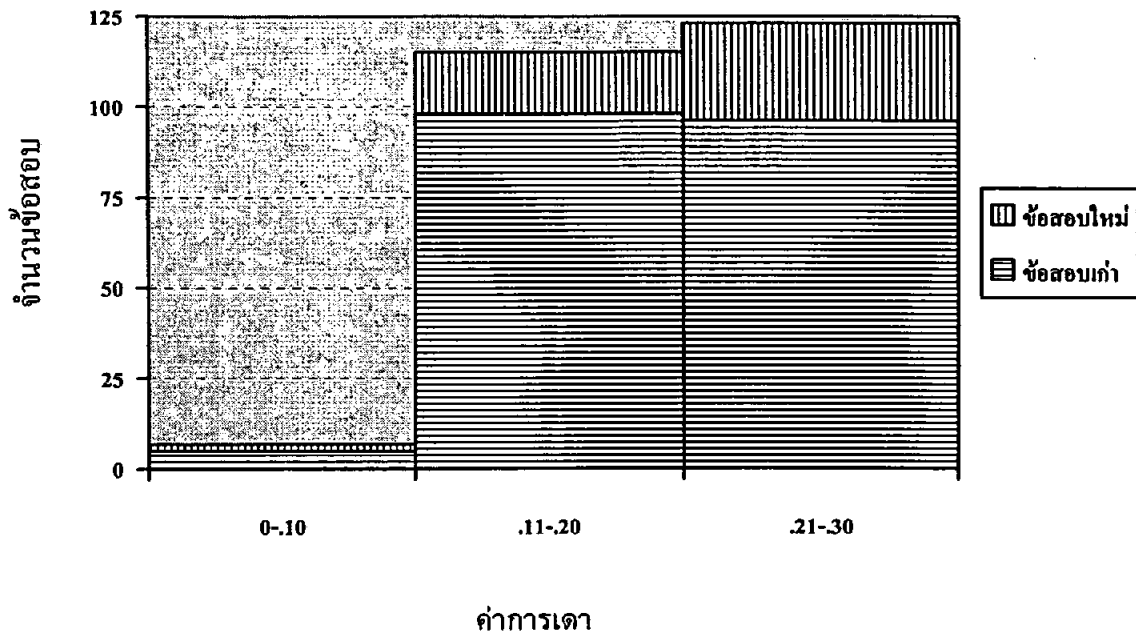


ภาพประกอบ 26 จำนวนข้อสอบวัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 ในคลังข้อสอบ
จำนวน 244 ข้อ จำแนกตามช่วงของค่าความยากง่าย

จากภาพประกอบ 26 จะเห็นว่าข้อสอบส่วนใหญ่ในคลังข้อสอบมีค่าความยากง่ายอยู่ในช่วง
-1.5 ถึง 1.5 มีข้อสอบประมาณ 193 ข้อ เป็นข้อสอบเก่า 160 ข้อ ข้อสอบใหม่ 35 ข้อ คิดเป็น
ร้อยละ 79.10 ของข้อสอบในคลัง จากลักษณะของการแจกแจงดังกล่าวมีลักษณะเบ้ไปทางลบ
เล็กน้อย แสดงว่าข้อสอบในคลังส่วนใหญ่เป็นข้อสอบที่ค่อนข้างยาก

ส่วนค่าความยากง่ายของแบบทดสอบเก่าเท่ากับ 0.80 ค่าความยากง่ายของแบบทดสอบ
ใหม่เท่ากับ 0.7080 และค่าความยากง่ายเฉลี่ยของแบบทดสอบทั้งสองฉบับเท่ากับ 0.783

เมื่อจำแนกข้อสอบในคลังข้อสอบออกตามค่าการเดาแล้ว สามารถจำแนกได้ ดัง
ภาพประกอบ 27



ภาพประกอบ 27 จำนวนข้อสอบวัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 ในคลังข้อสอบ 244 ข้อ จำแนกตามช่วงของค่าการเดา

จากภาพประกอบ 27 พบว่าข้อสอบส่วนใหญ่ในคลังข้อสอบมีค่าการเดาอยู่ในช่วง .11 ถึง .30 ประมาณ 239 ข้อ เป็นข้อสอบเก่า 195 ข้อ ข้อสอบใหม่ 44 ข้อ ซึ่งเป็นข้อสอบที่มีค่าการเดาอยู่ในช่วง .21 - .30 จำนวน 122 ข้อ และอยู่ในช่วง .11-.20 จำนวน 115 ข้อ

ส่วนค่าการเดาของแบบทดสอบเก่าเท่ากับ 0.20 ค่าการเดาของแบบทดสอบใหม่เท่ากับ 0.2310 และค่าการเดาเฉลี่ยของแบบทดสอบทั้งสองฉบับเท่ากับ 0.205

คลังข้อสอบที่ผู้วิจัยใช้ในการศึกษาวิจัยในครั้งนี้ เป็นข้อสอบวิชาคณิตศาสตร์ วัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 จำนวนรวมทั้งสิ้น 244 ข้อ มีค่าอำนาจจำแนกเฉลี่ย 1.006 ค่าความยากง่ายเฉลี่ย 0.783 และค่าการเดาเฉลี่ย 0.205 การสร้างข้อสอบเพิ่มเติมของผู้วิจัยในครั้งนี้ ผู้วิจัยได้สร้างข้อสอบขึ้นตามฟาเซ็ทการสร้างข้อสอบของรังสรรค์ มณีเล็ก โดยคำนึงถึงโครงสร้างลักษณะของข้อคำถาม และคำตอบ จึงได้ข้อสอบครอบคลุมและครบถ้วนตามโครงสร้างของฟาเซ็ท และทำการตรวจสอบความเที่ยงตรงตามโครงสร้างโดยผู้เชี่ยวชาญด้านการสอนคณิตศาสตร์ และผู้เชี่ยวชาญด้านการวัดผลการศึกษา ซึ่งผู้เชี่ยวชาญเห็นว่าข้อสอบที่ผู้วิจัยสร้างขึ้น สอดคล้องกับโครงสร้างหรือฟาเซ็ทดังกล่าว หลังจากที่ได้ศึกษาคุณภาพรายข้อและคุณภาพทั้งฉบับของแบบทดสอบที่ผู้วิจัยสร้างเพิ่มเติมขึ้นแล้ว ผู้วิจัยจึงนำข้อสอบเข้าไปรวมกับข้อสอบที่อยู่ในคลังเดิมของรังสรรค์ มณีเล็ก ซึ่งไม่สามารถที่จะทำการเทียบคะแนนของแบบทดสอบทั้งสองฉบับได้ เนื่องจากถ้า นำข้อสอบเดิมจำนวน 200 ข้อไปทำการทดสอบพร้อมกับแบบทดสอบฉบับใหม่จำนวน 54 ข้อ จะต้องใช้กลุ่มตัวอย่างและเวลาอีกมาก แต่ผู้วิจัยได้ใช้เวลาไปกับการพัฒนาโปรแกรมการทดสอบมาก

แล้วจึงมีเวลาจำกัดที่จะทำการทดลองข้อสอบใหม่เพื่อทำการเทียบคะแนนของแบบทดสอบทั้งสองฉบับได้ ดังนั้น ผู้วิจัยจึงพิจารณาความเป็นมิติเดียวของข้อสอบจากผลการตรวจสอบของผู้เชี่ยวชาญเพื่ออนุมานความเป็นมิติเดียวกันของข้อสอบที่ผู้วิจัยสร้างเพิ่มเติมขึ้นกับข้อสอบที่อยู่ในคลังเดิมดังกล่าว

2. การวิเคราะห์ข้อมูลเพื่อพัฒนาและการประเมินคุณภาพของโปรแกรมการทดสอบด้านความพึงพอใจในการสอบ

การวิเคราะห์ข้อมูลเพื่อพัฒนาโปรแกรมการทดสอบนั้น ผู้วิจัยยังขาดข้อมูลในการกำหนดเวลาให้กับข้อสอบแต่ละข้อ ผู้วิจัยจึงได้ทำการศึกษาข้อมูลเพื่อกำหนดเวลาที่ใช้ในการตอบข้อสอบแต่ละข้อดังกล่าว และเมื่อผู้วิจัยได้พัฒนาโปรแกรมการทดสอบเสร็จสิ้นแล้ว จึงมีความจำเป็นที่จะต้องประเมินคุณภาพของโปรแกรม ผู้วิจัยจึงได้พัฒนาเครื่องมือที่ใช้ในการประเมินโปรแกรมขึ้นและทำการประเมินคุณภาพของโปรแกรมการทดสอบ ผลปรากฏตามตาราง ดังนี้

2.1 ผลการกำหนดเวลาที่ใช้ในการตอบแต่ละข้อ

ตาราง 12 จำนวนข้อสอบ จำนวนผู้เข้าสอบ ค่าเฉลี่ย และความเบี่ยงเบนมาตรฐานของเวลาที่ใช้ในการทดสอบ จากการทดสอบจากกลุ่มตัวอย่าง เพื่อกำหนดเวลาที่ใช้ในการตอบแต่ละข้อของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

ชุดที่	ความยากของข้อสอบ	จำนวนข้อสอบ	จำนวนผู้สอบ	เวลาที่ใช้สอบ (นาที)	
				\bar{x}	SD
1	-1.50 ถึง -1.01	7	29	5.53	2.6359
2	-1.00 ถึง -0.05	20	31	19.45	6.1556
3	-0.05 ถึง -0.01	20	29	19.56	7.7111
4	0.00 ถึง 0.50	20	31	24.36	5.7470
5	0.51 ถึง 1.00	20	28	25.12	6.5315
6	1.01 ถึง 1.50	20	29	25.38	7.9364
7	1.51 ถึง 2.00	20	28	26.37	7.7685
8	2.01 ถึง 2.50	18	30	25.27	3.6177
รวม	-1.50 ถึง 2.50	135	235	21.12	8.8301

จากตาราง 12 พบว่าเวลาที่ใช้ในการสอบด้วยแบบทดสอบวัดความเข้าใจเรื่องเศษส่วนชั้นประถมศึกษาปีที่ 6 ทั้ง 8 ฉบับ ผู้สอบใช้เวลาในการทำแบบทดสอบแต่ละฉบับโดยเฉลี่ยแตกต่างกันไปตามความยากง่ายของข้อสอบ ข้อสอบชุดที่ 1 เป็นชุดของข้อสอบง่าย 7 ข้อ ใช้เวลาโดยเฉลี่ย 5.53 นาที (SD = 2.6359) ชุดที่ 2 ข้อสอบยากกว่าชุดแรกมีข้อสอบ 20 ข้อ ใช้เวลาเฉลี่ย 19.45 นาที (SD

= 6.1556) ชุดที่ 3 ใช้เวลาเฉลี่ย 19.56 นาที (SD = 7.7111) ชุดที่ 4 ใช้เวลาเฉลี่ย 24.36 นาที (SD = 5.7470) ชุดที่ 5 ใช้เวลาเฉลี่ย 25.12 นาที (SD = 6.5315) ชุดที่ 6 ใช้เวลาเฉลี่ย 25.38 นาที (SD = 7.9364) ชุดที่ 7 ใช้เวลาเฉลี่ย 26.37 นาที (SD = 7.7685) และชุดที่ 8 เป็นชุดของข้อสอบที่ยากที่สุดจำนวน 18 ข้อ ใช้เวลาเฉลี่ย 25.27 นาที (SD = 3.6177) และโดยรวมเฉลี่ยทั้ง 8 ฉบับ ใช้เวลา 21.12 นาที (SD = 8.8301)

เวลาที่ได้จากการทดลองสอบกับกลุ่มตัวอย่างครั้งนี้ เป็นเวลาที่ใช้สอบด้วยแบบทดสอบทั้งฉบับของแต่ละชุด แต่ผู้วิจัยต้องการกำหนดเวลาให้กับข้อสอบเป็นรายข้อมีหน่วยเป็นวินาที ดังนั้นผู้วิจัยจึงจะต้องหารเวลาที่ใช้ทั้งฉบับด้วยจำนวนข้อในฉบับนั้นและคูณด้วย 60 เพื่อให้เป็นวินาที แต่เนื่องจากการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ มีข้อตกลงว่า จะต้องไม่เป็นการสอบวัดความเร็วในการตอบ (Speed Test) หรือผู้สอบจะต้องมีเวลามากพอในการตอบข้อสอบแต่ละข้อ เพื่อให้เป็นไปตามข้อตกลงดังกล่าว ผู้วิจัยจึงเพิ่มเวลาให้กับข้อสอบมากขึ้น โดยนำค่าเฉลี่ยของเวลาการสอบในแต่ละฉบับบวกกับค่าเบี่ยงเบนมาตรฐานของฉบับนั้นก่อน แล้วจึงหาค่าเฉลี่ยเป็นรายข้อ และทำให้หน่วยเป็นวินาที โดยการคูณด้วย 60 ผลจากกระบวนการกำหนดเวลาให้กับข้อสอบในคลังข้อสอบครั้งนี้ ได้ผลแสดงดังตาราง 14

ตาราง 13 ชุดของข้อสอบ ความยากง่ายของข้อสอบ และเวลาที่ใช้ในการตอบข้อสอบแต่ละข้อ โดยการกำหนดเวลาให้กับข้อสอบในคลังแบ่งเป็น 12 ชุด ดังนี้

ชุดข้อสอบ	ความยากง่ายของข้อสอบ	จำนวนข้อสอบในคลัง	เวลาที่ใช้ตอบในแต่ละข้อ (วินาที)
1	-3.00 ถึง -2.51	0	0
2	-2.50 ถึง -2.01	1	60
3	-2.00 ถึง -1.51	5	65
4	-1.50 ถึง -1.01	7	70
5	-1.00 ถึง -0.51	21	77
6	-0.50 ถึง -0.01	37	82
7	0.00 ถึง 0.50	45	90
8	0.51 ถึง 1.00	40	95
9	1.01 ถึง 1.50	42	99
10	1.51 ถึง 2.00	24	103
11	2.01 ถึง 2.50	18	107
12	2.51 ถึง 3.00	4	112
รวม	-3.00 ถึง 3.00	244	90

จากตาราง 13 พบว่า ความยากง่ายของข้อสอบในคลังอยู่ในช่วง -2.50 ถึง 3.00 จำนวนข้อสอบทั้งสิ้น 244 ข้อ แบ่งข้อสอบออกเป็น 12 ชุด ตามช่วงความยากง่าย ข้อสอบชุดที่ 1 ไม่มีข้อสอบในช่วงนี้ ชุดที่ 2 ชุดที่ 3 และชุดที่ 12 ไม่ได้นำไปทดลองสอบกับกลุ่มตัวอย่าง จึงกำหนดเวลาโดยการคำนวณจากแนวโน้มของการลดลง เมื่อข้อสอบง่ายและแนวโน้มของการเพิ่มขึ้นเมื่อข้อสอบยากขึ้น จากกระบวนการกำหนดเวลาให้กับข้อสอบ จากตารางดังกล่าวจึงเริ่มกำหนดเวลาให้กับผู้สอบจากชุดที่ 4 ซึ่งมีจำนวนข้อสอบ 7 ข้อ กำหนดเวลาให้ตอบแต่ละข้อในชุดนี้ข้อละ 70 วินาที ชุดที่ 5 มีข้อสอบจำนวน 21 ข้อ กำหนดเวลาให้ตอบข้อละ 77 วินาที ชุดที่ 6 มีข้อสอบจำนวน 37 ข้อ กำหนดเวลาให้ตอบข้อละ 82 วินาที ชุดที่ 7 มีข้อสอบจำนวน 45 ข้อ กำหนดเวลาให้ตอบข้อละ 90 วินาที ชุดที่ 8 มีข้อสอบจำนวน 40 ข้อ กำหนดเวลาให้ตอบข้อละ 95 วินาที ชุดที่ 9 มีข้อสอบจำนวน 42 ข้อ กำหนดเวลาให้ตอบข้อละ 82 วินาที ชุดที่ 10 มีข้อสอบจำนวน 24 ข้อ กำหนดเวลาให้ตอบข้อละ 103 วินาที ชุดที่ 11 มีข้อสอบจำนวน 18 ข้อ กำหนดเวลาให้ตอบข้อละ 107 วินาที ส่วนชุดที่ 2 ชุดที่ 3 และชุดที่ 12 กำหนดเวลาให้ตอบข้อละ 60 วินาที 65 วินาที และ 112 วินาที ตามลำดับ และ การกำหนดเวลาให้ตอบข้อสอบในแต่ละข้อโดยเฉลี่ยเท่ากันทุกข้อ โดยหาจากค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานทั้งฉบับแล้วเฉลี่ยเป็นรายข้อ จำนวนข้อสอบในคลังทั้งสิ้น 244 ข้อ กำหนดเวลาให้ตอบแต่ละข้อเท่ากันทุกข้อ คือ 90 วินาที

2.2 ผลการวิเคราะห์ข้อมูลคุณภาพของแบบสอบวัดความพึงพอใจในการสอบ

คุณภาพของแบบสอบวัดความพึงพอใจในการสอบของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ผู้วิจัยทำการหาคุณภาพของแบบสอบวัดความพึงพอใจในการสอบ โดยการวิเคราะห์คุณภาพรายข้อ เพื่อหาค่าอำนาจจำแนก โดยใช้ t-test หาคความเที่ยงตรงเชิงโครงสร้างโดยการวิเคราะห์แบบ Item-Total Correlation และหาความเชื่อมั่นของแบบทดสอบวัด โดยใช้สูตรสัมประสิทธิ์อัลฟา (α -coefficient) ผลการวิเคราะห์ แสดงดังตาราง 14 - 16

ตาราง 14 ผลการวิเคราะห์ ค่าอำนาจจำแนกรายข้อของแบบสอบวัดความพอใจในการสอบด้วยการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

ข้อสอบ	กลุ่ม	N	\bar{x}	SD	t
1	ต่ำ	67	3.8060	.80225	6.358**
	สูง	68	4.5735	.58120	
2	ต่ำ	67	3.9552	.87789	5.534**
	สูง	68	4.6324	.48575	
3	ต่ำ	67	3.4627	.98977	5.153**
	สูง	68	4.2353	.73541	
4	ต่ำ	67	3.6119	.99932	5.062**
	สูง	68	4.3824	.75369	
5	ต่ำ	67	3.8657	.75670	5.260**
	สูง	68	4.5441	.74180	
6	ต่ำ	67	4.0746	.72433	4.010**
	สูง	68	4.5588	.67762	
7	ต่ำ	67	3.7463	.94297	5.319**
	สูง	68	4.5147	.72261	
8	ต่ำ	67	3.6866	1.04749	4.790**
	สูง	68	4.4118	.67438	
9	ต่ำ	67	3.2090	1.03773	7.755**
	สูง	68	4.3971	.71529	
10	ต่ำ	67	3.4627	.94273	6.792**
	สูง	68	4.3971	.62628	
11	ต่ำ	67	3.5821	.90703	5.837**
	สูง	68	4.3824	.66981	
12	ต่ำ	67	3.6119	.98404	4.265**
	สูง	68	4.2647	.78451	
13	ต่ำ	67	3.0597	.85068	9.779**
	สูง	68	4.3235	.63343	
14	ต่ำ	67	2.5821	1.08919	3.813**
	สูง	68	3.3529	1.25502	
15	ต่ำ	67	3.1045	.88992	7.474**
	สูง	68	4.1765	.77153	

ตาราง 14 (ต่อ)

ข้อสอบ	กลุ่ม	N	\bar{x}	SD	t
16	ต่ำ	67	3.5522	1.00429	4.604**
	สูง	68	4.2353	.69363	
17	ต่ำ	67	3.4478	.90927	7.380**
	สูง	68	4.4412	.63204	
18	ต่ำ	67	3.4478	1.11849	3.880**
	สูง	68	4.1618	1.01644	
19	ต่ำ	67	4.0597	.85068	5.999**
	สูง	68	4.7647	.46094	
20	ต่ำ	67	4.0000	1.04447	4.075**
	สูง	68	4.6029	.62628	
21	ต่ำ	67	3.3731	.98220	8.245**
	สูง	68	4.5441	.63325	
22	ต่ำ	67	3.2687	.78975	9.719**
	สูง	68	4.4853	.65774	
23	ต่ำ	67	3.3433	1.02325	8.590**
	สูง	68	4.5588	.55671	
24	ต่ำ	67	3.3284	.85967	9.344**
	สูง	68	4.5441	.63325	
25	ต่ำ	67	3.3731	1.04208	6.630**
	สูง	68	4.4118	.75775	

**มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 14 พบว่า ค่าอำนาจจำแนกรายข้อของแบบสอบวัดความพึงพอใจในการสอบ โดยใช้ค่า t-test ในการจำแนกกลุ่มผู้สอบนั้น พบว่าค่า t-test ที่ใช้ทดสอบมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกข้อ แสดงว่าข้อสอบมีค่าอำนาจจำแนกรายข้อทุกข้อ เมื่อนำคะแนนสอบของแต่ละข้อไปหาค่าสัมประสิทธิ์สหสัมพันธ์รายข้อกับคะแนนการสอบวัดความพึงพอใจในการสอบทั้งฉบับ เพื่อตรวจสอบความเที่ยงตรงเชิงโครงสร้าง (Construct Validity) ผลการวิเคราะห์ดังกล่าว แสดงได้ดังตาราง 16 และหาค่าความเชื่อมั่นของแบบสอบวัดความพึงพอใจในการสอบ ดังนี้

ตาราง 15 ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนรายข้อกับคะแนนรวมทั้งฉบับของแบบสอบวัดความพึงพอใจในการสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

ข้อสอบ	ค่าสหสัมพันธ์รายข้อ	ข้อสอบ	ค่าสหสัมพันธ์รายข้อ
1	.5006	14	.2168
2	.5356	15	.6522
3	.4790	16	.4568
4	.4501	17	.6007
5	.5459	18	.3854
6	.4226	19	.5628
7	.5439	20	.4662
8	.5388	21	.6638
9	.6254	22	.6428
10	.5692	23	.6623
11	.4835	24	.5687
12	.4347	25	.5509
13	.6304		

$r_{tt} = .9139$

จากตาราง 15 พบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนรายข้อกับคะแนนของแบบสอบวัดความพึงพอใจในการสอบทั้งฉบับ มีค่าสัมประสิทธิ์สหสัมพันธ์ทางบวกทุกข้ออยู่ระหว่าง 0.2168 ถึง 0.6638 และค่าความเชื่อมั่นของแบบสอบวัดเท่ากับ 0.9139 จากลักษณะของความสัมพันธ์ดังกล่าวพออนุมานได้ว่า ข้อสอบทุกข้อวัดโครงสร้างเดียวกัน และมีความเชื่อมั่นในการวัดสูง

2.3 ผลการวิเคราะห์ข้อมูลเพื่อประเมินคุณภาพของโปรแกรมการทดสอบ

การวิเคราะห์ข้อมูลเพื่อประเมินคุณภาพของโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ผู้วิจัยทำการประเมินโปรแกรมโดยพิจารณาจากผู้ประเมิน 2 กลุ่ม ได้แก่ กลุ่มของผู้เชี่ยวชาญด้านการวัดผลและผู้เชี่ยวชาญด้านการใช้โปรแกรมการทดสอบ เพื่อตรวจสอบความถูกต้องแม่นยำของการประมวลผลที่ได้จากการใช้โปรแกรมการทดสอบ และกลุ่มผู้ใช้โปรแกรมการทดสอบ ซึ่งประกอบด้วยผู้ดำเนินการสอบ เพื่อประเมินประสิทธิภาพของโปรแกรมการทดสอบด้านความสะดวกรวดเร็วในการใช้โปรแกรม ความแม่นยำของโปรแกรม และความผิดพลาดของโปรแกรม และนักเรียนกลุ่มตัวอย่างที่ทำการทดสอบ เพื่อประเมินคุณภาพโปรแกรมด้านความพึงพอใจในการทดสอบจากการทดสอบโดยการใช้โปรแกรม ซึ่งผลการวิเคราะห์ข้อมูลการประเมินคุณภาพของโปรแกรมปรากฏผล ได้ดังนี้

2.3.1 ผลการวิเคราะห์ข้อมูลการประเมินคุณภาพของโปรแกรมโดยกลุ่มผู้เชี่ยวชาญ พบว่า

1) โปรแกรมสามารถทำงานภายใต้เงื่อนไขของตัวแปรที่ผู้วิจัยกำหนด ได้แก่การจัดชุดข้อสอบให้ตอบในแต่ละชั้น การกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อ และการกำหนดรูปแบบของข้อสอบแบบมีสื่อประสม และแบบไม่มีสื่อประสม ได้ถูกต้อง

2) โปรแกรมสามารถทำการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์โดยใช้รูปแบบแยกทางแปรผัน ประมาณค่าความสามารถของผู้สอบโดยวิธีของเบส์ คัดเลือกข้อสอบข้อต่อไปจากค่าฟังก์ชันสารสนเทศของข้อสอบที่สูงที่สุด และยุติการทดสอบเมื่อค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบเท่ากับหรือที่ต่ำกว่า 0.3 หรือผู้สอบทำการทดสอบครบ 36 ข้อได้อย่างถูกต้องและรวดเร็ว

3) โปรแกรมสามารถบันทึกเวลาที่ใช้ในการสอบ จำนวนข้อที่ตอบถูก จำนวนข้อที่ใช้ในการทดสอบ และบอกตำแหน่งเปอร์เซ็นต์ไทล์ของผู้สอบได้ถูกต้อง

4) โปรแกรมสามารถคำนวณประมาณค่าความสามารถของผู้สอบ ค่าความคลาดเคลื่อนในการประมาณต่อความสามารถ ค่าฟังก์ชันสารสนเทศของข้อสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบได้ถูกต้องและรวดเร็ว

2.3.2 การประเมินคุณภาพ โดยกลุ่มผู้ใช้โปรแกรมการทดสอบ

1) การประเมินคุณภาพของโปรแกรมโดยผู้ดำเนินการสอบ พบว่า

1.1) การติดตั้งโปรแกรมสามารถทำได้ง่าย การใช้โปรแกรมในการทดสอบมีความสะดวกรวดเร็ว

1.2) การดำเนินการสอบใช้เวลาเฉลี่ยน้อยลง และสามารถรายงานผลการสอบได้ทันที

1.3) โปรแกรมสามารถคำนวณประมาณค่าความสามารถของผู้สอบ ความคลาดเคลื่อนในการประมาณค่าความสามารถ ค่าฟังก์ชันสารสนเทศของข้อสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ ได้ถูกต้องแม่นยำ เมื่อทำการตรวจสอบกับการคำนวณด้วยมือ

1.4) โปรแกรมสามารถบันทึกเวลาที่ใช้ในการสอบ จำนวนข้อที่ตอบถูกและจำนวนข้อที่ทำการทดสอบทั้งหมด ได้ถูกต้องครบถ้วน

1.5) ไม่พบความผิดพลาดในการทำงานของโปรแกรม และโปรแกรมมีความทนทานต่อการใช้งาน ถึงแม้ว่าผู้ใช้งานจะใช้โปรแกรมลัดขั้นตอนก็ตาม

2) การประเมินคุณภาพของโปรแกรมโดยนักเรียนกลุ่มตัวอย่าง

การประเมินโปรแกรมโดยนักเรียนกลุ่มตัวอย่างนั้น เป็นการประเมินคุณภาพของโปรแกรมด้านความพึงพอใจในการสอบ พบว่า โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อนำไปใช้สอบกับนักเรียนกลุ่มตัวอย่างแล้วนักเรียนกลุ่มตัวอย่างมีความพึงพอใจในการสอบในระดับมาก เมื่อเปรียบเทียบความพึงพอใจในการสอบระหว่างรูปแบบของข้อสอบที่มีสื่อประสมกับไม่มีสื่อประสม พบว่าผู้สอบมีความพึงพอใจในการสอบจากรูปแบบของข้อสอบที่มีสื่อประสมสูงกว่ารูปแบบของข้อสอบที่ไม่มีสื่อประสม ดังนั้นโปรแกรมการทดสอบที่

ผู้วิจัยพัฒนาขึ้นสามารถสร้างความพึงพอใจให้กับผู้สอบได้ดียิ่งขึ้นนำแบบสอบวัดความพึงพอใจในการสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ไปทดสอบหลังการทดลองใช้โปรแกรมการทดสอบ เพื่อศึกษาคุณภาพของโปรแกรมด้านความพึงพอใจในการสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ที่ผู้วิจัยพัฒนาขึ้น ได้ผลการทดสอบดังตาราง 17

ตาราง 16 จำนวนผู้สอบ ค่าเฉลี่ย และค่าความเบี่ยงเบนมาตรฐานของคะแนนความพึงพอใจในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ รูปแบบที่ไม่มีสื่อประสม และรูปแบบที่มีสื่อประสม

รูปแบบของข้อสอบ	N	\bar{x}	SD	t
ไม่มีสื่อประสม	61	90.4426	11.2894	-8.161**
มีสื่อประสม	74	106.1216	10.9581	
รวมเฉลี่ย	135	99.0370	13.5583	

** มีนัยสำคัญทางสถิติที่ระดับ .01

ตาราง 16 พบว่าการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ รูปแบบที่ไม่มีสื่อประสมมีผู้สอบจำนวน 61 คน ได้คะแนนเฉลี่ยความพึงพอใจในการสอบ 90.4426 คะแนน ความเบี่ยงเบนมาตรฐาน 11.2894 คะแนน ส่วนรูปแบบของข้อสอบที่มีสื่อประสมมีผู้สอบจำนวน 74 คน ได้คะแนนเฉลี่ยความพึงพอใจในการสอบ 106.1216 คะแนน ความเบี่ยงเบนมาตรฐาน 10.9581 คะแนน และคะแนนโดยเฉลี่ยของแบบสอบวัดความพึงพอใจในการสอบทั้งสองรูปแบบเท่ากับ 99.0370 คะแนน ความเบี่ยงเบนมาตรฐาน 13.5583 คะแนน เมื่อนำค่าเฉลี่ยของคะแนนสอบวัดความพึงพอใจในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ รูปแบบที่ไม่มีสื่อประสมกับรูปแบบที่มีสื่อประสมมาทดสอบความแตกต่าง โดยใช้ t-test พบว่า ค่าเฉลี่ยของคะแนนความพึงพอใจในการทดสอบแบบปรับเหมาะกับความสามารถของผู้ทดสอบด้วยคอมพิวเตอร์รูปแบบที่มีสื่อประสมมีค่าสูงกว่าค่าเฉลี่ยของคะแนนความพึงพอใจในการสอบรูปแบบที่ไม่มีสื่อประสมอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จากผลการทดสอบดังกล่าวจึงสรุปได้ว่าผู้สอบมีความพึงพอใจในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์รูปแบบที่มีสื่อประสม มากกว่ารูปแบบการทดสอบที่ไม่มีสื่อประสม หรือการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์รูปแบบที่มีสื่อประสมมีประสิทธิภาพด้านความพึงพอใจในการสอบมากกว่า

ตอนที่ 2 ผลการวิเคราะห์ข้อมูลในขั้นตอนหลัก

1. ผลการวิเคราะห์ค่าสถิติพื้นฐาน

ผลการวิเคราะห์ค่าเฉลี่ย และค่าความเบี่ยงเบนมาตรฐานของเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามอิทธิพลหลัก ปฏิสัมพันธ์ 2 ตัวแปร และ ปฏิสัมพันธ์ 3 ตัวแปร ผลการวิเคราะห์ข้อมูลได้ผลตามตาราง 17-19

ตาราง 17 ผลการวิเคราะห์ จำนวนผู้เข้าสอบ ค่าเฉลี่ย และค่าความเบี่ยงเบนมาตรฐานของเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ จำแนกตามระดับของอิทธิพลหลักของตัวแปรต้นแต่ละตัว

อิทธิพลหลัก 1 ตัวแปร	N	เวลาที่ใช้ ในการสอบ		จำนวนข้อของ แบบทดสอบ		สารสนเทศของ แบบทดสอบ	
		\bar{x}	SD	\bar{x}	SD	\bar{x}	SD
จำนวนข้อในแต่ละชั้น							
ชั้นละ 1 ข้อ	205	14.07	7.353	22.78	7.123	11.4618	0.2920
ชั้นละ 2 ข้อ	196	12.14	6.699	22.61	7.089	11.8176	0.5312
ชั้นละ 3 ข้อ	197	14.03	7.441	24.23	7.102	12.0453	0.8069
เวลาที่ใช้ในแต่ละข้อ							
ไม่จำกัดเวลา	201	15.41	7.933	24.13	6.969	11.7571	0.6480
เวลาเฉลี่ยเท่ากันทุกข้อ	200	12.32	5.796	23.00	6.885	11.7515	0.6193
เวลาตามค่าความยาก	197	12.12	6.625	22.45	7.461	11.8038	0.6155
รูปแบบของข้อสอบ							
ไม่มีสื่อประสม	300	14.26	7.097	23.69	7.229	11.7403	0.5981
มีสื่อประสม	298	12.05	6.777	22.70	7.006	11.8011	0.6548
รวม	598	13.16	7.022	23.20	7.130	11.7706	0.6272

จากตาราง 17 พบว่าอิทธิพลของจำนวนข้อสอบในแต่ละชั้นที่ทำให้ใช้เวลาในการทดสอบน้อยที่สุด ได้แก่ จำนวนข้อสอบชั้นละ 2 ข้อ จะใช้เวลาในการทดสอบเฉลี่ย 2.14 นาที (SD = 6.699) มีอิทธิพลต่อจำนวนข้อของแบบทดสอบที่ทำให้จำนวนข้อของแบบทดสอบน้อยที่สุด ได้แก่ จำนวนข้อสอบชั้นละ 2 ข้อ มีจำนวนข้อของแบบทดสอบเฉลี่ย 22.61 ข้อ (SD = 7.089) และมีอิทธิพลต่อค่าฟังก์ชันสารสนเทศของแบบทดสอบมากที่สุด ได้แก่ จำนวนข้อสอบชั้นละ 3 ข้อ ทำให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงที่สุดเฉลี่ย 12.0453 (SD = 0.8069)

อิทธิพลของการกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อที่ทำให้ใช้เวลาในการทดสอบน้อยที่สุด ได้แก่ การกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อตามค่าความยากของข้อสอบ จะใช้เวลาในการทดสอบน้อยที่สุดเฉลี่ย 12.12 นาที (SD = 6.625) มีอิทธิพลต่อจำนวนข้อของแบบทดสอบที่ทำให้จำนวนของแบบทดสอบน้อยที่สุด ได้แก่ การกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อตามค่าความยากของข้อสอบ มีจำนวนข้อของแบบทดสอบน้อยที่สุดเฉลี่ย 22.45 ข้อ (SD = 7.461) และมีอิทธิพลต่อค่าฟังก์ชันสารสนเทศของแบบทดสอบมากที่สุดคือ การกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อตามค่าความยากของข้อสอบ มีค่าฟังก์ชันสารสนเทศของแบบทดสอบเฉลี่ย 11.8038 (SD = 0.6155)

อิทธิพลรูปแบบของข้อสอบที่ไม่มีสื่อประสมกับรูปแบบที่มีสื่อประสมที่ทำให้ใช้เวลาในการทดสอบน้อยที่สุด ได้แก่ รูปแบบของข้อสอบแบบที่มีสื่อประสม ใช้เวลาในการทดสอบน้อยที่สุดเฉลี่ย 12.05 นาที (SD = 6.777) มีอิทธิพลต่อจำนวนข้อของแบบทดสอบที่ทำให้จำนวนข้อของแบบทดสอบน้อยที่สุด ได้แก่ รูปแบบของข้อสอบแบบที่มีสื่อประสมเฉลี่ย 22.70 ข้อ (SD = 7.006) และมีอิทธิพลต่อค่าฟังก์ชันสารสนเทศของแบบทดสอบมากที่สุดคือรูปแบบของข้อสอบที่มีสื่อประสมมีค่าฟังก์ชันสารสนเทศของแบบทดสอบเฉลี่ย 11.8011 (SD = 0.6548)

ตาราง 18 ผลการวิเคราะห์ข้อมูล จำนวนผู้เข้าสอบ ค่าเฉลี่ย และความเบี่ยงเบนมาตรฐานของเวลาที่ใช้ในการทำแบบทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จากอิทธิพลร่วมของระดับต่างๆ ของตัวแปรต้นร่วมกัน 2 ระดับ

อิทธิพลร่วม 2 ตัวแปร	N	เวลาที่ใช้ในการสอบ		จำนวนข้อของแบบทดสอบ		สารสนเทศของแบบทดสอบ	
		\bar{x}	SD	\bar{x}	SD	\bar{x}	SD
ข้อสอบชั้นละ 1 ข้อ x							
ไม่จำกัดเวลา	72	16.32	8.6930	23.78	6.863	11.4223	0.2469
เวลาเท่ากันทุกข้อ	69	12.53	6.2388	22.43	7.269	11.5214	0.3361
เวลาตามความยาก	64	12.35	5.9557	22.02	7.232	11.4419	0.2817
ไม่มีสื่อประสม	102	14.49	7.0149	23.12	7.088	11.4428	0.2758
มีสื่อประสม	103	13.28	7.6622	22.44	7.175	11.4806	0.3074
ข้อสอบชั้นละ 2 ข้อ x							
ไม่จำกัดเวลา	62	14.30	6.1118	23.48	7.403	11.8507	0.5037
เวลาเท่ากันทุกข้อ	63	11.46	5.1748	22.16	6.149	11.6698	0.4608
เวลาตามความยาก	71	11.21	6.5828	22.24	7.598	11.9197	0.5879
ไม่มีสื่อประสม	97	13.22	6.3245	23.21	7.426	11.8273	0.5469
มีสื่อประสม	99	11.09	5.7075	22.02	6.728	11.8080	0.5179

ตาราง 18 (ต่อ)

อิทธิพลร่วม 2 ตัวแปร	N	เวลาที่ใช้ในการสอบ		จำนวนข้อของแบบทดสอบ		สารสนเทศของแบบทดสอบ	
		\bar{x}	SD	\bar{x}	SD	\bar{x}	SD
ข้อสอบชั้นละ 3 ข้อ x							
ไม่จำกัดเวลา	67	16.22	8.4682	25.10	6.661	12.0302	0.8752
เวลาเท่ากันทุกข้อ	68	13.31	5.8653	24.35	7.022	12.0607	0.8185
เวลาตามความยาก	62	12.07	7.3632	23.15	7.607	12.0446	0.7257
ไม่มีสื่อประสม	101	15.04	7.7980	24.74	7.134	11.9572	0.7488
มีสื่อประสม	96	12.15	6.7749	23.69	7.066	12.1379	0.8579
ไม่จำกัดเวลา x							
ไม่มีสื่อประสม	103	17.14	8.1247	24.82	7.098	11.7332	0.5379
มีสื่อประสม	98	14.00	7.5108	23.41	6.793	11.7823	0.7485
เวลาเท่ากันทุกข้อ x							
ไม่มีสื่อประสม	100	13.43	5.6733	23.40	7.100	11.6673	0.5179
มีสื่อประสม	100	11.21	5.7323	22.60	6.674	11.8357	0.6804
เวลาตามความยาก x							
ไม่มีสื่อประสม	97	12.50	6.5661	22.80	7.417	11.8232	0.7011
มีสื่อประสม	100	11.37	6.6295	22.11	7.525	11.7850	0.5221
รวม	598	13.16	7.0223	23.20	7.130	11.7706	0.6272

จากตาราง 18 พบว่า อิทธิพลร่วมของระดับต่างๆ ของตัวแปรต้นร่วมกัน 2 ระดับ ที่มีต่อการใช้เวลาในการทดสอบ คือทำให้ใช้เวลาในการทดสอบน้อยที่สุดได้แก่ การจัดชุดข้อสอบให้ตอบชั้นละ 2 ข้อ และรูปแบบของข้อสอบแบบที่มีสื่อประสมจะใช้เวลาในการทดสอบน้อยที่สุดเฉลี่ย 11.09 นาที (SD = 5.7075) และทำให้ใช้เวลาในการทดสอบมากที่สุดได้แก่ ไม่จำกัดเวลาในการตอบข้อสอบในแต่ละข้อ และรูปแบบของข้อสอบแบบที่มีสื่อประสม จะใช้เวลาในการสอบมากที่สุดเฉลี่ย 17.14 นาที (SD = 8.1247)

อิทธิพลร่วมของระดับต่างๆ ของตัวแปรต้นร่วมกัน 2 ระดับ ที่มีต่อจำนวนข้อของแบบทดสอบ คือทำให้จำนวนข้อของแบบทดสอบน้อยที่สุดได้แก่ การจัดชุดข้อสอบให้ตอบชั้นละ 2 ข้อ และรูปแบบของข้อสอบแบบที่มีสื่อประสม ทำให้จำนวนข้อของแบบทดสอบน้อยที่สุดเฉลี่ย 22.02 ข้อ (SD = 6.728) และทำให้จำนวนข้อของแบบทดสอบมากที่สุดได้แก่ การจัดชุดข้อสอบให้ตอบชั้นละ 3 ข้อ และไม่จำกัดเวลาในการตอบข้อสอบในแต่ละข้อ ทำให้จำนวนของแบบทดสอบมากที่สุดเฉลี่ย 25.10 ข้อ (SD = 6.661)

ส่วนอิทธิพลร่วมของระดับต่างๆ ของตัวแปรต้นร่วมกัน 2 ระดับที่มีต่อค่าฟังก์ชันสารสนเทศของแบบทดสอบ คือทำให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบมีค่ามากที่สุดได้แก่ การจัดชุดข้อสอบให้ตอบชั้นละ 3 ข้อ และรูปแบบของข้อสอบแบบที่มีสื่อประสม ทำให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงที่สุดเฉลี่ย 12.1379 (SD = 0.8579) และทำให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบต่ำที่สุดได้แก่ การจัดชุดข้อสอบให้ตอบชั้นละ 1 ข้อ และไม่จำกัดเวลาในการตอบข้อสอบในแต่ละข้อ ได้ค่าฟังก์ชันสารสนเทศของแบบทดสอบเฉลี่ย 11.4223 (SD = 0.2469)

ตาราง 19 ผลการวิเคราะห์ข้อมูล จำนวนผู้เข้าสอบ ค่าเฉลี่ย และค่าความเบี่ยงเบนมาตรฐานของเวลาที่ใช้ทำแบบทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จากระดับต่าง ๆ ของตัวแปรต้นร่วมกัน 3 ระดับ

อิทธิพลร่วม 3 ตัวแปร	N	เวลาที่ใช้ในการสอบ		จำนวนข้อของแบบทดสอบ		สารสนเทศของแบบทดสอบ	
		\bar{x}	SD	\bar{x}	SD	\bar{x}	SD
		ไม่มีสื่อประสม					
ชั้นละ 1 ข้อ ไม่จำกัดเวลา	37	16.54	8.5861	23.84	6.747	11.4705	0.2501
ชั้นละ 1 ข้อ เวลาเท่ากันทุกข้อ	34	13.27	6.2098	22.18	7.171	11.4790	0.3259
ชั้นละ 1 ข้อ เวลาตามความยาก	31	13.37	5.1596	23.29	7.502	11.3701	0.2370
ชั้นละ 2 ข้อ ไม่จำกัดเวลา	31	16.40	5.8998	24.90	7.997	11.9910	0.5106
ชั้นละ 2 ข้อ เวลาเท่ากันทุกข้อ	31	13.25	5.5005	22.52	6.923	11.6335	0.4493
ชั้นละ 2 ข้อ เวลาตามความยาก	35	11.09	6.6125	22.31	7.287	11.9338	0.6208
ชั้นละ 3 ข้อ ไม่จำกัดเวลา	35	18.02	9.3765	25.77	6.673	11.8622	0.6737
ชั้นละ 3 ข้อ เวลาเท่ากันทุกข้อ	35	14.10	5.3671	25.37	6.954	11.8801	0.7003
ชั้นละ 3 ข้อ เวลาตามความยาก	31	13.21	7.6256	22.87	7.684	12.1515	0.8617
มีสื่อประสม							
ชั้นละ 1 ข้อ ไม่จำกัดเวลา	35	16.08	8.9236	23.71	7.082	11.3714	0.2362
ชั้นละ 1 ข้อ เวลาเท่ากันทุกข้อ	35	12.21	6.2714	22.69	7.459	11.5626	0.3455
ชั้นละ 1 ข้อ เวลาตามความยาก	33	11.02	6.4077	20.82	6.867	11.5094	0.3064

ตาราง 19 (ต่อ)

อิทธิพลร่วม 3 ตัวแปร	N	เวลาที่ใช้ในการสอบ		จำนวนข้อของแบบทดสอบ		สารสนเทศของแบบทดสอบ	
		\bar{x}	SD	\bar{x}	SD	\bar{x}	SD
ชั้นละ 2 ข้อ ไม่จำกัดเวลา	31	17.39	5.8998	24.90	7.997	11.9007	0.5106
ชั้นละ 2 ข้อ เวลาเท่ากันทุกข้อ	32	10.12	4.5268	21.81	5.385	11.7049	0.4761
ชั้นละ 2 ข้อ เวลาตามความยาก	36	11.33	6.6454	22.17	11.3294	11.9061	0.5626
ชั้นละ 3 ข้อ ไม่จำกัดเวลา	32	14.25	6.9722	24.37	6.676	12.2140	1.0327
ชั้นละ 3 ข้อ เวลาเท่ากันทุกข้อ	33	12.05	6.1839	23.27	7.037	12.2523	0.8990
ชั้นละ 3 ข้อ เวลาตามความยาก	31	11.33	7.0301	23.42	7.645	11.9377	0.5521
รวมทั้งสิ้น	598	13.16	7.0223	23.20	7.130	11.7706	0.6272

จากตาราง 19 พบว่าอิทธิพลร่วม 3 ตัวแปร ที่ส่งผลต่อเวลาที่ใช้ในการทดสอบ คือ ทำให้ใช้เวลาในการทดสอบน้อยที่สุดได้แก่ การจัดชุดข้อสอบให้ตอบชั้นละ 2 ข้อ การกำหนดเวลาให้ตอบเท่ากันทุกข้อ และรูปแบบของข้อสอบแบบที่มีสื่อประสมจะใช้เวลาในการสอบน้อยที่สุดเฉลี่ย 10.12 นาที (SD = 4.5268) ส่งผลต่อจำนวนข้อของแบบทดสอบ คือ จำนวนข้อของแบบทดสอบน้อยที่สุดได้แก่ จัดชุดข้อสอบให้ตอบชั้นละ 1 ข้อ กำหนดเวลาให้ตอบตามค่าความยากของข้อสอบและรูปแบบของข้อสอบแบบที่มีสื่อประสม ทำให้จำนวนข้อของแบบทดสอบน้อยที่สุดเฉลี่ย 20.82 ข้อ (SD = 6.867) และส่งผลต่อค่าฟังก์ชันสารสนเทศของแบบทดสอบมากที่สุดคือ จัดชุดข้อสอบให้ตอบชั้นละ 3 ข้อ กำหนดเวลาให้ตอบเท่ากันทุกข้อ และรูปแบบของข้อสอบแบบที่มีสื่อประสม ให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงที่สุดเฉลี่ย 12.2523 (SD = 0.8990)

2. ผลการวิเคราะห์ข้อมูลเพื่อศึกษาผลของตัวแปรอิสระที่มีต่อตัวแปรตาม

การวิเคราะห์คุณภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์นั้น จะต้องคำนึงถึงค่าฟังก์ชันสารสนเทศของแบบทดสอบกล่าวคือ การทดสอบจะต้องมีความคลาดเคลื่อนมาตรฐานในการวัดค่า ค่าฟังก์ชันสารสนเทศของการทดสอบจะต้องมีค่าสูง และจะต้องใช้เวลาในการสอบและจำนวนข้อสอบน้อยลง หรือยังคงเท่าเดิม ในการวิจัยในครั้งนี้ผู้วิจัยเห็นว่าตัวแปรการจัดชุดข้อสอบให้ตอบในแต่ละชั้น การกำหนดเวลาที่ใช้ในการสอบแต่ละข้อ และรูปแบบของข้อสอบที่มีสื่อประสม กับไม่มีสื่อประสม น่าจะมีอิทธิพลต่อคุณภาพดังกล่าว ดังนั้นในการศึกษาวิจัยครั้งนี้ ผู้วิจัยจึงสนใจศึกษาอิทธิพลของตัวแปรต่าง ๆ ที่มีต่อเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ ดังผลการวิเคราะห์ตามตาราง ต่อไปนี้

ตาราง 20 ผลการวิเคราะห์ความแปรปรวนของเวลาที่ใช้ในการทดสอบ จำนวนข้อสอบของการทดสอบและค่าฟังก์ชันสารสนเทศของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามแหล่งความแปรปรวน

แหล่งความแปรปรวน/ ตัวแปรตาม	SS	df	MS	F
ผลหลัก				
จำนวนข้อที่ใช้ในแต่ละชั้น (A)/				
เวลาที่ใช้ในการสอบ	254.138	2	127.069	2.771
จำนวนข้อของแบบทดสอบ	292.832	2	146.416	2.891
สารสนเทศของแบบทดสอบ	35.308	2	17.654	53.487**
เวลาที่ใช้ในการตอบแต่ละข้อ (B)/				
เวลาที่ใช้ในการสอบ	1462.908	2	731.454	15.953**
จำนวนข้อของแบบทดสอบ	278.112	2	139.056	2.746
สารสนเทศของแบบทดสอบ	.245	2	.123	.374
รูปแบบของข้อสอบ (C)/				
เวลาที่ใช้ในการสอบ	724.770	1	724.770	15.807**
จำนวนข้อของแบบทดสอบ	139.947	1	139.947	2.763
สารสนเทศของแบบทดสอบ	.613	1	.613	1.858
ปฏิสัมพันธ์ 2 ทาง				
(AxB) / เวลาที่ใช้ในการสอบ				
จำนวนข้อของแบบทดสอบ	54.509	4	13.627	.297
สารสนเทศของแบบทดสอบ	28.248	4	7.062	.139
สารสนเทศของแบบทดสอบ	2.362	4	.591	1.789
(AxC) / เวลาที่ใช้ในการสอบ				
จำนวนข้อของแบบทดสอบ	39.956	2	19.978	.436
สารสนเทศของแบบทดสอบ	7.154	2	3.577	.071
สารสนเทศของแบบทดสอบ	.914	2	.457	1.385
(BxC) / เวลาที่ใช้ในการสอบ				
จำนวนข้อของแบบทดสอบ	36.003	2	18.002	.393
จำนวนข้อของแบบทดสอบ	17.608	2	8.804	.174
สารสนเทศของแบบทดสอบ	1.103	2	.552	1.671
ปฏิสัมพันธ์ 3 ทาง (AxBxC)/				
เวลาที่ใช้ในการสอบ	201.714	4	50.428	1.100
จำนวนข้อของแบบทดสอบ	191.099	4	47.775	.943
สารสนเทศของแบบทดสอบ	3.242	4	.811	2.456
ความคลาดเคลื่อน /				
เวลาที่ใช้ในการสอบ	26592.873	580	45.850	
จำนวนข้อของแบบทดสอบ	29375.111	580	50.647	
สารสนเทศของแบบทดสอบ	191.436	580	.330	

จากตาราง 20 เมื่อพิจารณาอิทธิพลหลักของตัวแปรต้น ต่อเวลาที่ใช้ในการทดสอบจำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ โดยไม่พิจารณาปฏิสัมพันธ์ระหว่างตัวแปรอิสระแล้ว พบว่า จำนวนข้อสอบที่ใช้ในการตอบในแต่ละชั้นส่งผลต่อค่าฟังก์ชันสารสนเทศของแบบทดสอบ เวลาที่ใช้ในการตอบแต่ละข้อส่งผลต่อเวลาที่ใช้ในการทดสอบ และรูปแบบของข้อสอบส่งผลต่อเวลาที่ใช้ในการทดสอบ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

หากพิจารณาปฏิสัมพันธ์ร่วมกันของตัวแปรต้น 2 ตัว ต่อเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ พบว่าปฏิสัมพันธ์ร่วมของตัวแปรอิสระ 2 ตัว ไม่ส่งผลต่อเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ

และเมื่อพิจารณาปฏิสัมพันธ์ร่วมกันของตัวแปรต้นทั้ง 3 ตัวต่อเวลาที่ใช้ในการทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ พบว่าปฏิสัมพันธ์ร่วมของตัวแปรอิสระ 3 ตัว ไม่ส่งผลต่อเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ

เมื่อเปรียบเทียบค่าเฉลี่ยของเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ เป็นรายคู่ ตามระดับของตัวแปรอิสระทั้ง 3 ตัวแปร คือ จำนวนข้อสอบในแต่ละชั้นเวลาที่ใช้ในการตอบแต่ละข้อ และรูปแบบของข้อสอบแบบที่มีสื่อประสมกับแบบที่ไม่มีสื่อประสม ได้ผลการวิเคราะห์ ดังตาราง 21-24

ตาราง 21 ความแตกต่างค่าเฉลี่ยของเวลาที่ใช้ในการทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามตัวแปรจำนวนข้อสอบในแต่ละชั้น

ค่าเฉลี่ย	\bar{X} ชั้นละ 1 ข้อ	\bar{X} ชั้นละ 2 ข้อ	\bar{X} ชั้นละ 3 ข้อ
เวลาที่ใช้ในการสอบ	14.07	12.14	14.03
\bar{X} ชั้นละ 1 ข้อ	-	1.93	0.04
\bar{X} ชั้นละ 2 ข้อ			-1.89
\bar{X} ชั้นละ 3 ข้อ			
จำนวนข้อของแบบทดสอบ	22.78	22.61	24.23
\bar{X} ชั้นละ 1 ข้อ	-	0.17	-1.45
\bar{X} ชั้นละ 2 ข้อ			-1.62
\bar{X} ชั้นละ 3 ข้อ			-
สารสนเทศของแบบทดสอบ	11.462	11.818	12.045
\bar{X} ชั้นละ 1 ข้อ	-	-0.356**	-0.583**
\bar{X} ชั้นละ 2 ข้อ		-	-0.227**
\bar{X} ชั้นละ 3 ข้อ			-

** มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 21 พบว่าการจัดชุดข้อสอบให้ตอบชั้นละ 1 ข้อ กับชั้นละ 2 ข้อ การจัดชุดข้อสอบให้ตอบชั้นละ 1 ข้อ กับชั้นละ 3 ข้อ และการจัดชุดข้อสอบให้ตอบชั้นละ 2 ข้อ กับชั้นละ 3 ข้อ จะใช้เวลาในการทดสอบ และใช้จำนวนข้อสอบในการทดสอบไม่ต่างกัน ส่วนการทดสอบด้วยการจัดชุดข้อสอบให้ตอบชั้นละ 1 ข้อ กับชั้นละ 2 ข้อ การจัดชุดข้อสอบให้ตอบชั้นละ 1 ข้อกับชั้นละ 3 ข้อ และการจัดชุดข้อสอบให้ตอบชั้นละ 2 ข้อ กับชั้นละ 3 ข้อ ทำให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ตาราง 22 ความแตกต่างค่าเฉลี่ยของเวลาที่ใช้ในการทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ จำแนกตามตัวแปรเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อ

ค่าเฉลี่ย	\bar{X} ไม่จำกัดเวลา	\bar{X} เวลาเท่ากันทุกข้อ	\bar{X} เวลาตามความยาก
เวลาที่ใช้ในการสอบ	15.41	12.32	12.12
\bar{X} ไม่จำกัดเวลา	-	3.09**	3.29**
\bar{X} เวลาเท่ากันทุกข้อ		-	0.20
\bar{X} เวลาตามความยาก			-
จำนวนข้อของแบบทดสอบ	24.13	23.00	22.45
\bar{X} ไม่จำกัดเวลา	-	1.13	1.68
\bar{X} เวลาเท่ากันทุกข้อ		-	0.55
\bar{X} เวลาตามความยาก			-
สารสนเทศของแบบทดสอบ	11.757	11.752	11.804
\bar{X} ไม่จำกัดเวลา	-	0.005	-0.047
\bar{X} เวลาเท่ากันทุกข้อ		-	-0.052
\bar{X} เวลาตามความยาก			-

** มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 22 พบว่าการทดสอบด้วยข้อสอบที่ไม่จำกัดเวลาในการตอบ จะใช้เวลาในการทดสอบต่างจากการทดสอบด้วยข้อสอบที่จำกัดเวลาให้ตอบแต่ละข้อเท่ากันทุกข้อ และต่างจากการทดสอบด้วยข้อสอบที่จำกัดเวลาให้ตอบตามความยากง่ายของข้อสอบ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 การทดสอบด้วยข้อสอบที่จำกัดเวลาให้ตอบเท่ากันทุกข้อกับจำกัดเวลาให้ตอบตามความยากง่ายของข้อสอบจะใช้เวลาในการทดสอบไม่ต่างกัน ส่วนการทดสอบด้วยข้อสอบที่ไม่จำกัดเวลาในการตอบกับข้อสอบที่จำกัดเวลาให้ตอบเท่ากันทุกข้อ การทดสอบด้วยข้อสอบที่ไม่จำกัดเวลาในการตอบกับข้อสอบที่จำกัดเวลาให้ตอบตามความยากง่ายของข้อสอบ และการสอบด้วยข้อสอบที่จำกัดเวลาให้ตอบเท่ากันทุกข้อกับข้อสอบที่จำกัดเวลาให้ตอบตามความยากง่ายของข้อสอบ จะใช้ข้อสอบจำนวนไม่ต่างกัน และค่าฟังก์ชันสารสนเทศของแบบทดสอบก็ไม่ต่างกัน

ตาราง 23 ความแตกต่างของค่าเฉลี่ยของเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์จำแนกตามตัวแปรรูปแบบของข้อสอบ

ค่าเฉลี่ย	\bar{X} ไม่มีสื่อประสม	\bar{X} มีสื่อประสม
เวลาที่ใช้ในการสอบ	14.26	12.05
\bar{X} ไม่มีสื่อประสม	-	2.21**
\bar{X} มีสื่อประสม		
จำนวนข้อของแบบทดสอบ	23.69	22.70
\bar{X} ไม่มีสื่อประสม	-	0.99
\bar{X} มีสื่อประสม		
สารสนเทศของแบบทดสอบ	11.740	11.801
\bar{X} ไม่มีสื่อประสม	-	-0.061
\bar{X} มีสื่อประสม		

**มีนัยสำคัญทางสถิติที่ระดับ .01

จากตาราง 23 พบว่า การทดสอบด้วยข้อสอบรูปแบบที่ไม่มีสื่อประสมจะใช้เวลาในการสอบต่างจากการทดสอบด้วยข้อสอบรูปแบบที่มีสื่อประสม อย่างมีนัยสำคัญที่ระดับ .01 ส่วนการทดสอบด้วยข้อสอบรูปแบบที่ไม่มีสื่อประสมกับข้อสอบรูปแบบที่มีสื่อประสม จำนวนข้อสอบที่ใช้ทำการทดสอบไม่ต่างกัน และค่าฟังก์ชันสารสนเทศของแบบทดสอบก็ไม่ต่างกัน

เพื่อให้เห็นความแตกต่าง ของเวลาที่ใช้ในการทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ จำแนกตามอิทธิพลของตัวแปรต้น ได้แก่การจัดชุดข้อสอบให้ตอบในแต่ละชั้น การกำหนดเวลาให้ตอบในแต่ละข้อ และการจัดรูปแบบของข้อสอบรูปแบบที่ไม่มีสื่อประสมกับรูปแบบที่มีสื่อประสม ดังตาราง 24

ตาราง 24 เวลาที่ใช้ในการทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ จำแนกตามอิทธิพลของตัวแปรต้น ได้แก่ การจัดชุดข้อสอบให้ตอบในแต่ละชั้น การกำหนดเวลาให้ตอบในแต่ละข้อ และการจัดรูปแบบของข้อสอบรูปแบบที่ไม่มีสื่อประสมกับรูปแบบที่มีสื่อประสม

ตัวแปรอิสระ	เวลาที่ใช้ในการสอบ	จำนวนข้อของแบบทดสอบ	ฟังก์ชันสารสนเทศของแบบทดสอบ
1. จำนวนข้อในแต่ละชั้น	ไม่ต่าง	ไม่ต่าง	ต่าง
- ชั้นละ 1 ข้อ	14.07	22.78	11.462
- ชั้นละ 2 ข้อ	12.14	22.61	11.818
- ชั้นละ 3 ข้อ	14.03	24.23	12.045
2. เวลาที่ใช้ในแต่ละข้อ	ต่าง	ไม่ต่าง	ไม่ต่าง
- ไม่จำกัดเวลา	15.41	24.13	11.757
- เวลาเฉลี่ยเท่ากันทุกข้อ	12.32	23.00	11.752
- เวลาตามค่าความยาก	12.12	22.45	11.804
3. รูปแบบของข้อสอบ	ต่าง	ไม่ต่าง	ไม่ต่าง
- ไม่มีสื่อประสม	14.26	23.69	11.740
- มีสื่อประสม	12.05	22.70	118.01

จากตาราง 24 พบว่า เมื่อจัดชุดข้อสอบให้ตอบในแต่ละชั้นต่างกัน ไม่ทำให้เวลาที่ใช้ในการทดสอบ และจำนวนของแบบทดสอบต่างกัน แต่ทำให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบต่างกัน โดยการจัดชุดข้อสอบให้ตอบชั้นละ 3 ข้อ ทำให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงที่สุด รองลงมาคือจัดชุดข้อสอบให้ตอบชั้นละ 2 ข้อ และจัดชุดข้อสอบให้ตอบชั้นละ 1 ข้อ ให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบต่ำที่สุด

เมื่อกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อต่างกัน ทำให้เวลาที่ใช้ในการทดสอบต่างกัน โดยที่การกำหนดเวลาในการตอบข้อสอบในแต่ละข้อตามน้ำหนักค่าความยากของข้อสอบ และการกำหนดเวลาในการตอบแต่ละข้อเฉลี่ยเท่ากันทุกข้อ ต่างจากการไม่กำหนดเวลาในการตอบข้อสอบในแต่ละข้อ ซึ่งการกำหนดเวลาในการตอบข้อสอบในแต่ละข้อตามน้ำหนักค่าความยากของข้อสอบใช้เวลาสอบน้อยที่สุด ส่วนการไม่กำหนดเวลาในการตอบข้อสอบในแต่ละข้อจะใช้เวลาในการทดสอบมากที่สุด ส่วนจำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบไม่ต่างกัน

เมื่อจัดรูปแบบของข้อสอบแบบที่ไม่มีสื่อประสมกับรูปแบบของข้อสอบแบบที่มีสื่อประสม ประกอบการทดสอบทำให้เวลาที่ใช้ในการทดสอบต่างกัน โดยการจัดรูปแบบข้อสอบแบบที่ไม่มีสื่อประสมใช้เวลาในการสอบมากกว่ารูปแบบที่มีสื่อประสม ส่วนจำนวนข้อของแบบทดสอบและค่าฟังก์ชันสารสนเทศของแบบทดสอบไม่ต่างกัน

บทที่ 5

สรุปผล อภิปราย และข้อเสนอแนะ

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้ มีวัตถุประสงค์เพื่อพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พร้อมทั้งศึกษาประสิทธิภาพของโปรแกรมการทดสอบ และเพื่อเปรียบเทียบเวลาที่ใช้ในการทดสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ เมื่อมีการจัดชุดข้อสอบให้ตอบในแต่ละชั้นต่างกัน การกำหนดเวลาที่ใช้ในการตอบแต่ละข้อต่างกัน และการจัดรูปแบบของข้อสอบที่ไม่มีสื่อประสมและรูปแบบที่มีสื่อประสม

สมมติฐานในการวิจัย

1. เมื่อจัดชุดข้อสอบให้ตอบในแต่ละชั้นต่างกันค่าฟังก์ชันสารสนเทศของแบบทดสอบน่าจะต่างกัน
2. เมื่อกำหนดเวลาที่ใช้ในการตอบข้อสอบแต่ละข้อตามน้ำหนักค่าความยากของข้อสอบฟังก์ชันสารสนเทศของแบบทดสอบจะมีค่ามากที่สุด
3. ฟังก์ชันสารสนเทศของแบบทดสอบที่รูปแบบของข้อสอบมีสื่อประสม น่าจะมีค่ามากกว่ารูปแบบของข้อสอบที่ไม่มีสื่อประสม
4. เมื่อจัดชุดข้อสอบให้ตอบในแต่ละชั้นและเมื่อกำหนดเวลาให้ตอบข้อสอบในแต่ละข้อต่างกันจะร่วมกันส่งผลทำให้ฟังก์ชันสารสนเทศของแบบทดสอบต่างกัน
5. เมื่อกำหนดเวลาให้ตอบข้อสอบในแต่ละข้อและการจัดรูปแบบของข้อสอบที่มีสื่อประสมต่างกันจะร่วมกันส่งผลทำให้ฟังก์ชันสารสนเทศของแบบทดสอบต่างกัน

วิธีดำเนินการวิจัย

การดำเนินการวิจัยในครั้งนี้ผู้วิจัยแบ่งการดำเนินการวิจัยออกเป็น 2 ระยะ ซึ่งมีรายละเอียดในการวิจัย ดังต่อไปนี้

ระยะที่ 1 เป็นการศึกษาเพื่อพัฒนาเครื่องมือที่ใช้ในการวิจัย ประกอบด้วย

1.1 การสร้างและพัฒนาคลังข้อสอบ เป็นการสร้างข้อสอบเพิ่มเติมขึ้นเพื่อวัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 ข้อสอบที่สร้างขึ้น จำนวน 54 ข้อ กลุ่มตัวอย่างคือ นักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 1 ปีการศึกษา 2546 โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษา ปทุมธานี เขตที่ 1 และ 2 จำนวนนักเรียนกลุ่มตัวอย่าง 597 คน

1.2 การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ โปรแกรมการทดสอบที่ผู้วิจัยพัฒนาขึ้นนี้ ผู้วิจัยพัฒนาโปรแกรมให้สามารถทำงานบนระบบปฏิบัติการวินโดวส์ 98 ขึ้นไป โดยใช้โปรแกรมบอร์แลนเดลฟายในการเขียนโปรแกรม ซึ่งโปรแกรม

สามารถใช้งานกับเครื่องไมโครคอมพิวเตอร์ได้ โปรแกรมการทดสอบที่ผู้วิจัยพัฒนาขึ้น มีจำนวนทั้งสิ้น 18 โปรแกรม ซึ่งได้จากการจัดกลุ่มของตัวแปรต้น และแต่ละโปรแกรมจะประกอบด้วย โปรแกรมย่อย สำหรับการพิมพ์เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบ กำหนดค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบ และการกำหนดคุณลักษณะของแบบทดสอบตามชนิดของตัวแปรต้น ได้แก่การจัดชุดข้อสอบให้ตอบในแต่ละชั้น การกำหนดเวลาที่ใช้ในการตอบแต่ละข้อ และการกำหนดรูปแบบของข้อสอบที่มีสื่อประสม และแบบไม่มีสื่อประสม และโปรแกรมสำหรับการทดสอบ ซึ่งเป็นโปรแกรมสำเร็จรูปสำหรับการทดสอบบนหน้าจอบนคอมพิวเตอร์ ผู้สอบจะใช้โปรแกรมการทดสอบในการทดสอบ โดยผู้สอบจะต้องเริ่มการทดสอบด้วยการกรอกข้อมูลส่วนตัวให้สมบูรณ์จึงจะสามารถเริ่มทำการทดสอบได้ ขณะทำการทดสอบ ผู้สอบจะต้องทำข้อสอบและเลือกตอบคำถามบนหน้าจอบนคอมพิวเตอร์ เมื่อยุติการทดสอบตามเกณฑ์ ผู้สอบจะทราบผลการสอบทันที และเครื่องคอมพิวเตอร์ จะบันทึกผลการทดสอบของผู้สอบทุกคนไว้ ทั้งนี้ในการกำหนดเวลาที่ใช้ในการตอบแต่ละข้อนั้น ผู้วิจัยได้นำโปรแกรมการทดสอบไปทดสอบกับนักเรียนกลุ่มตัวอย่างเพื่อศึกษาข้อมูลในการกำหนดเวลาให้กับข้อสอบแต่ละข้อ กับกลุ่มตัวอย่าง นักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 1 ปีการศึกษา 2546 โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาปทุมธานี เขตที่ 1 และ 2 จำนวนนักเรียนกลุ่มตัวอย่าง 235 คน

1.3 การศึกษาประสิทธิภาพของโปรแกรมการทดสอบ เมื่อผู้วิจัยพัฒนาโปรแกรมการทดสอบเสร็จสิ้นลง ผู้วิจัยได้นำโปรแกรมไปให้ผู้เชี่ยวชาญตรวจสอบประสิทธิภาพในการทำงานของโปรแกรม และตรวจสอบความถูกต้องแม่นยำในการประมวลผลการสอบ หลังจากนั้น จึงนำโปรแกรมการทดสอบที่ผ่านการพิจารณาของผู้เชี่ยวชาญไปทดลองใช้โดยกลุ่มผู้ใช้โปรแกรมการทดสอบ ซึ่งประกอบด้วย ผู้ดำเนินการสอบ เพื่อประเมินประสิทธิภาพของโปรแกรมการทดสอบด้านความสะดวกรวดเร็วในการใช้โปรแกรม ความแม่นยำของโปรแกรม และความทนทานต่อความผิดพลาดในการใช้งานของโปรแกรม และทดลองใช้โดยนักเรียนกลุ่มตัวอย่างที่ทำการทดสอบ เพื่อศึกษาประสิทธิภาพของโปรแกรมการทดสอบด้านความพึงพอใจในการสอบ กับกลุ่มตัวอย่างนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2546 โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาปทุมธานี เขตที่ 1 และ 2 จำนวนนักเรียนกลุ่มตัวอย่าง 135 คน

ระยะที่ 2 เป็นการศึกษาวิจัยผลของตัวแปรอิสระต่อตัวแปรตาม

การศึกษาวิจัยผลของตัวแปรอิสระ 3 ตัว ต่อตัวแปรตามได้แก่ เวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ โดยใช้คลังข้อสอบวิชาคณิตศาสตร์ วัดความเข้าใจเรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 จำนวนข้อสอบ 244 ข้อ ทำการทดสอบผ่านโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ กับกลุ่มตัวอย่างนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2546 โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาปทุมธานี เขตที่ 1 และ 2 จำนวนนักเรียนกลุ่มตัวอย่าง 598 คน

สรุปผลการวิจัย

จากผลการวิเคราะห์ข้อมูลสามารถสรุปผลการวิจัยแบ่งออกได้เป็น 2 ประเด็น ตามวัตถุประสงค์ของการวิจัย ดังต่อไปนี้

1. สรุปผลการพัฒนาโปรแกรม

ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์

วัตถุประสงค์ของการวิจัยในประเด็นที่ 1 เพื่อพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ รูปแบบหลายขั้นตอนแบบแยกทางแปรผัน โดยการจัดชุดข้อสอบให้ตอบในแต่ละชั้นต่างกัน การกำหนดเวลาที่ใช้ในการตอบแต่ละข้อต่างกัน และการจัดรูปแบบของข้อสอบที่มีสื่อประสมและไม่มีสื่อประสม และศึกษาประสิทธิภาพของโปรแกรมการทดสอบ

โปรแกรมการทดสอบที่ผู้วิจัยพัฒนาขึ้นในครั้งนี้ ประกอบด้วยโปรแกรมสำหรับพิมพ์เพิ่ม-ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบ ซึ่งมี 2 รูปแบบ คือ แบบที่มีสื่อประสมและแบบที่ไม่มีสื่อประสม และโปรแกรมสำหรับการทดสอบ ซึ่งสามารถจัดโปรแกรมสำหรับการทดสอบได้ 18 โปรแกรม ซึ่งเกิดจากการจัดหมู่ของระดับของตัวแปรอิสระ 3 ตัว ได้แก่ ตัวแปรการจัดชุดข้อสอบให้ตอบในแต่ละชั้นต่างกัน 3 แบบ การกำหนดเวลาให้ตอบในแต่ละข้อต่างกัน 3 แบบ และการจัดรูปแบบของข้อสอบต่างกัน 2 รูปแบบ ซึ่งสามารถสรุปผลการวิจัยในประเด็นนี้ ได้ดังนี้

1.1 โปรแกรมสำหรับการพิมพ์เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบ

โปรแกรมสำหรับการพิมพ์เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบ มีความสามารถในการรองรับการทำงาน ดังต่อไปนี้

1.1.1 โปรแกรมมีขีดความสามารถในการทำงานบนระบบปฏิบัติการของวินโดวส์ 98 ขึ้นไป

1.1.2 โปรแกรมมีระบบป้องกันการเข้าสู่การพิมพ์เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบ โดยใช้รหัสผ่าน

1.1.3 โปรแกรมสามารถพิมพ์เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบได้ทั้งข้อสอบรูปแบบที่มีสื่อประสม และรูปแบบที่ไม่มีสื่อประสม ซึ่งข้อสอบที่จะนำมาพิมพ์ จะต้องมีคุณลักษณะดังนี้

1) ข้อสอบเป็นข้อสอบแบบเลือกตอบ ซึ่งสามารถกำหนดจำนวนตัวเลือกได้ตั้งแต่ 2 ถึง 5 ตัวเลือก

2) รูปแบบของตัวเลือกสามารถกำหนดชนิดของตัวเลือกได้ 3 ชนิด คือ ก, ข, ค,....; A,B,C,... และ 1,2,3...

3) ข้อสอบที่นำมาจัดพิมพ์ไม่ได้จำกัดจำนวนข้อ แต่จะต้องมีข้อสอบจำนวนมากจึงจะสามารถใช้กับโปรแกรมการทดสอบได้อย่างมีประสิทธิภาพ

4) ข้อสอบจะต้องผ่านการวิเคราะห์หาคุณภาพของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) แบบ 3 พารามิเตอร์

5) ค่าพารามิเตอร์ของข้อสอบจะต้องมีค่าพารามิเตอร์ประจำข้อ ดังนี้

- ค่าความยาก (b) อยู่ระหว่าง -3.0 ถึง 3.0
- ค่าอำนาจจำแนก (a) อยู่ระหว่าง 0.2 ถึง 2.5
- ค่าการเดา (c) มีค่าเท่ากับหรือน้อยกว่า 0.3

6) ข้อสอบจะเป็นวิชาใด หรือระดับใดก็ได้ โดยสามารถกำหนดชื่อของแบบทดสอบลงในช่องกำหนดชื่อได้

1.1.4 โปรแกรมสามารถกำหนดจำนวนข้อสูงสุดในการทดสอบแต่ละครั้งได้

1.1.5 โปรแกรมสามารถกำหนดจำนวนข้อในแต่ละชั้นของการสอบได้ 3 แบบ คือ ชั้นละ 1 ข้อ ชั้นละ 2 ข้อ และชั้นละ 3 ข้อ ซึ่งวิธีการคัดเลือกข้อสอบในแต่ละชั้นนั้น คัดเลือกจากค่าฟังก์ชันสารสนเทศของข้อสอบที่สูงที่สุด ณ ระดับความสามารถของผู้สอบในขณะทำข้อสอบข้อนั้น ๆ กรณีกำหนดข้อสอบชั้นละ 2 และ 3 ข้อ จะเลือกข้อสอบที่มีค่าฟังก์ชันสารสนเทศสูงสุดเท่ากัน หรือใกล้เคียงกันที่สุดจำนวน 2 และ 3 ข้อ ตามลำดับ

1.1.6 โปรแกรมสามารถกำหนดเวลาให้กับข้อสอบได้ 3 แบบ คือไม่กำหนดเวลาในการตอบข้อสอบแต่ละข้อ กำหนดเวลาในการตอบข้อสอบแต่ละข้อเท่ากันทุกข้อ และกำหนดเวลาในการตอบข้อสอบแต่ละข้อตามค่าความยากได้ 12 ค่า ซึ่งแบ่งค่าตามช่วงของความยากของข้อสอบ 12 ช่วง ความกว้างของแต่ละช่วงเท่ากับ 0.5 ซึ่งความยากจะมีค่าอยู่ระหว่าง -3.0 ถึง 3.0 การกำหนดเวลาสำหรับการวิจัยในครั้งนี้ ผู้วิจัยได้กำหนดเวลาที่ใช้ในการตอบในแต่ละข้อจากผลการวิจัยที่ได้จากการวิจัยนำร่อง ซึ่งช่วงแรกเป็นข้อสอบที่ง่ายที่สุดแต่ไม่มีข้อสอบในช่วงนี้จึงไม่ได้กำหนดเวลาไว้ การกำหนดเวลาที่ใช้ในการตอบแต่ละข้อจึงเริ่มกำหนดตั้งแต่ช่วงที่ 2 ดังนี้

- ช่วงความยาก -3.00 ถึง -2.51 ข้อสอบ 0 ข้อ เวลาข้อละ – วินาที
- ช่วงความยาก -2.50 ถึง -2.01 ข้อสอบ 1 ข้อ เวลาข้อละ 60 วินาที
- ช่วงความยาก -2.00 ถึง -1.51 ข้อสอบ 5 ข้อ เวลาข้อละ 65 วินาที
- ช่วงความยาก -1.50 ถึง -1.01 ข้อสอบ 7 ข้อ เวลาข้อละ 70 วินาที
- ช่วงความยาก -1.00 ถึง -0.51 ข้อสอบ 21 ข้อ เวลาข้อละ 77 วินาที
- ช่วงความยาก -0.50 ถึง -0.01 ข้อสอบ 37 ข้อ เวลาข้อละ 82 วินาที
- ช่วงความยาก 0.00 ถึง 0.50 ข้อสอบ 45 ข้อ เวลาข้อละ 90 วินาที
- ช่วงความยาก 0.51 ถึง 1.00 ข้อสอบ 40 ข้อ เวลาข้อละ 95 วินาที
- ช่วงความยาก 1.01 ถึง 1.50 ข้อสอบ 42 ข้อ เวลาข้อละ 99 วินาที
- ช่วงความยาก 1.51 ถึง 2.00 ข้อสอบ 24 ข้อ เวลาข้อละ 103 วินาที
- ช่วงความยาก 2.01 ถึง 2.50 ข้อสอบ 18 ข้อ เวลาข้อละ 107 วินาที
- ช่วงความยาก 2.51 ถึง 3.00 ข้อสอบ 4 ข้อ เวลาข้อละ 112 วินาที
- เวลาโดยเฉลี่ยเท่ากันทุกข้อจำนวนข้อสอบ 244 ข้อ เวลาข้อละ 90 วินาที

1.1.7 โปรแกรมสามารถใช้สื่อประสมภาพนิ่ง ภาพวีว หรือภาพทิวทัศน์ และเสียงเพลงประกอบ ร่วมกับการนำเสนอข้อสอบในแต่ละข้อได้

1.1.8 โปรแกรมสามารถเพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบได้โดยวิธีการพิมพ์ และโดยวิธีการใช้ไฟล์รูปภาพได้

1.1.9 โปรแกรมสามารถเลือกสีพื้น สีตัวอักษร ขนาด และลักษณะของอักษรได้

1.2 โปรแกรมสำหรับการทำการทดสอบ

โปรแกรมสำหรับการทำการทดสอบ มีความสามารถในการทำงานได้ดังนี้

1.2.1 โปรแกรมสามารถทำงานบนระบบปฏิบัติการของวินโดว 98 ขึ้นไป

1.2.2 โปรแกรมสามารถให้ผู้สอบพิมพ์ข้อมูลส่วนตัวได้แก่ ชื่อ สกุล ชั้นเรียน โรงเรียน วันที่สอบ และเวลาเริ่มต้นการทดสอบได้

1.2.3 โปรแกรมการทดสอบสามารถทำการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ รูปแบบแยกทางแปรผัน เริ่มต้นสอบด้วยข้อสอบที่มีความยากง่ายปานกลาง ประเมินค่าความสามารถของผู้สอบด้วยวิธีของเบส์ คัดเลือกข้อสอบข้อต่อไปด้วยการเลือกข้อที่มีค่าฟังก์ชันสารสนเทศของข้อสอบที่สูงที่สุด ณ ระดับความสามารถขณะนั้นของผู้สอบ และยุติการทดสอบเมื่อค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบน้อยกว่าหรือเท่ากับ 0.3 หรือทำข้อสอบครบตามจำนวนข้อที่กำหนด

1.2.4 โปรแกรมสามารถทำการทดสอบด้วยข้อสอบทั้งรูปแบบที่มีสื่อประสมและรูปแบบข้อสอบที่ไม่มีสื่อประสมได้

1.2.5 โปรแกรมสามารถแสดงเวลาที่ใช้ในการตอบข้อสอบแต่ละข้อได้โดยเวลาจะบอกเป็น 3 ช่วง ได้แก่ ช่วงแรกช่องแสดงเวลาเป็นสีเขียวแสดงว่ายังมีเวลาในการทำข้อสอบมากพอในการทำข้อสอบข้อนั้น ช่วงที่สองช่องแสดงเวลาเป็นสีเหลืองแสดงว่าเหลือเวลาทำข้อสอบข้อนั้น 30 วินาที และช่วงเวลาสุดท้ายช่องแสดงเวลาเป็นสีแดง แสดงว่าเหลือเวลาทำข้อสอบข้อนั้นเพียง 10 วินาที

1.2.6 โปรแกรมสามารถเลือกข้อสอบ หรือชุดของข้อสอบได้ตรงกับความสามารถของผู้สอบ โดยเลือกข้อสอบข้อต่อไปจากค่าฟังก์ชันสารสนเทศของข้อสอบที่สูงที่สุด ณ ระดับความสามารถของผู้สอบในขณะนั้น

1.2.7 โปรแกรมสามารถแสดงค่าความสามารถของผู้สอบ และค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบ เมื่อผู้สอบทำข้อสอบเสร็จในแต่ละข้อ

1.2.8 โปรแกรมสามารถหยุดเวลาของการทำแบบทดสอบเมื่อหมดเวลาในการตอบข้อสอบในแต่ละข้อ และเริ่มจับเวลาเพิ่มเมื่อผู้สอบเลือกทำข้อสอบข้อต่อไป

1.2.9 เมื่อการทดสอบยุติลงตามเกณฑ์ที่กำหนด โปรแกรมสามารถให้ผู้สอบทำการบันทึกผลการสอบ และโปรแกรมสามารถรายงานผลการสอบได้ทันที

1.2.10 โปรแกรมสามารถรายงานผลการสอบได้ดังนี้

- ข้อมูลส่วนตัวของผู้สอบ
- รูปแบบหรือชนิดของแบบทดสอบ
- เวลาเริ่มต้นการทดสอบ และเวลารวมทั้งหมดที่ใช้ ในการทดสอบ
- จำนวนข้อที่ตอบถูก และจำนวนข้อสอบทั้งหมดที่ใช้ในการทดสอบ
- ตำแหน่งเปอร์เซ็นต์ไทล์ของผู้สอบ

- ระดับความสามารถและความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบ
- ค่าฟังก์ชันสารสนเทศของข้อสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ

1.2.11 ผลการทดสอบสามารถพิมพ์ออกมาทางเครื่องพิมพ์ได้

1.3 ผลการประเมินคุณภาพของโปรแกรม

การประเมินคุณภาพของโปรแกรมการทดสอบในครั้งนี้ ผู้วิจัยทำการประเมินโปรแกรมโดยพิจารณาจากผู้ประเมิน 2 กลุ่ม ได้แก่ กลุ่มของผู้เชี่ยวชาญด้านการวัดผลและผู้เชี่ยวชาญด้านการใช้โปรแกรมการทดสอบ เพื่อตรวจสอบความถูกต้องแม่นยำของการประมวลผลที่ได้จากการใช้โปรแกรมการทดสอบ และกลุ่มผู้ใช้โปรแกรมการทดสอบ ซึ่งประกอบด้วยผู้ดำเนินการสอบ เพื่อประเมินประสิทธิภาพของโปรแกรมการทดสอบด้านความสะดวกรวดเร็วในการใช้โปรแกรม ความแม่นยำของโปรแกรม และความผิดพลาดของโปรแกรม และนักเรียนกลุ่มตัวอย่างที่ทำการทดสอบ เพื่อประเมินคุณภาพโปรแกรมด้านความพึงพอใจในการทดสอบจากการทดสอบโดยการใช้โปรแกรม ซึ่งผลการประเมินคุณภาพของโปรแกรมสามารถสรุป ได้ดังนี้

1.3.1 การประเมินคุณภาพของโปรแกรมโดยกลุ่มผู้เชี่ยวชาญ พบว่า

- 1) โปรแกรมสามารถทำงานภายใต้เงื่อนไขของตัวแปรที่ผู้วิจัยกำหนด ได้แก่ การจัดชุดข้อสอบให้ตอบในแต่ละชั้น การกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อ และการกำหนดรูปแบบของข้อสอบแบบมีสื่อประสม และแบบไม่มีสื่อประสม ได้ถูกต้อง
- 2) โปรแกรมสามารถทำการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์โดยใช้รูปแบบแยกทางแปรผัน ประมาณค่าความสามารถของผู้สอบโดยวิธีของเบส์ คัดเลือกข้อสอบข้อต่อไปจากค่าฟังก์ชันสารสนเทศของข้อสอบที่สูงที่สุด และยุติการทดสอบเมื่อค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบเท่ากับหรือที่ต่ำกว่า 0.3 หรือผู้สอบทำการทดสอบครบ 36 ข้อได้อย่างถูกต้องและรวดเร็ว
- 3) โปรแกรมสามารถบันทึกเวลาที่ใช้ในการสอบ จำนวนข้อที่ตอบถูก จำนวนข้อที่ใช้ในการทดสอบ และบอกตำแหน่งเปอร์เซ็นต์ไทล์ของผู้สอบได้ถูกต้อง
- 4) โปรแกรมสามารถคำนวณประมาณค่าความสามารถของผู้สอบ ค่าความคลาดเคลื่อนในการประมาณค่าความสามารถ ค่าฟังก์ชันสารสนเทศของข้อสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบได้ถูกต้องและรวดเร็ว

1.3.2 การประเมินคุณภาพ โดยกลุ่มผู้ใช้โปรแกรมการทดสอบ

- 1) การประเมินคุณภาพของโปรแกรมโดยผู้ดำเนินการสอบ พบว่า
 - 1.1) การติดตั้งโปรแกรมสามารถทำได้ง่าย การใช้โปรแกรมในการทดสอบมีความสะดวกรวดเร็ว
 - 1.2) การดำเนินการสอบใช้เวลาน้อยลง และสามารถรายงานผลการสอบได้ทันที

1.3) โปรแกรมสามารถคำนวณประมาณค่าความสามารถของผู้สอบ ความคลาดเคลื่อนในการประมาณค่าความสามารถ ค่าฟังก์ชันสารสนเทศของข้อสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ ได้ถูกต้องแม่นยำ เมื่อทำการตรวจสอบกับการคำนวณด้วยมือ

1.4) โปรแกรมสามารถบันทึกเวลาที่ใช้ในการสอบ จำนวนข้อที่ตอบถูก และจำนวนข้อที่ทำการทดสอบทั้งหมด ได้ถูกต้องครบถ้วน

1.5) ไม่พบความผิดพลาดในการทำงานของโปรแกรม และโปรแกรมมีความทนทานต่อการใช้งาน ถึงแม้ว่าผู้ใช้งานจะใช้โปรแกรมลัดขั้นตอนนี้ก็ตาม

2) การประเมินคุณภาพของโปรแกรมโดยนักเรียนกลุ่มตัวอย่าง

การประเมินโปรแกรมโดยนักเรียนกลุ่มตัวอย่างนั้น เป็นการประเมินคุณภาพของโปรแกรมด้านความพึงพอใจในการสอบ พบว่า โปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อนำไปใช้สอบกับนักเรียนกลุ่มตัวอย่างแล้วนักเรียนกลุ่มตัวอย่างมีความพึงพอใจในการสอบในระดับมาก เมื่อเปรียบเทียบความพึงพอใจในการสอบระหว่างรูปแบบของข้อสอบที่มีสื่อประสมกับไม่มีสื่อประสม พบว่าผู้สอบมีความพึงพอใจในการสอบจากรูปแบบของ ข้อสอบที่มีสื่อประสมสูงกว่ารูปแบบของข้อสอบที่ไม่มีสื่อประสม ดังนั้นโปรแกรมการทดสอบที่ผู้วิจัยพัฒนาขึ้นสามารถสร้างความพึงพอใจให้กับผู้สอบได้เป็นอย่างดี

2. สรุปผลการวิจัยอิทธิพลของตัวแปรอิสระที่มีต่อตัวแปรตาม

วัตถุประสงค์ของการวิจัยในประเด็นที่ 2 เพื่อเปรียบเทียบเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อมีการจัดชุดข้อสอบให้ตอบในแต่ละชั้นต่างกัน การกำหนดเวลาให้ตอบในแต่ละข้อต่างกัน และเมื่อมีการจัดรูปแบบของข้อสอบที่ไม่มีสื่อประสมกับรูปแบบที่มีสื่อประสม ผลการวิเคราะห์ข้อมูลสามารถสรุปผลได้ดังนี้

2.1 เมื่อจัดชุดข้อสอบให้ตอบในแต่ละชั้นต่างกัน ไม่ทำให้เวลาที่ใช้ในการทดสอบ และจำนวนของแบบทดสอบต่างกัน แต่ทำให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบต่างกัน โดยการจัดชุดข้อสอบให้ตอบชั้นละ 3 ข้อ ทำให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงที่สุด รองลงมาคือจัดชุดข้อสอบให้ตอบชั้นละ 2 ข้อ และจัดชุดข้อสอบให้ตอบชั้นละ 1 ข้อ ให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบต่ำที่สุด

2.2 เมื่อกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อต่างกัน ทำให้เวลาที่ใช้ในการทดสอบต่างกัน โดยที่การกำหนดเวลาในการตอบข้อสอบในแต่ละข้อตามค่าความยากของข้อสอบ และการกำหนดเวลาในการตอบแต่ละข้อเฉลี่ยเท่ากันทุกข้อ ต่างจากการไม่กำหนดเวลาในการตอบข้อสอบในแต่ละข้อ ซึ่งการกำหนดเวลาในการตอบข้อสอบในแต่ละข้อตามค่าความยากของข้อสอบใช้เวลาสอบน้อยที่สุด ส่วนการไม่กำหนดเวลาในการตอบข้อสอบในแต่ละข้อจะใช้เวลาในการทดสอบมากที่สุด ส่วนจำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบไม่ต่างกัน

2.3 เมื่อจัดรูปแบบของข้อสอบที่ไม่มีสื่อประสมกับรูปแบบของข้อสอบที่มีสื่อประสม ประกอบการทดสอบทำให้เวลาที่ใช้ในการทดสอบต่างกัน โดยการจัดรูปแบบข้อสอบที่ไม่มีสื่อประสมใช้เวลาในการสอบมากกว่ารูปแบบที่มีสื่อประสม ส่วนจำนวนข้อของแบบทดสอบและค่าฟังก์ชันสารสนเทศของแบบทดสอบไม่ต่างกัน

ถ้าพิจารณาปฏิสัมพันธ์ของตัวแปรอิสระหรือการร่วมกันส่งผลของตัวแปรอิสระ 2 ตัว และ ตัวแปรอิสระ 3 ตัว ต่อเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ พบว่า

2.4 เมื่อจัดชุดข้อสอบให้ตอบในแต่ละขั้นกับการกำหนดเวลาที่ใช้ในการตอบแต่ละข้อ ไม่ทำให้เวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแตกต่างกัน

2.5 เมื่อจัดชุดข้อสอบให้ตอบในแต่ละขั้นกับรูปแบบของข้อสอบ ไม่ทำให้เวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแตกต่างกัน

2.6 เมื่อกำหนดเวลาที่ใช้ในการตอบแต่ละข้อกับรูปแบบของข้อสอบ ไม่ทำให้เวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแตกต่างกัน

2.7 เมื่อจัดชุดข้อสอบให้ตอบในแต่ละขั้นกับการกำหนดเวลาที่ใช้ในการตอบแต่ละข้อและรูปแบบของข้อสอบ ไม่ทำให้เวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบแตกต่างกัน

อภิปรายผลการวิจัย

1. ผลการพัฒนาโปรแกรม และผลของการประเมินคุณภาพของโปรแกรม

1.1 ผลของการพัฒนาโปรแกรม

การพัฒนาโปรแกรมการทดสอบในครั้งนี้ ผู้วิจัยพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ประกอบด้วยโปรแกรมสำหรับพิมพ์เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบ และโปรแกรมสำหรับการทดสอบซึ่งมีทั้งหมด 18 รูปแบบ ที่เกิดจากการจัดหมวดหมู่ของตัวแปรอิสระ การพัฒนาโปรแกรมการทดสอบ อภิปรายผล ได้ดังต่อไปนี้

1.1.1 ลักษณะของโปรแกรมสำหรับการพิมพ์เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบ เป็นโปรแกรมที่ทำงานบนระบบปฏิบัติการของวินโดวส์ 98 ขึ้นไป มีรหัสผ่านสำหรับการป้องกันการเข้าสู่ตัวข้อสอบ โปรแกรมใช้ได้กับรูปแบบของข้อสอบที่ไม่มีสื่อประสม และรูปแบบของข้อสอบที่มีสื่อประสม ข้อสอบที่ใช้กับโปรแกรมจะต้องเป็นข้อสอบเลือกตอบจำนวนตัวเลือก 2 ถึง 5 ตัวเลือก ชนิดของตัวเลือกเป็นแบบ ก,ข,ค,..., A,B,C,..., หรือแบบ 1,2,3,... โปรแกรมจะต้องมีข้อสอบจำนวนมาก โดยผ่านการวิเคราะห์คุณภาพ ตามทฤษฎีการตอบสนองของข้อสอบแบบ 3 พารามิเตอร์ ค่าความยากอยู่ระหว่าง -3.0 ถึง 3.0 ค่าอำนาจจำแนกอยู่ระหว่าง 0.2 ถึง 2.5 และค่าการเดามีค่าเท่ากับหรือน้อยกว่า 0.3 จะเป็นข้อสอบวิชาใดหรือระดับใดก็ได้ โปรแกรมสามารถกำหนดคุณลักษณะของการทดสอบได้แก่ กำหนดจำนวนข้อสูงที่สุดในการทดสอบแต่ละครั้ง กำหนดจำนวนข้อที่ใช้ในการตอบในแต่ละขั้นได้ 3 แบบ

กำหนดเวลาที่ใช้ในการตอบแต่ละข้อได้ 3 รูปแบบ จำนวน 12 คำ การเพิ่ม ลด หรือเปลี่ยนแปลงแก้ไข ข้อสอบสามารถทำได้โดยการพิมพ์หรือวิธีการใช้ไฟล์รูปภาพในการนำข้อสอบเข้าในคลัง และโปรแกรมสามารถเลือกสีพื้น สีตัวอักษร ขนาดและลักษณะของตัวอักษรได้

1.1.2 โปรแกรมการทดสอบประกอบด้วยโปรแกรมย่อย 18 โปรแกรม ซึ่งเกิดจากการจัดหมวดหมู่ของตัวแปรอิสระ เป็นโปรแกรมที่ทำงานบนระบบปฏิบัติการของวินโดวส์ 98 ขึ้นไป โปรแกรมการทดสอบเริ่มต้นจากพิมพ์ข้อมูลส่วนตัวของผู้สอบ ได้แก่ ชื่อ สกุล ชั้นเรียน โรงเรียน วันที่ สอบ และเวลาเริ่มต้นการทดสอบ จึงเริ่มทำการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วย คอมพิวเตอร์รูปแบบแยกทางแปรผัน เริ่มต้นการสอบด้วยข้อสอบที่มีความยากปานกลาง ประมาณค่าความสามารถของผู้สอบโดยวิธีของเบส์ คัดเลือกข้อสอบข้อต่อไปจากค่าฟังก์ชันสารสนเทศของข้อสอบที่สูงที่สุด ยุติการทดสอบเมื่อค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบเท่ากับหรือน้อยกว่า 0.3 หรือทำข้อสอบครบตามจำนวนที่กำหนด โปรแกรมสามารถแสดงบอกเวลาที่ใช้ในการตอบแต่ละข้อโดยบอกเวลาเป็น 3 ช่วง แสดงค่าความสามารถของผู้สอบและค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบเมื่อทำข้อสอบเสร็จในแต่ละข้อ โปรแกรมสามารถหยุดเวลาของการทำแบบทดสอบเมื่อหมดเวลาในการตอบข้อสอบในแต่ละข้อ และเริ่มจับเวลาเพิ่มเมื่อผู้สอบเลือกทำข้อสอบข้อต่อไป เมื่อการทดสอบยุติลงตามเกณฑ์ โปรแกรมจะทำการบันทึกผลการทดสอบและรายงานผลการสอบทันที โดยที่ผลจากการทดสอบประกอบด้วย ผลข้อมูลส่วนตัวของผู้สอบ รูปแบบหรือชนิดของแบบทดสอบ เวลาเริ่มต้นสอบ และเวลารวมทั้งหมดที่ใช้ในการสอบ จำนวนข้อที่ตอบถูกและจำนวนข้อทั้งหมดที่ทำ ค่าความสามารถและค่าความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบ ค่าฟังก์ชันสารสนเทศของข้อสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ และผลการสอบทั้งหมดสามารถพิมพ์ออกมาทางเครื่องพิมพ์ได้

1.2 ผลการประเมินคุณภาพของโปรแกรม

การประเมินคุณภาพของโปรแกรมการทดสอบนั้น ผู้วิจัยพิจารณาจากผู้ประเมินคุณภาพ 2 กลุ่ม ซึ่งกลุ่มแรกได้แก่ผู้เชี่ยวชาญด้านการวัดผลและผู้เชี่ยวชาญด้านโปรแกรมการทดสอบ ส่วนผู้ประเมินกลุ่มที่สองได้แก่กลุ่มผู้ใช้โปรแกรมการทดสอบ การประเมินผลสามารถอภิปรายผลได้ดังนี้

1.2.1 การประเมินคุณภาพ โดยกลุ่มผู้เชี่ยวชาญ

การประเมินคุณภาพ โดยกลุ่มผู้เชี่ยวชาญด้านการวัดผลและผู้เชี่ยวชาญด้านโปรแกรมการทดสอบ ผู้ประเมินกลุ่มนี้จะพิจารณาประสิทธิภาพด้านความแม่นยำของโปรแกรม ผลจากการประเมินพบว่า โปรแกรมการทดสอบสามารถทำงานภายใต้เงื่อนไขของตัวแปรที่ผู้วิจัยกำหนด ได้แก่ การจัดชุดข้อสอบให้ตอบในแต่ละชั้น การกำหนดเวลาที่ใช้ในการตอบแต่ละข้อ และการกำหนดรูปแบบของข้อสอบที่ไม่มีสื่อประสมกับรูปแบบที่มีสื่อประสม โปรแกรมคำนวณค่าความสามารถของผู้สอบด้วยวิธีของเบส์ คัดเลือกข้อสอบจากค่าฟังก์ชันสารสนเทศของข้อสอบที่สูงที่สุด ยุติการทดสอบเมื่อค่าความคลาดเคลื่อนในการประมาณค่าเท่ากับหรือน้อยกว่า 0.3 หรือผู้สอบทำข้อสอบครบ 36 ข้อ ซึ่งโปรแกรมสามารถทำงานตามเงื่อนไขดังกล่าวได้อย่างมีประสิทธิภาพ ส่วนการคำนวณผลจากการทดสอบนั้นปรากฏว่าโปรแกรมคำนวณประมาณค่าความสามารถของผู้สอบ ความคลาดเคลื่อนในการประมาณค่า

ความสามารถ ค่าฟังก์ชันสารสนเทศของข้อสอบ ค่าฟังก์ชันสารสนเทศของแบบทดสอบ เวลาที่ใช้ในการสอบ จำนวนข้อที่ตอบถูกและจำนวนข้อที่ทำการทดสอบทั้งหมด ได้ถูกต้องแม่นยำ และมีความรวดเร็วมากในการคำนวณค่าต่าง ๆ ดังกล่าว เมื่อเปรียบเทียบกับ การคำนวณด้วยมือ จากผลการประเมินจึงอนุมานได้ว่า โปรแกรมที่ผู้วิจัยพัฒนาขึ้นมีคุณภาพ และมีประสิทธิภาพในการทำงาน

1.2.2 การประเมินคุณภาพ โดยกลุ่มผู้ใช้โปรแกรมการทดสอบ

1) กลุ่มผู้ดำเนินการสอบ ประเมินประสิทธิภาพด้านความสะดวกรวดเร็วในการใช้โปรแกรม ความแม่นยำของโปรแกรม และความทนทานต่อความผิดพลาดของโปรแกรม พบว่าการติดตั้งโปรแกรมสามารถทำได้ง่าย การใช้โปรแกรมในการทดสอบมีความสะดวกรวดเร็ว ใช้เวลาในการดำเนินการสอบน้อยลง และสามารถรายงานผลการสอบได้ทันที โปรแกรมสามารถคำนวณประมาณค่าความสามารถของผู้สอบ ความคลาดเคลื่อนในการประมาณค่าความสามารถ ค่าฟังก์ชันสารสนเทศของข้อสอบ ค่าฟังก์ชันสารสนเทศของแบบทดสอบ ได้ถูกต้องแม่นยำ โปรแกรมสามารถบันทึกเวลาที่ใช้ในการสอบ จำนวนข้อที่ตอบถูกและจำนวนข้อที่ทำการทดสอบทั้งหมด ได้ถูกต้องครบถ้วน และไม่พบความผิดพลาดในการทำงานของโปรแกรม

2) กลุ่มนักเรียนตัวอย่างที่ทำการทดสอบ ทำการประเมินความพึงพอใจในการทดสอบพบว่า ผู้สอบมีความพึงพอใจในการสอบในระดับมาก ซึ่งสอดคล้องกับผลการวิจัยของ ดริงใจ พูนผลอำนาย (2534) และสายชล อบทม (2539) ที่พบว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ สร้างความดึงดูดใจให้ผู้สอบสนใจในการทำข้อสอบเพิ่มขึ้น ผู้สอบมีความกระตือรือร้นที่จะทำการทดสอบ และมีทัศนคติที่ดีในการทดสอบ และยังพบอีกว่า ความพึงพอใจของผู้สอบต่อโปรแกรมการทดสอบรูปแบบของข้อสอบที่มีสื่อประสมสูงกว่าโปรแกรมการทดสอบรูปแบบที่ไม่มีสื่อประสม ซึ่งสอดคล้องกับแนวคิดของเมเจอร์ (Meijer and Michael.1999) ; ริงสรรค์ มณีเล็ก (2540) และสตรีทแมน และเอจเจน (Straetmens and Eggen.1998) ที่กล่าวไว้ว่า น่าจะนำภาพกราฟฟิก เสียง ภาพเคลื่อนไหว หรือสื่อประสมมาจัดหารูปแบบของข้อสอบ เพื่อใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ซึ่งจะทำให้การทดสอบน่าสนใจ ช่วยผ่อนคลาย ในขณะที่ทำการทดสอบ ไม่ทำให้ผู้สอบเกิดความเบื่อหน่าย ผู้สอบจะมีความพึงพอใจในการสอบมากยิ่งขึ้น

2. ผลของตัวแปรอิสระที่มีอิทธิพลต่อตัวแปรตาม

การวิจัยในครั้งนี้ ผู้วิจัยได้ตั้งสมมติฐานตามที่ผู้วิจัยคาดว่าจะมีอิทธิพลของตัวแปรอิสระต่อตัวแปรตามไว้ 5 ประเด็น จากการศึกษาเอกสาร และงานวิจัยที่เกี่ยวข้องที่ผ่านมา ผลจากการวิจัยสามารถอภิปรายผลตามสมมติฐานของการวิจัยได้ ดังนี้

2.1 เมื่อจัดชุดข้อสอบให้ตอบในแต่ละชั้นต่างกันในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ พบว่า ค่าฟังก์ชันสารสนเทศของแบบทดสอบต่างกัน ซึ่งผลจากการวิจัยนี้ตรงตามสมมติฐานที่ตั้งไว้ และตรงตามเหตุผลที่ผู้วิจัยเสนอไว้ ในสมมติฐานที่ว่าถ้าจัดชุดข้อสอบให้ตอบในแต่ละชั้นเพิ่มมากขึ้น และมีจำนวนข้อสอบไม่มากเกินไปในแต่ละชั้น จะส่งผลให้ค่าฟังก์ชัน สารสนเทศของแบบทดสอบสูงขึ้น ซึ่งผลจากการวิจัยเมื่อจัดชุดข้อสอบให้ตอบชั้นละ 3 ข้อ ทำให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงที่สุด รองลงมาคือชั้นละ 2 ข้อ และชั้นละ 1 ข้อ

ตามลำดับ ส่วนเวลาที่ใช้ในการสอบและจำนวนข้อของแบบทดสอบไม่ต่างกัน ซึ่งสอดคล้องกับทฤษฎีของกุลลิคเซน ลอร์ดและโนวิด (ต่าย เชียงฉิ.2534; อ้างอิงมาจาก Gulliksen.1967; Lord and Novick.1968) เสนอไว้ว่าโดยทฤษฎีแล้วจำนวนข้อสอบยิ่งมากขึ้นค่าความเชื่อมั่นก็ยิ่งสูงขึ้น แต่ถ้าข้อสอบมากเกินไปก็จะทำให้เสียเวลาในการสอบ ผู้สอบเกิดความเบื่อหน่าย และจะทำให้ความเชื่อมั่นลดลงได้ จากแนวคิดของทฤษฎีดังกล่าวทำให้ผู้วิจัยนำมาเป็นเหตุผลในการเพิ่มจำนวนข้อสอบในแต่ละชั้น จึงทำให้ค่าความเชื่อมั่นในแต่ละชั้นของการทดสอบมีค่าสูง และสามารถคัดเลือกชุดของข้อสอบในชั้นต่อไปได้ถูกต้องตรงกับความสามารถของผู้สอบยิ่งขึ้น ซึ่งสอดคล้องกับผลการวิจัยของนันทิยา ฟิงคำ (2531) ที่พบว่า ถ้าค่าฟังก์ชันสารสนเทศของข้อสอบในแต่ละชั้นมีความน่าเชื่อถือมากยิ่งขึ้น ค่าความคลาดเคลื่อนในการประมาณค่าความสามารถจะต่ำลง การทดสอบจะยุติเร็วขึ้น ด้วยเหตุนี้การจัดชุดข้อสอบให้ตอบชั้นละ 3 ข้อ และชั้นละ 2 ข้อ จึงทำให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงขึ้น แต่ไม่เพิ่มจำนวนข้อของแบบทดสอบ และเวลาที่ใช้ในการสอบ ดังผลการวิจัยที่กล่าวไว้ข้างต้น

2.2 เมื่อกำหนดเวลาที่ใช้ในการตอบแต่ละข้อต่างกันในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบแล้ว พบว่า จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบไม่ต่างกัน ซึ่งไม่เป็นไปตามสมมติฐานที่ตั้งไว้ แต่มีความสอดคล้องกับสมมติฐานที่ตั้งไว้อยู่บางประเด็นที่เกี่ยวข้องกันได้แก่ เวลาที่ใช้ในการทดสอบต่างกัน ทั้งนี้ เนื่องจากว่าคุณภาพของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบนั้นจะต้องคำนึงถึงเรื่องเวลาที่ใช้ในการสอบ จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ ผลจากการวิจัยในครั้งนี้ถึงแม้ว่า การกำหนดเวลาที่ใช้ในการตอบแต่ละข้อจะไม่ทำให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงขึ้น และจำนวนของแบบทดสอบก็ไม่เพิ่มขึ้น แต่ใช้เวลาในการทดสอบน้อยลง แสดงว่าการกำหนดเวลาที่ใช้ในการตอบแต่ละข้อมีผลต่อคุณภาพของแบบทดสอบ ซึ่งสอดคล้องกับแนวความคิดของ สตีเวน (Steven. <http://ercir.syr.edu/plweb/cgi/fastweb>) ที่กล่าวว่า องค์ประกอบสำคัญอีกประการหนึ่งที่ควรคำนึงถึงในการทดสอบแบบปรับเหมาะคือ การกำหนดระยะเวลาที่ใช้สอบ เพราะการทดสอบด้วยคอมพิวเตอร์สามารถกำหนดเวลาให้กับข้อสอบตามที่ต้องการได้ การกำหนดเวลาจะทำให้ประหยัดเวลาของการทดสอบลง และรังสรรค์ มณีเล็ก (2540) เสนอแนะว่า ควรศึกษาการกำหนดเวลาที่ใช้ในการตอบข้อสอบแต่ละข้อและศึกษาผลของการกำหนดเวลานั้นด้วย ซึ่งผลจากการวิจัยครั้งนี้ พบว่าการกำหนดเวลาที่ใช้ในการตอบเฉลี่ยเท่ากันทุกข้อ และการกำหนดเวลาที่ใช้ในการตอบตามค่าความยากของข้อสอบ ทำให้ลดเวลาที่ใช้ในการทดสอบลง ซึ่งสอดคล้องกับผลการวิจัยของ ศิริชัย กาญจนวาสี (2538) และผลการวิจัยของไวส์ โมรีโน และคนอื่นๆ (ต่าย เชียงฉิ; อ้างอิงจาก Weiss.1982, Morino and others.1984) ที่พบว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ทำการประมาณค่าฟังก์ชันสารสนเทศของแบบทดสอบได้เที่ยงตรงและน่าเชื่อถือ โดยใช้จำนวนข้อสอบและเวลาที่ใช้ในการทดสอบน้อยลง

2.3 เมื่อจัดรูปแบบของข้อสอบที่ไม่มีสื่อประสมกับรูปแบบของข้อสอบที่มีสื่อประสมให้กับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แล้ว พบว่าเวลาที่ใช้ในการทดสอบต่างกัน แต่จำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบไม่ต่างกัน ผลการวิจัยนี้ไม่ตรงตามสมมติฐานการวิจัยที่ตั้งไว้ว่า เมื่อจัดรูปแบบของข้อสอบที่มีสื่อประสมค่าฟังก์ชันสารสนเทศ

ของแบบทดสอบจะสูงกว่ารูปแบบของข้อสอบที่ไม่มีสื่อประสม แต่ถึงอย่างไรก็ตามถึงแม้ว่าจำนวนข้อของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบจะไม่แตกต่างกันก็ตาม แต่เวลาที่ใช้ในการทดสอบด้วยรูปแบบของข้อสอบที่มีสื่อประสมน้อยกว่าเวลาการทดสอบด้วยรูปแบบของข้อสอบที่ไม่มีสื่อประสม แสดงว่าการจัดรูปแบบของข้อสอบที่มีสื่อประสม มีประสิทธิภาพสูงกว่า เนื่องจากประหยัดเวลาที่ใช้ในการทดสอบได้มากกว่า ทั้งนี้เนื่องจากว่า การจัดรูปแบบของข้อสอบที่มีสื่อประสมจะสร้างความสนใจของผู้สอบ ชวนให้ผู้สอบอยากติดตามว่าข้อสอบข้อต่อไปจะมีสื่อประสมรูปแบบใด ทำให้ผู้สอบตั้งใจในการทำข้อสอบมากขึ้น ซึ่งสอดคล้องกับแนวความคิดของ เมเยอร์ (Meijer and Michael.1999) ; รังสรรค์ มณีเล็ก(2540) สตรีทแมน และเอจเจน (Straetmens and Eggen.1998) ที่กล่าวว่า ให้นำคอมพิวเตอร์มาใช้ในการทดสอบจะได้ข้อสอบตรงกับความต้องการและสามารถนำเสนอวิธีการทดสอบที่รวมเอาภาพกราฟฟิค เสียง ภาพเคลื่อนไหว หรือสื่อประสม และเนื้อหาสาระของข้อสอบเข้าไว้ด้วยกันจะทำให้การทดสอบน่าสนใจ ผู้สอบมีความกระตือรือร้นที่จะทำการทดสอบ จึงทำให้ผู้สอบใช้เวลาในการทดสอบน้อยลง

สำหรับปฏิสัมพันธ์ระหว่างตัวแปรนั้นผู้วิจัยไม่พบว่ามีเอกสาร หรือผลงานวิจัยใดกล่าวถึงอิทธิพลร่วมของตัวแปรที่ผู้วิจัยศึกษาในครั้งนี้อาไว้ แต่ผู้วิจัยได้ตั้งสมมติฐานไว้เพื่อเป็นการพิสูจน์ว่าจะมีอิทธิพลร่วมระหว่างตัวแปรอิสระต่อตัวแปรตามหรือไม่อย่างไร ดังนั้น ผู้วิจัยจึงตั้งสมมติฐานไว้เพียง 2 ข้อ ดังกล่าว ซึ่งสามารถอภิปรายผลตามสมมติฐานได้ดังนี้

เมื่อพิจารณาปฏิสัมพันธ์หรืออิทธิพลร่วมระหว่างตัวแปร 2 ตัว ซึ่งประกอบด้วย การจัดชุดข้อสอบให้ตอบในแต่ละขั้นกับการกำหนดเวลาที่ใช้ในการตอบแต่ละข้อ การจัดชุดข้อสอบให้ตอบในแต่ละขั้นกับการจัดรูปแบบของข้อสอบ การกำหนดเวลาที่ใช้ในการตอบในแต่ละข้อกับรูปแบบของข้อสอบ และอิทธิพลร่วมระหว่างตัวแปร 3 ตัวแปร ได้แก่ การจัดชุดข้อสอบให้ตอบในแต่ละขั้นกับการกำหนดเวลาที่ใช้ในการตอบแต่ละข้อและรูปแบบของข้อสอบ ไม่ทำให้เวลาที่ใช้ในการทดสอบ จำนวนของแบบทดสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบต่างกัน ซึ่งไม่เป็นไปตามสมมติฐานการวิจัยข้อที่ 4 และสมมติฐานการวิจัยข้อที่ 5 จากผลการวิจัยดังกล่าว จึงพออภิปรายผลได้ว่าปฏิสัมพันธ์ร่วมระหว่างตัวแปรอิสระที่ผู้วิจัยศึกษาในครั้งนี้อาไม่มีผลต่อตัวแปรตาม ทั้งนี้อาจเนื่องมาจากตัวแปรที่ผู้วิจัยศึกษามีความเป็นอิสระต่อกัน และมีอิทธิพลต่อตัวแปรตามแตกต่างกัน กล่าวคือ การจัดชุดข้อสอบให้ตอบในแต่ละขั้น มีผลต่อค่าฟังก์ชันสารสนเทศของแบบทดสอบ ส่วนการกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อและการจัดรูปแบบของข้อสอบมีผลต่อเวลาที่ใช้ในการทดสอบ

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลการทำวิจัยไปใช้

การพิจารณาตัวแปรเพื่อที่จะนำไปใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ ให้เกิดประโยชน์ และมีประสิทธิภาพสูงสุดนั้น ผู้วิจัยขอเสนอผลของตัวแปรที่ได้จากข้อค้นพบจากการวิจัยในครั้งนี้ไปใช้ในการทดสอบแบบปรับเหมาะ ดังนี้

1.1 การจัดชุดข้อสอบให้ตอบในแต่ละขั้นของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์นั้น ควรจัดชุดข้อสอบให้ตอบขั้นละ 3 ข้อ จะทำให้ค่าฟังก์ชัน

สารสนเทศของแบบทดสอบสูงที่สุด รองลงมาคือจัดชุดข้อสอบให้ตอบชั้นละ 2 ข้อ ส่วนการจัดชุดข้อสอบให้ตอบชั้นละ 1 ข้อ ค่าฟังก์ชันสารสนเทศของแบบทดสอบจะต่ำที่สุด

1.2 การกำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อของการทดสอบแบบปรับเหมาะ กับความสามารถของผู้สอบด้วยคอมพิวเตอร์นั้น ควรกำหนดเวลาให้ตอบข้อสอบในแต่ละข้อเฉลี่ยเท่ากันทุกข้อ หรือกำหนดเวลาให้ตอบข้อสอบในแต่ละข้อตามน้ำหนักค่าความยากของข้อสอบ กล่าวคือถ้าข้อสอบง่ายก็ให้เวลาตอบน้อย ข้อสอบยากก็ให้เวลาตอบมากกว่าจะทำให้การทดสอบใช้เวลาเฉลี่ยนลง ส่วนการทดสอบที่ไม่กำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อจะใช้เวลาสอบมากที่สุด

1.3 การสร้างรูปแบบของข้อสอบในการทดสอบแบบปรับเหมาะ กับความสามารถของผู้สอบด้วยคอมพิวเตอร์นั้น ควรสร้างรูปแบบของข้อสอบที่มีสื่อประสมประกอบการทดสอบจะทำให้การทดสอบดึงดูดใจ ผู้สอบมีความกระตือรือร้นในการสอบ ผู้สอบมีความพึงพอใจในการสอบมากขึ้น และใช้เวลาในการทดสอบน้อยลง ส่วนการทดสอบด้วยรูปแบบของข้อสอบที่ไม่มีสื่อประสม ทำให้การทดสอบไม่ดึงดูดความสนใจเท่าที่ควร ผู้สอบขาดความกระตือรือร้นในการสอบ และยังใช้เวลาในการทดสอบมากกว่าอีกด้วย

1.4 การทดสอบแบบปรับเหมาะ กับความสามารถของผู้สอบด้วยคอมพิวเตอร์ที่ ผู้วิจัยพัฒนาขึ้นในครั้งนั้น สามารถนำไปใช้ในการทดสอบเพื่อวัดผลการเรียนรู้ตามจุดประสงค์ การทดสอบเพื่อค้นหาข้อบกพร่องของผู้เรียน การทดสอบวัดผลสัมฤทธิ์ เพื่อจัดตำแหน่งของผู้เรียน หรือใช้เป็น ข้อมูลในการประเมินผู้เรียนได้อย่างมีความน่าเชื่อถือ ดังนั้นโปรแกรมการทดสอบที่ผู้วิจัยพัฒนาขึ้นนี้ครู ผู้สอนสามารถนำไปใช้ทำการทดสอบกับนักเรียนในห้องเรียนได้กับทุกวิชา

2. ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

2.1 ควรมีการศึกษาสื่อประสมที่จะนำมาใช้ในการจัดรูปแบบของข้อสอบว่าจะใช้สื่อประสมรูปแบบใด จึงจะเหมาะสมที่สุด เช่น ควรมีการเปรียบเทียบระหว่างรูปภาพที่นำมาใช้ สี สัน ขนาดของรูปภาพ หรือตัวอักษร และชนิดของเสียง ว่าสื่อประสมรูปแบบใดจะทำให้การทดสอบมีประสิทธิภาพมากที่สุด

2.2 ควรมีการศึกษาสื่อประสมที่จะนำมาใช้ในการจัดรูปแบบของข้อสอบกับนักเรียนในระดับที่แตกต่างกันว่าจะใช้สื่อประสมรูปแบบใด จึงจะเหมาะสมกับนักเรียนในแต่ละระดับนั้นๆ มากที่สุด

2.3 การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะนี้ มีค่าใช้จ่ายค่อนข้างสูงจึงควรมีการประเมินค่าใช้จ่ายในการพัฒนาโปรแกรมอย่างละเอียด เพื่อศึกษาความคุ้มค่าของโปรแกรม

2.4 ควรมีการศึกษาและประเมินผลการใช้โปรแกรมในระยะยาวเมื่อมีการนำโปรแกรมการทดสอบไปใช้ในสถานการณ์สอบจริง

บรรณานุกรม

บรรณานุกรม

- คณะกรรมการการศึกษาแห่งชาติ , สำนักงาน. (2542). พระราชบัญญัติการศึกษาแห่งชาติ พุทธศักราช 2542. กรุงเทพฯ
- คณิต ไช้มุกต์. (2534). การพัฒนาเกณฑ์ในการเลือกข้อสอบที่ปรับเหมาะกับความสามารถของผู้สอบ. วิทยานิพนธ์ ค.ด. (การวัดและประเมินผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, อัดสำเนา.
- จักรกฤษณ์ ส้าราญใจ. (2531). ประสิทธิภาพเชิงสัมพัทธ์ของข้อสอบเลือกตอบชนิดตัดสินคำตอบทุกตัวเลือกเทียบกับข้อสอบเลือกตอบชนิดแบบฉบับในแบบสอบผลสัมฤทธิ์. วิทยานิพนธ์ ค.ด.(การวัดและประเมินผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, อัดสำเนา.
- ชาลิต โพธิ์นคร. (2540). การศึกษาอิทธิพลของการให้คะแนนและการคัดเลือกข้อสอบที่มีต่อคุณภาพแบบทดสอบเลือกตอบ. ปริญญาานิพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ, อัดสำเนา.
- ✓ คริ่งใจ พูลผลอำนวยการ. (2534). การพัฒนาแบบสอบเฉพาะบุคคลในวิชาวิทยาศาสตร์สำหรับ นักเรียนชั้นมัธยมศึกษาปีที่ 3. วิทยานิพนธ์ ค.ด. (การวัดและประเมินผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, อัดสำเนา.
- ถ่าย เชียงฉนิ. (2534). การศึกษาเปรียบเทียบประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบเทเลอรรูปปรามิตที่มีรูปแบบจำนวนชั้นและวิธีการให้คะแนนที่แตกต่างกันโดยใช้วิธีมอนติคาร์โล. ปริญญาานิพนธ์ กศ.ด.(การวิจัยและพัฒนาหลักสูตร) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ, อัดสำเนา.
- ณัฐพงษ์ งามแสง. (2539). การเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบสอบแบบตัวเลือกซ้อนระหว่างประเภทที่กำหนดและไม่กำหนดสถานการณ์คำตอบ. ปริญญาานิพนธ์ ค.ม. (การวัดและประเมินผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, อัดสำเนา.
- ทัศนีย์ ชาติไทย. (2539). การประยุกต์ทฤษฎีการตอบสนองข้อสอบเพื่อศึกษา คุณภาพและประสิทธิภาพของแบบทดสอบอิงเกณฑ์ที่มีความยากและความยาวต่างกัน. ปริญญาานิพนธ์ กศ.ด. (การวิจัยและพัฒนาหลักสูตร) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ, อัดสำเนา.
- ✓ นันทิยา ฟิ่งคำ. (2531). การเปรียบเทียบคุณภาพการทดสอบแบบซี เอ ที และแบบประเพณีนิยมในการวัดความสามารถด้านคำศัพท์ภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3. ปริญญาานิพนธ์ ค.ด. (การวัดและประเมินผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, อัดสำเนา.

- บุญเชิด ภิญญอนันตพงษ์. (2545). การประเมินผลการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ : แนวคิดและวิธีการ. กรุงเทพฯ : สำนักงานคณะกรรมการการศึกษาแห่งชาติ.
- _____. (2544). รายงานการวิจัยฉบับสมบูรณ์เรื่องการพัฒนามาตรฐานประเมินคุณลักษณะความสามารถพิเศษของเด็กนักเรียนอายุ 6 – 12 ปี ตามแนวคิดของ Penzulli. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ผจงจิต อินทสุวรรณ. (2545). การวิเคราะห์ความแปรปรวนหลายตัวแปร. กรุงเทพฯ : พัฒนา จำกัด.
- _____. (2528). ทฤษฎีการตอบข้อคำถาม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ.
- มยุรี ศรีชัย. (2538). เทคนิคการสุ่มกลุ่มตัวอย่าง. กรุงเทพฯ : วีเจ. พรีเมียม.
- ระพีพรรณ ศรีวิเชียร. (2535). การพัฒนาแบบสอบตามระดับความสามารถในวิชาคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3. วิทยานิพนธ์ ค.ด. (การวัดและประเมินผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัยจุฬาลงกรณ์มหาวิทยาลัย, อัดสำเนา.
- ✓ รังสรรค์ มณีเล็ก. (2540). ผลของตัวแปรบางตัวต่อความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์. ปรินญาณิพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัยมหาวิทยาลัยศรีนครินทรวิโรฒ, อัดสำเนา.
- รัตนา ศรีเหรียญ. (2539). การเปรียบเทียบประสิทธิภาพของวิธีประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซ์ิมันไลด์ลิสต์ วิธีฮิวริสติก และวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกัน ด้วยเทคนิคมอนติคาร์โล. ปรินญาณิพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ, อัดสำเนา.
- เรวดี อินทสระ. (2539). ผลการตรวจสอบความลำเอียงของข้อสอบต่อการศึกษาคความเที่ยงตรงเชิงพยากรณ์ของแบบทดสอบคัดเลือกที่คิดคะแนนต่างกัน. ปรินญาณิพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ, อัดสำเนา.
- วดาภรณ์ พูลผลอำนาจ. (2534). ผลของรูปแบบตัวเลือกต่างกันที่มีต่อรูปแบบการตอบสนองข้อสอบไค้งลักษณะของข้อสอบ ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ. วิทยานิพนธ์ ค.ด. (การวัดและประเมินผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, อัดสำเนา.
- วิรัช วรรณรัตน์. (2535). วิธีการทางสถิติสำหรับการวิจัย. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ศิริชัย กาญจนวาสี. (2538). การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive Testing). กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย.
- ศิริพงษ์ ภูพินนา. (2546). ผลการใช้บทเรียนคอมพิวเตอร์มัลติมีเดียแบบฝึกที่มีการควบคุมการเรียนรู้ 3 แบบกับผู้เรียนที่มีระดับความสามารถทางการเรียนต่างกัน. ปรินญาณิพนธ์ กศ.ม. (เทคโนโลยีทางการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ, อัดสำเนา.

- ส.วาสนา ประवालพฤษ์. (2543). รายงานการวิจัยเรื่องระบบการวัดและประเมินผลผู้เรียนระดับอุดมศึกษาตามแนวพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ.2542. กรุงเทพฯ : ทบวงมหาวิทยาลัย.
- _____. (2545). หลักการและเทคนิคการประเมินทางการศึกษา. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ.
- สมหวัง บุญสิทธิ์. (2542). ปัจจัยบางประการที่ส่งผลต่อความแม่นยำในการวัดของแบบทดสอบอัตโนมัติที่สร้างโดยการคัดเลือกข้อคำถามวิธี *Binary Programming* ปริญญาโท กศ.ด. (การทดสอบและวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ, อัดสำเนา.
- สมศักดิ์ ลีลา. (2539). การพัฒนาระบบคลังข้อสอบเพื่อการเรียนการสอนด้วยไมโครคอมพิวเตอร์. ปริญญาโท กศ.ด. (การทดสอบและวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ, อัดสำเนา.
- สายชล ออบทม. (2539). การพัฒนาโปรแกรมสำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์. ปริญญาโท ค.ม. (การวัดและประเมินผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, อัดสำเนา.
- สำเร็จ บุญเรืองรัตน์. (2529, มกราคม – เมษายน). "การพัฒนาทฤษฎีลาแทนเทรทเพื่อวิเคราะห์ข้อสอบ" *วารสารการวัดผลการศึกษา*. 7(21) : 41 – 66.
- อัญชิกา ดิษเจริญ. (2536). การเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบสอบเลือกตอบต่างแบบ. ปริญญาโท ค.ม. (การวัดและประเมินผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, อัดสำเนา.
- Allen, Mary J. and Yen, Wendy M. (1979). *Introduction to Measurement Theory*. Monterey : Brooks Cole Publishing Company.
- Almond, Russell G. and Mislevy, Robert J. (1999, September.) "Graphical Models and Computerized Adaptive Testing," *Applied Psychological Measurement*. 23(3) : 223 - 237
- Baker, Frank B. (1992). *Item Response Theory : Parameter Estimation Techniques*. New York : Marcel Dekker, Inc.
- Berger, Martijn P. F. (1994, June). "A General Approach to Algorithmic Design of Fixed-Form Tests, Adaptive Tests, and Testlets," *Applied Psychological Measurement*. 18(2) : 141 – 153.
- Bizot, Elizabeth B ; Goldman, Steven H. (1994, April). *The Practical Impact of IRT Models and Parameters When Converting a Test to Adaptive Format*.
<http://ercir.syr.edu/plweb-cgi/fastweb>.

- Chang, Hua-Hua and Ying, Zhiliang. (1996, September). "A Global Information Approach to Computerized Adaptive Testing," *Applied Psychological Measurement*. 20(3) : 213 – 229.
- Chang, Hua-Hua and Ying, Zhiliang. (1999, September). " α -Stratified Multistage Computerized Adaptive Testing," *Applied Psychological Measurement*. 23(3): 211 - 222.
- Chen, Shu-Ying and Others. (2000, September). "A Comparison of Item Selection Rules at the Early Stages of Computerized Adaptive Testing," *Applied Psychological Measurement*. 24(3) : 241 – 255.
- Cheng, Philip E. and Liou, Michelle. (2000, September). "Estimation of Trait Level in Computerized Adaptive Testing," *Applied Psychological Measurement*. 24(3) : 257 – 265.
- Dood, Barbara G. and Koch, William R. (1993). "Computerized Adaptive Testing Using the partial Credit Model : Effects of Item Pool Characteristics and Different Stopping Rules," *Educational and Psychological Measurement*. 53 : 61–77.
- Dodd, Barbara G. and Others. (1995, March). "Computerized Adaptive Testing With Polytomous Items," *Applied psychological Measurement*. 19(1) : 5 – 22.
- Eggen, T. J. H. M. (1999, March). "Item Selection in Adaptive Testing With the Sequential probability Ratio Test," *Applied psychological Measurement*. 23(3) : 249 – 261.
- Eggen, T. J. H. M. and Straetmans, F. J. J. M. (2000, October). "Computerized Adaptive Testing for Classifying Examinees into Three Categories," *Educational and Psychological Measurement*. 60(5) : 713 – 734.
- Estes, Gary D., Ed. *Examples of Item Banks to Support Local Test Development : Two Case Studies With Reactions*. <http://ercir.syr.edu/plweb-cgi/fastweb>.
- Green, B. F. And Others. (1984). "Tech Guidelines for Assessing Computerized Adaptive Tests" *Journal of Educational Measurement*. 21(4) : 347 – 360.
- Gruijter, Dato N. M. De and Kamp, Leo J. Th. Van der. (1976). *Advances in Psychological and Educational Measurement*. London : John Wiley & Sons.
- Gulliksen, Harold. (1950). *Theory of Mental Tests*. New York : John Wiley & Sons, Inc.
- ✓ Hambleton, Ronald K. and Swaminathan, Hariharan. (1985). *Item Response Theory*. U.S.A. : Kluwer - Nijhoff Publishing.
- Hambleton, R. K., and Cook, L. L. (1977). "Latent Trait Models and Their Use in the Analysis of Educational Test Data" *Journal of Educational Measurement*. 14.
- Hetter, Rebecca D. and Others. (1994, October). "A Comparison of Item Calibration Media in Computerized Adaptive Test," *Applied Psychological Measurement*. 18(3) : 197 – 204.

- Kirk, Roger E. (1995). *Experimental Design : Procedures for the Behavioral Sciences*. U.S.A. : Brooks/Cole Publishing Company.
- Linder, Wim J. van der. and Others. (1999, September). "Using Response-Time Constraints to Control for Differential Speededness in Computerized Adaptive Testing," *Applied Psychological Measurement*. 23(3) : 187 – 194.
- Linn, Robert L. (1989). *Educational Measurement*. New York : Macmillan Publishing Company.
- Lord, F. M. (1980). *Applicational of Item Response Theory to Practical Testing Problems*. Hillsdale, New York : Lawrence Erlbaum Associates.
- Lork, Frederic M. (1980). *Applications of Item Response Theory to Practical Testing Problems*. New Jersey : Lawrence Erlbaum Associates, Publishers.
- Lord, Frederic M. and Novick, Melvin R. (1968). *Statistical Theories of Mental Test Scores*. Canada : Addison-Wesley Publishing Company, Inc.
- McDonald, R. P. (1982). "Linear Versus Non-linear Models in Item Response Theory" *Applied Psychological Measurement*. 6 : 379 – 396.
- McDonald, R. P. And K. S. Ahlawat. (1974). "Difficulty Factors in Binary Data" *British Journal of Mathematical and Statistical Psychology*. 27 : 82 – 89.
- McKinley, Robert L.;Reckase, Mark D. *An Evaluation of One–and Three- Parameter Logistic Tailord Testing Procedures for Use with Small Item Pools*.
<http://ercir.syr.edu/plweb-cgi/fastweb>.
- Meijer, Rob R. and Nering, Michael L. (1999, September). "Computerized Adaptive Testing : Overview and Introduction," *Applied psychological Measurement*. 23(3) : 187 – 194.
- Myers, Jerome L. (1972). *Fundamentals of Experimental Design*. Boston :Allyn and Bacon, Inc.
- Nicewander, W. Alan and Thomasson, Gary L. (1999, September). "Som Reliability Estimates for Conmputerized Adaptive Tests," *Applied psychological Measurement*. 23(3) : 239 – 247.
- Owen, R. (1975). *A Bayesian sequential procedure for quantal response in the context of adaptive mental testing*. *Journal of the American Statistical Association*. 70, 351 - 356.
- Ponsoda, Vicente and Others. (1999). "The Effects of Test Difficulty Manipulation in Computerized Adaptive Testing and Self-Adapted Testing," *Applied Measurement in Education*. 12(2) : 167 – 184.
- Revuelta, Javier and Ponsoda, Vicente. (1998). "A Comparison of Item Exposure Control methods in Computerized Adaptive Testing," *Journal of Educational Measurement*. 35(4) : 311 – 327.

Stocking, Martha L. and Swanson, Len. (1998, September). "Optimal Design of Item Banks for Computerized Adaptive Tests," *Applied Psychological Measurement*. 22(3) : 271 – 279.

✓ Straetmans, Gerard J. J. M. and Eggen, Theo J. H. M. (1998, January-February).

"Computerized Adaptive Testing : What It Is and How It Works," *Educational Technology*. 38(1) : 45 – 52.

Tatsuoka, Maurice M. and Lohnes, Paul R. (1988). *Multivariate Analysis*. New York : Macmillan publishing Company.

Traub, Ross E. (1994). *Reliability for the Social Sciences : Theory and Applications*. California : SAGE Publications.

Urry, V. W. (1977). "Tailored Testing : A Successful Application of Latent Trait Theory" *Journal of Educational Measurement*. 14(2) : 181 – 196.

Wang, Tianyou and Hanson, Bradley A. (1999, September,). "Reducing Bias in CAT Trait Estimation : A Comparison of Approaches," *Applied Psychological Measurement*. 23(3) : 263 – 278.

Wang, Tianyou and Vispoel, Walter P. (1998). "Properties of Ability Estimation Methods in Computerized Adaptive Testing," *Journal of Educational Measurement*. 35(2) : 109 – 135.

Weiss, David J. (1973). *The Stratified Adaptive Computerized Ability Test*. Minnesota : Department of Psychology, University of Minnesota.

_____. (1974). *Strategies of Adaptive Ability Measurement*. Minnesota : Department of Psychology , University of Minnesota.

Wise, Steven L. (Online). *Comparison of Stratum Scored and Maximum-Likelihood Scored CATs*. <http://ercir.syr.edu/plweb-cgi/fastweb>.

Wise, Steven L. *Examinee Issues in CATs*. <http://ercir.syr.edu/plweb-cgi/fastweb>.

Wise, Steven L. and Others. (1999). "Examinee Judgments of Changes in Item Difficulty : Implications for item Review in Computerized Adaptive Testing," *Applied Measurement in Education*. 12(2) : 185 – 198.

Xiao, Beiling. (1999, June). "Strategies for Computerized Adaptive Grading Testing," *Applied Psychological Measurement*. 23(2) : 136 - 146.

Yan, Duanli ; Lewis, Charles and Stocking, Martha. (Online). *Adaptive Testing Without IRT*. <http://ercir.syr.edu/plweb-cgi/fastweb>

Zwick, Rebecca and Others. (1994, June). "A Simulation Study of Methods for Assessing Differential Item Functioning in Computerized Adaptive Tests," *Applied Psychological Measurement*. 18(2) : 121 – 140.

ภาคผนวก ก
โปรแกรมคอมพิวเตอร์สำหรับการทดสอบ

โปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะ

```

unit TestUtil;
interface
uses Classes, Graphics, ItemTime;
const
  ProgramStr = 'The Test data file 95N v';
  VersionStr = '2.2';
  DataSignature: String[Length(ProgramStr + VersionStr) + 1] =
    ProgramStr + VersionStr + #26;
  Choices: array[0..2] of String[5] = (
    '๓๔๕๖๗', 'ABCDE', '12345');
type
  TTestPassword = String[8];
  TTestObj = class(TObject)
    PassWord: TTestPassword;
    TestName: String[40];
    ItemTimeRec: TSetItemTimeRec;
    ChoiceType: Byte;
    NumChoice: Byte;
    MaxItem: Integer;
    DoTime: Integer;
    CurlItem: 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;
  end;

```

```

    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;
{ Stream Routines }
function GetTestFromStream(S: TStream): Boolean;
procedure PutTestToStream(S: TStream);
{ 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 GetSEZeta(IZeta: 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;
```

```
  CurrentRowSqr: Double = 1.0;
```

```
  SumIZetaMax: Double = 0.0;
```

```
  ZetaAvg: Double = 0.0;
```

```
  ZetaSum: Double = 0.0;
```

```
  StartTime: TDateTime = 0;
```

```
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;
{ 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;
  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

```



```
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;
    end;
    end;
end;

```

```

CurlItem := 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(CurlItem));
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;
{ 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.NumItems + 1)
    + IntToStr(DoTestRec.TimeUsed + 1)
    + IntToStr(DoTestRec.TestType + 1);
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);
    end;
  end;

```

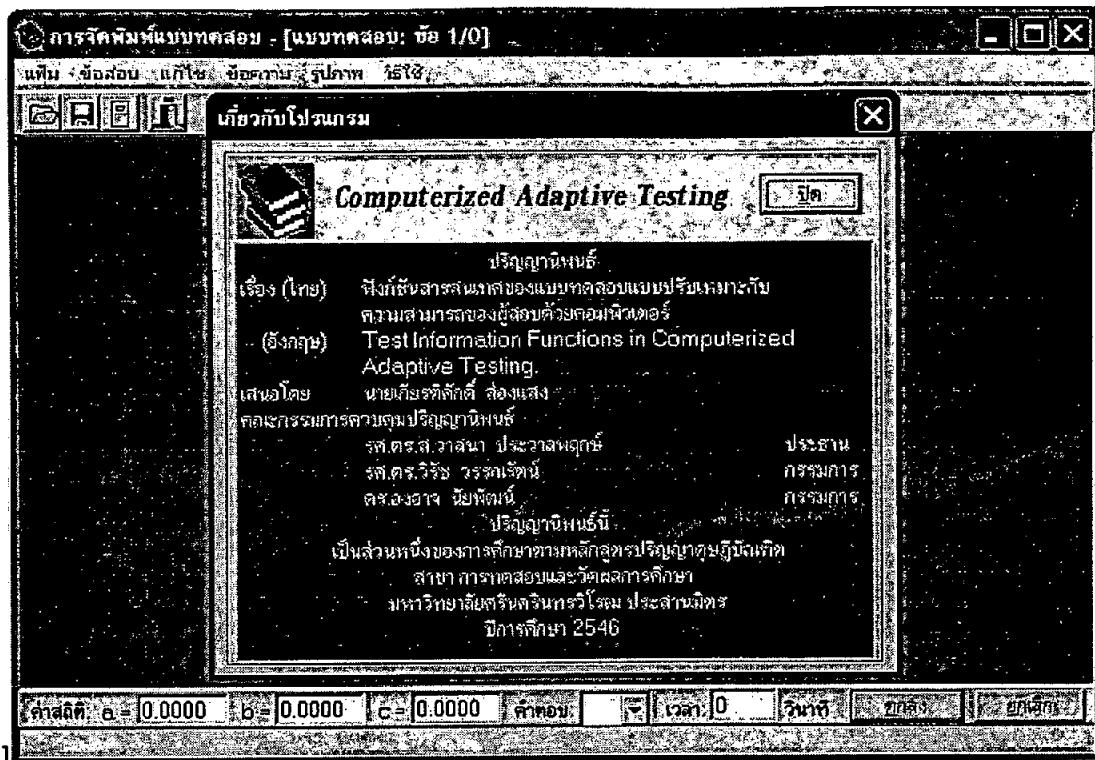
```
    Result := RowSqr * (1 - Result);
end else
begin
    Result := (O_D / Tmp1) * (O_D / A_Dp + D);
    Result := RowSqr * (1 - Result / A_Dp);
end;
end;
function GetIZetaMax(A, B, C: Single; Zeta: Double): Double;
var
    Tmp: Double;
begin
    Tmp := 1.7 * A * (Zeta - B);
    Result := (C + Exp(Tmp)) * Power((1 + Exp(-Tmp)), 2);
    Result := Power(1.7 * A, 2) * (1 - C) / Result;
end;
function GetSEZeta(IZeta: Double): Double;
begin
    if IZeta > 0 then
        Result := 1 / Sqrt(IZeta)
    else
        Result := 0;
    end;
end;
end.
```

ภาคผนวก ข

คู่มือโปรแกรมพิมพ์ข้อสอบ

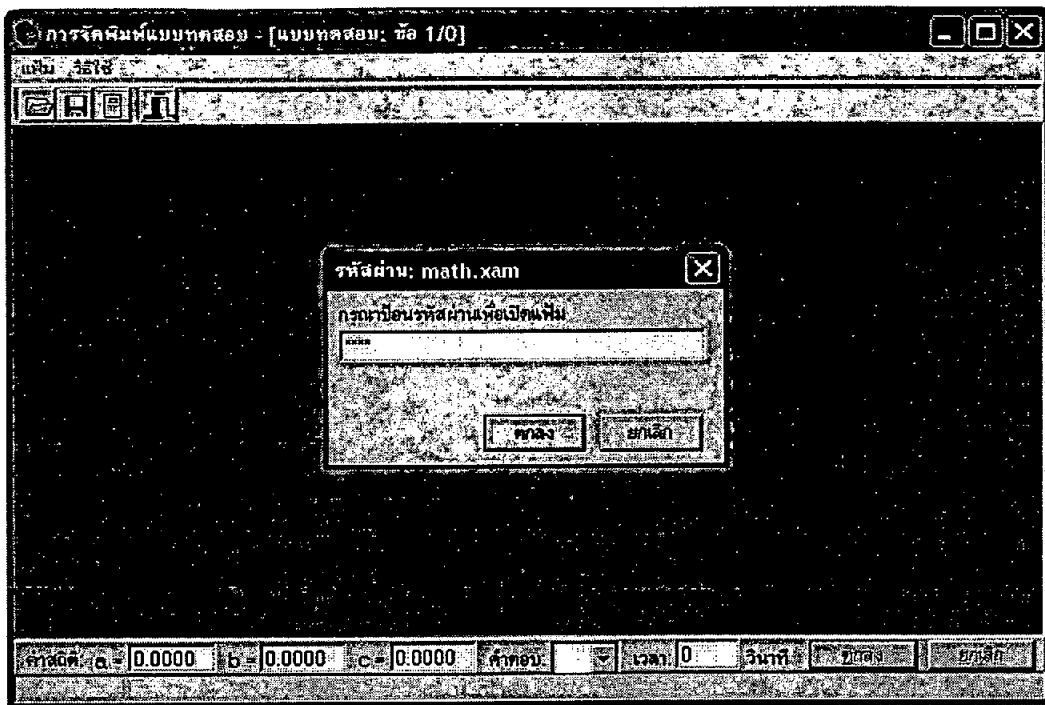
คู่มือโปรแกรมการทำข้อสอบ

**คู่มือโปรแกรม
สำหรับพิมพ์ข้อสอบ**



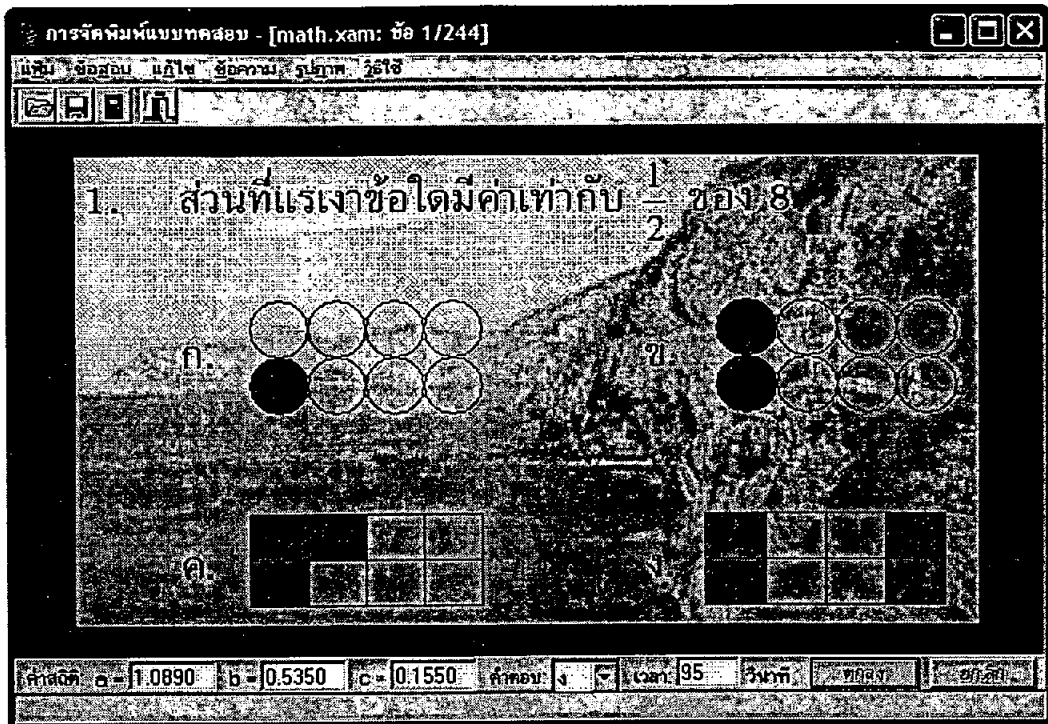
ภาพประกอบ 1 โลโก้หน้าแรกแสดงโปรแกรมวัดพิมพ์แบบทดสอบ

เมื่อคลิกปุ่มปิด หรือกด Enter บนแป้นพิมพ์ โปรแกรมจะแสดงหน้าจอถัดไปให้ป้อนรหัสผ่าน เพื่อเข้าสู่โปรแกรมการวัดพิมพ์เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบ ดังภาพประกอบ 2

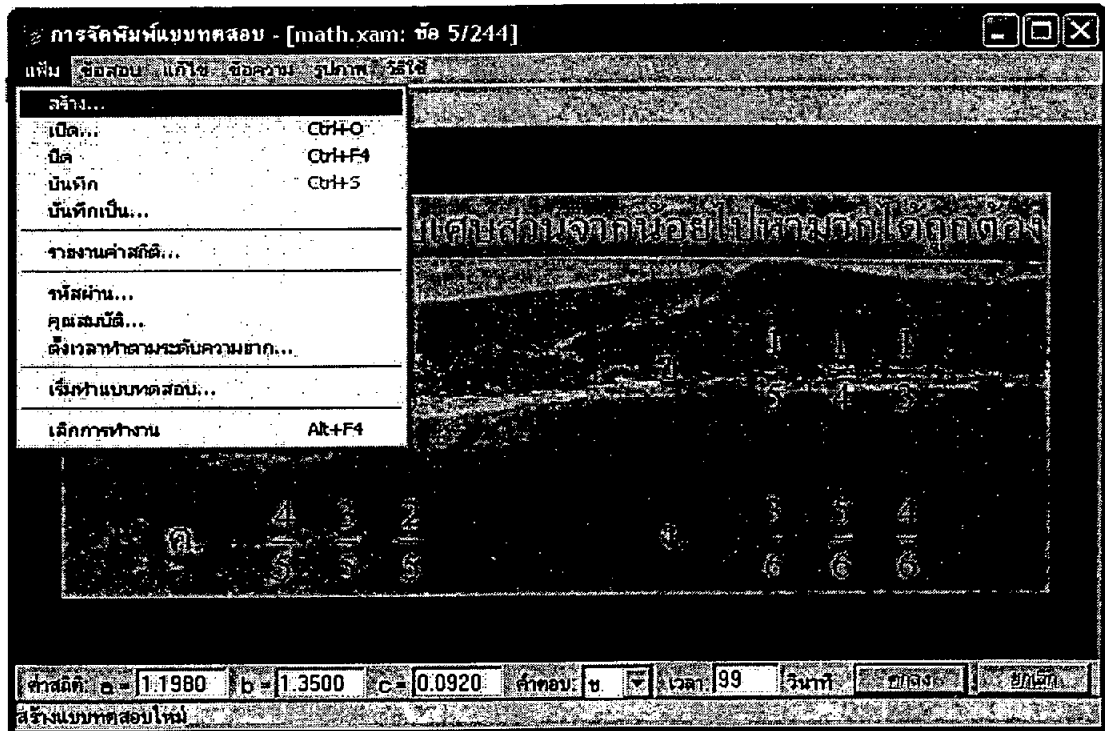


ภาพประกอบ 2 หน้าจอของโปรแกรมเพื่อป้อนรหัสผ่าน

เมื่อป้อนรหัสผ่านได้ถูกต้อง และคลิกเมาส์ที่ปุ่มตกลงโปรแกรมจะให้ผ่านเข้าสู่โปรแกรม ซึ่งสามารถพิมพ์เพิ่ม ลด เปลี่ยนแปลงแก้ไขข้อสอบ และค่าพารามิเตอร์ของข้อสอบได้ตามที่ต้องการ ดังภาพประกอบ 3

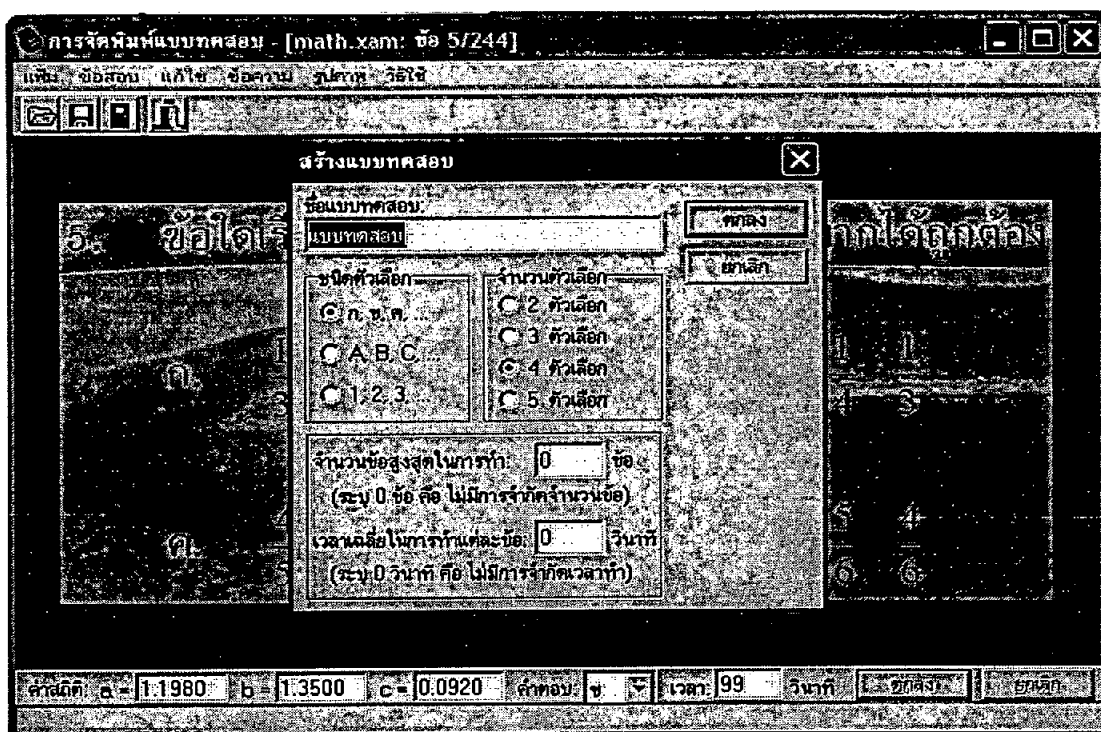


ภาพประกอบ 3 หน้าจอแสดงโปรแกรมเพิ่ม ลด เปลี่ยนแปลงแก้ไขข้อสอบ

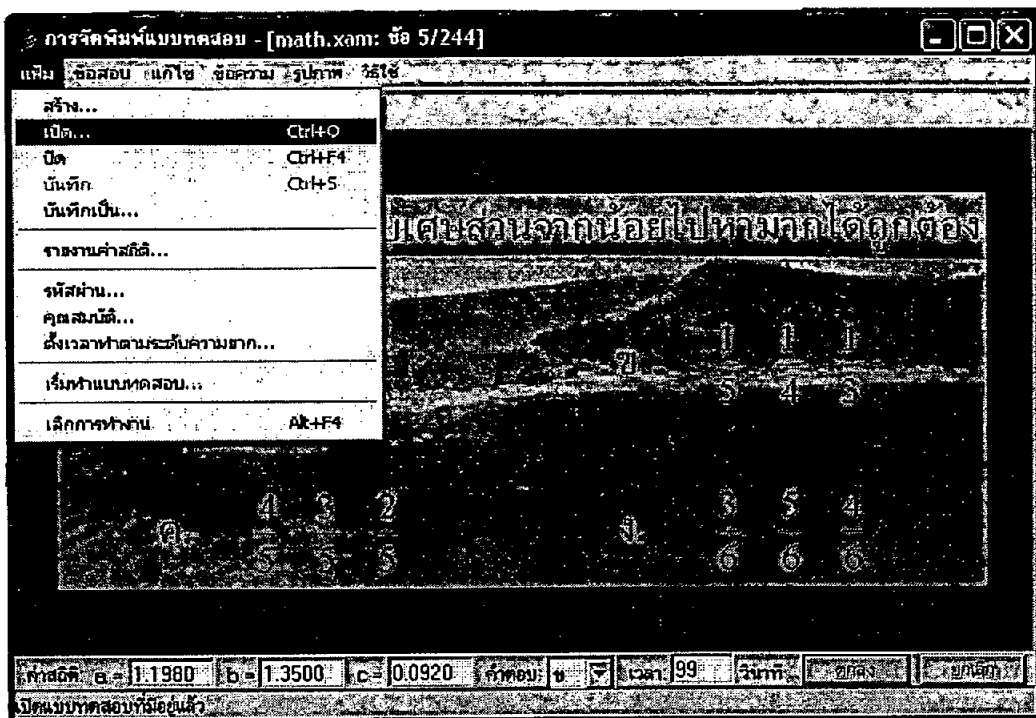


ภาพประกอบ 4 หน้าจอโปรแกรมคำสั่งเพิ่ม - สร้าง

เมื่อเลือกคำสั่งแฟ้ม และคลิกที่คำสั่งสร้าง โปรแกรมจะยอมให้ผู้สร้างแบบทดสอบกำหนด ชื่อแบบทดสอบ ชนิดของตัวเลือก จำนวนตัวเลือก จำนวนข้อสูงสุดของแบบทดสอบ เวลาเฉลี่ยในการตอบข้อสอบแต่ละข้อ ดังภาพประกอบ 5

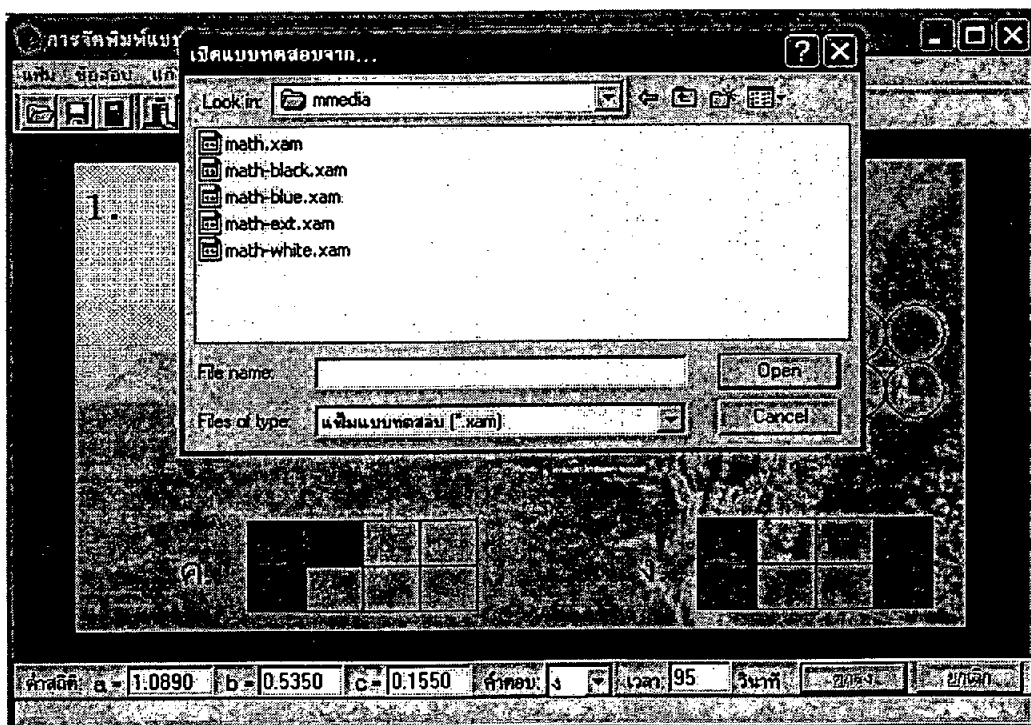


ภาพประกอบ 5 หน้าจอโปรแกรมสร้างแบบทดสอบ

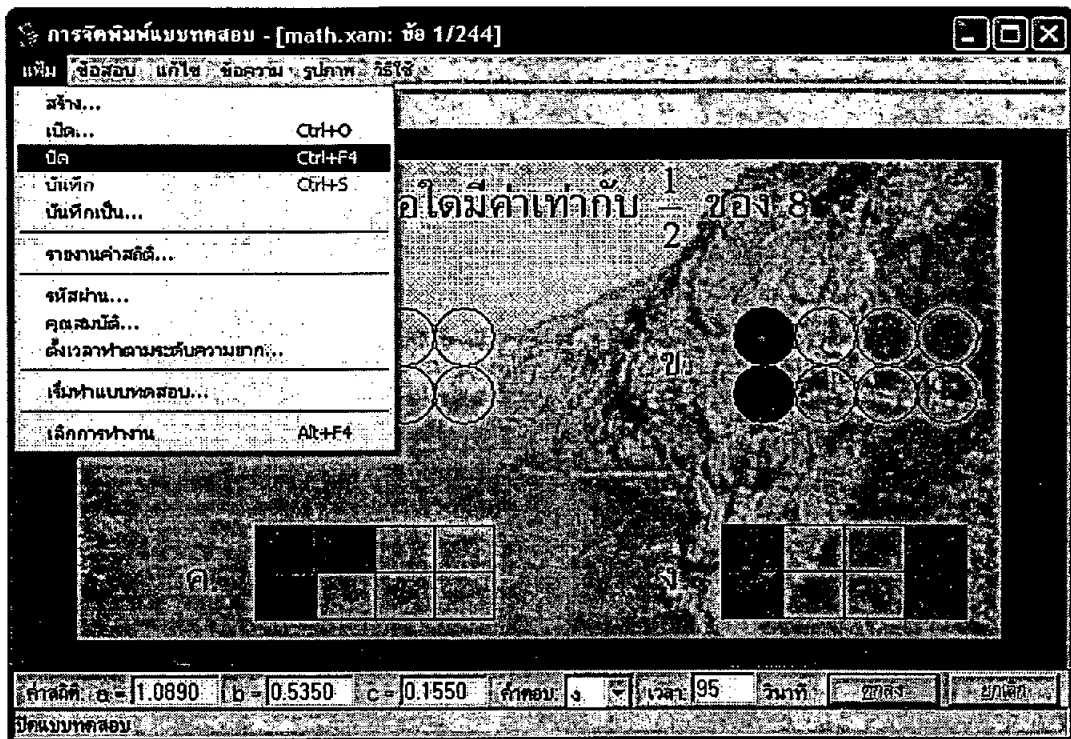


ภาพประกอบ 6 หน้าจอโปรแกรมคำสั่งแฟ้ม-เปิด

เมื่อเลือกคำสั่งแฟ้ม และคลิกที่คำสั่งเปิด โปรแกรมจะให้เลือกแบบทดสอบจากชื่อของแบบทดสอบที่มีอยู่ในโปรแกรมเดิม เพื่อนำแบบทดสอบฉบับนั้นมา เพิ่ม ลด เปลี่ยนแปลง แก้ไข ข้อสอบและค่าพารามิเตอร์ของข้อสอบ ดังภาพประกอบ 7

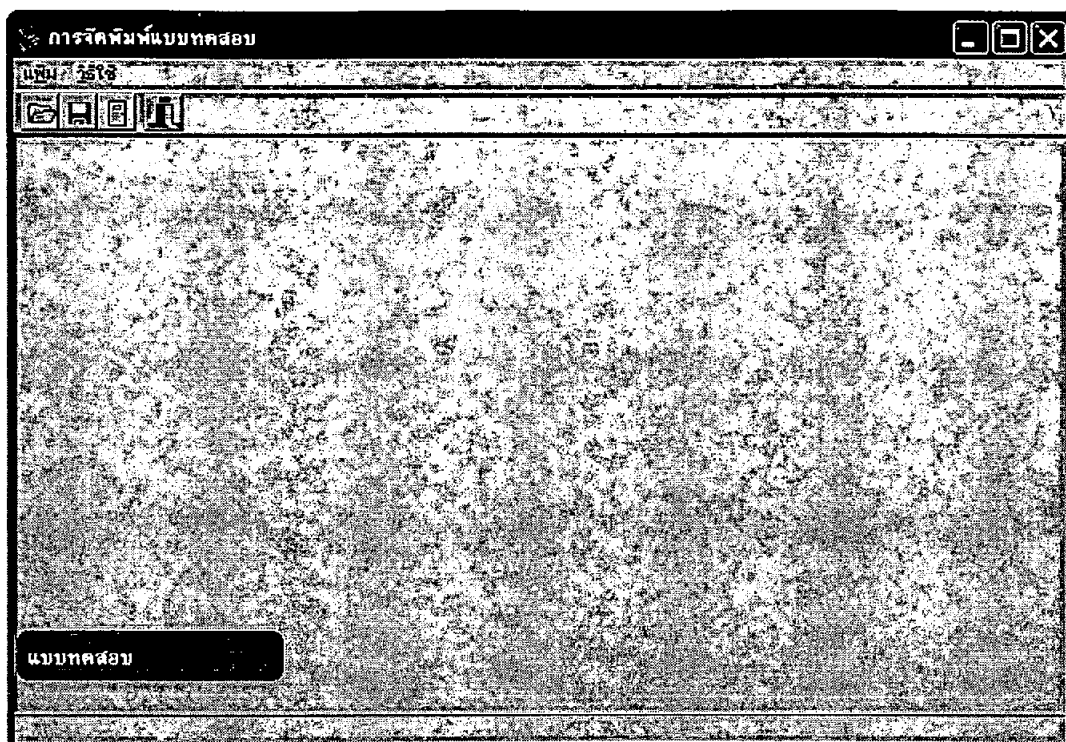


ภาพประกอบ 7 หน้าจอโปรแกรมให้เลือกเปิดแบบทดสอบที่มีอยู่เดิม

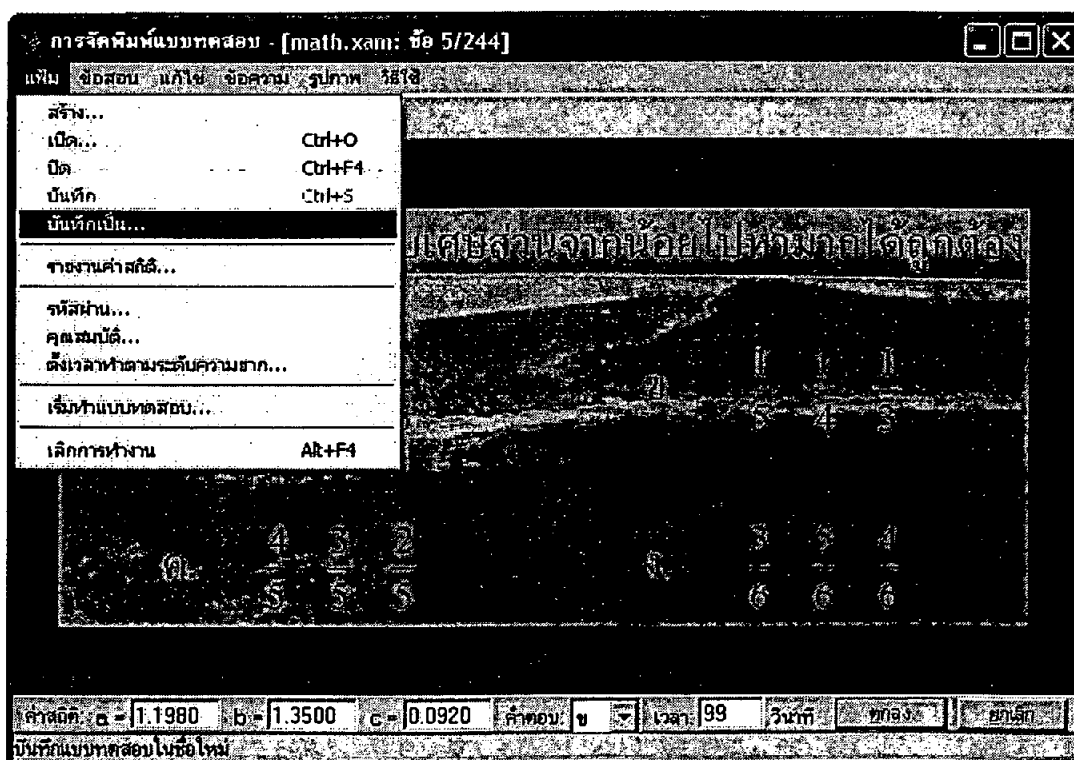


ภาพประกอบ 8 หน้าจอโปรแกรมคำสั่งแฟ้ม – ปิด

เมื่อเลือกคำสั่งแฟ้ม และคลิกที่คำสั่งปิด โปรแกรมจะปิดแบบทดสอบที่กำลังสร้างหรือแบบทดสอบที่กำลังแก้ไขเพิ่มเติมอยู่ในขณะนั้นลง หรือจะใช้คำสั่งนี้เมื่อสิ้นสุดการทำงาน เมื่อคลิกคำสั่งปิดจะได้หน้าจอตั้งภาพประกอบ 9

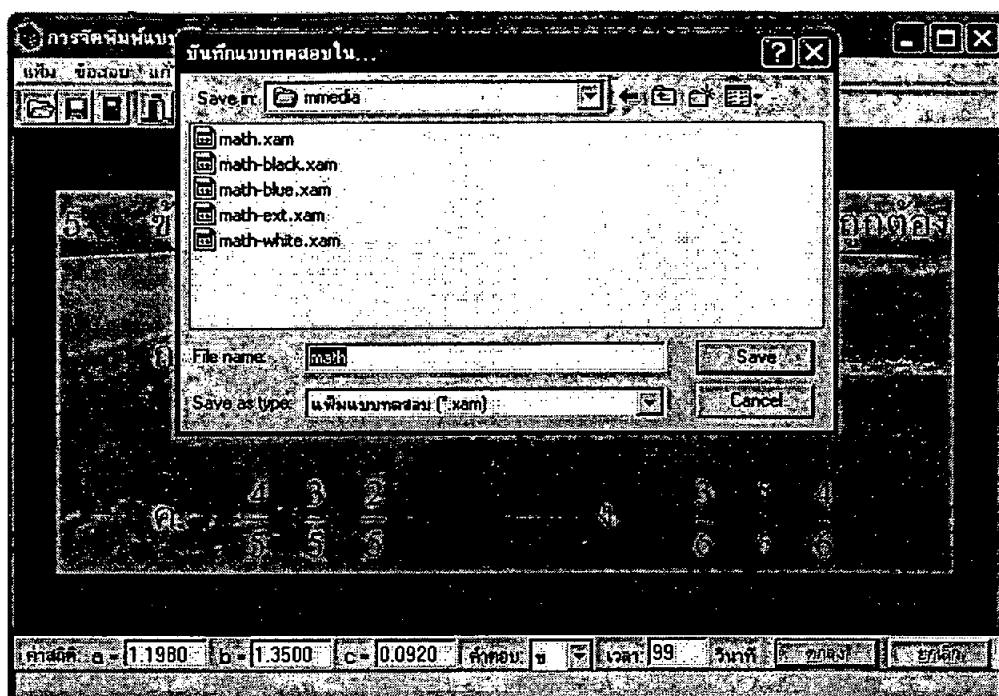


ภาพประกอบ 9 หน้าจอโปรแกรมเมื่อเลือกคำสั่งปิด

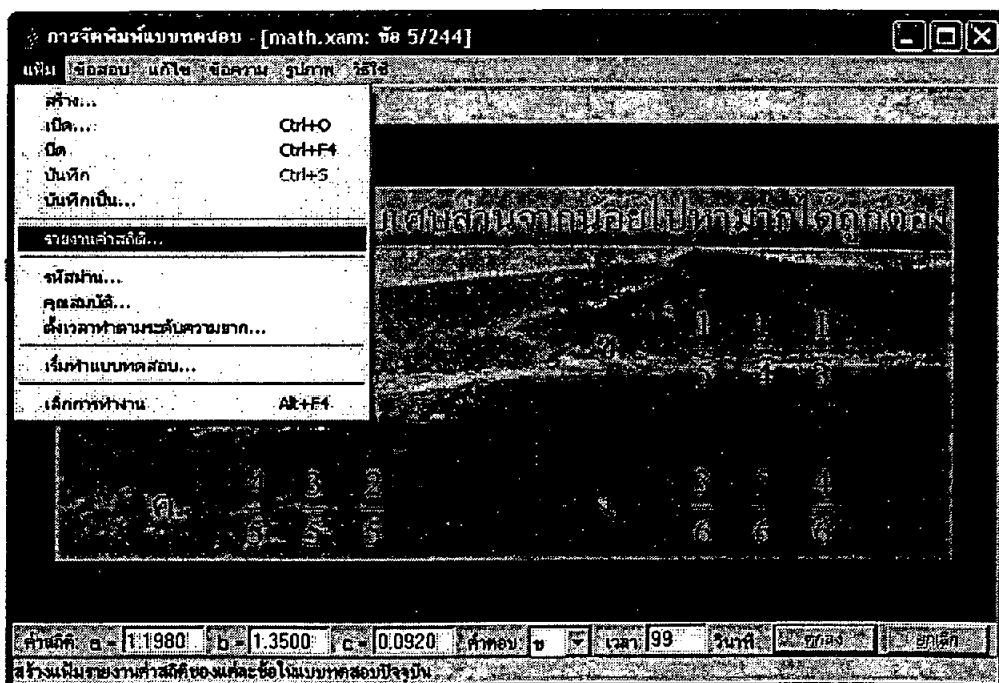


ภาพประกอบ 10 หน้าจอคำสั่ง แฟ้ม-บันทึก และบันทึกเป็น

เมื่อเลือกคำสั่งเพิ่ม และคลิกที่ปุ่มบันทึก หลังจากเพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขข้อสอบแล้ว โปรแกรมจะบันทึกสิ่งที่เปลี่ยนแปลงให้ หรือถ้าคลิกที่คำสั่งบันทึกเป็น โปรแกรมจะให้เลือกชื่อไฟล์หรือทำการบันทึก ดังภาพประกอบ 11

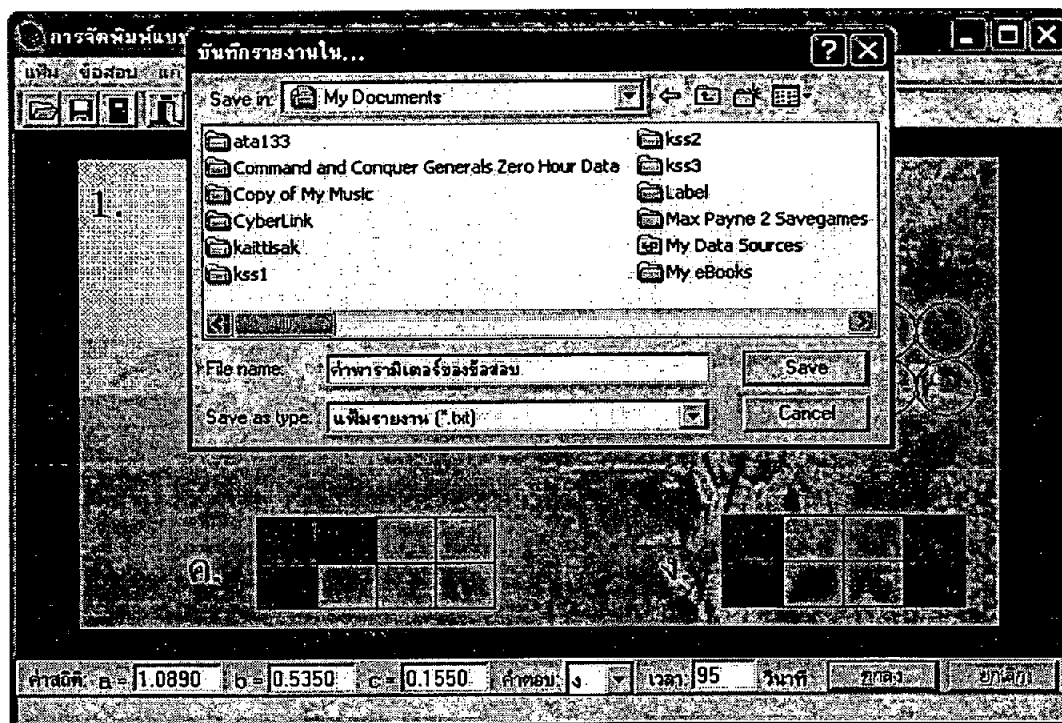


ภาพประกอบ 11 หน้าจอสำหรับการบันทึกลงในชื่อไฟล์ที่ต้องการ



ภาพประกอบ 12 หน้าจอโปรแกรมคำสั่งเพิ่ม-รายงานค่าสถิติ

เมื่อเลือกคำสั่งแฟ้ม และคลิกเมาส์ที่คำสั่งรายงานค่าสถิติ โปรแกรมจะให้เลือกไฟล์ และชื่อของไฟล์ที่ทำการบันทึกค่าสถิติของข้อสอบทุกข้อในแบบทดสอบฉบับนั้น ดังภาพประกอบ 13 เมื่อเลือกไฟล์ และชื่อของไฟล์ตามภาพประกอบ 13 แล้วค่าสถิติของข้อสอบจะปรากฏดังภาพประกอบ 14



ภาพประกอบ 13 หน้าจอโปรแกรมสำหรับเลือกไฟล์ที่จะทำการบันทึก

ค่าพารามิเตอร์ของข้อสอบ - Notepad

File Edit Format View Help

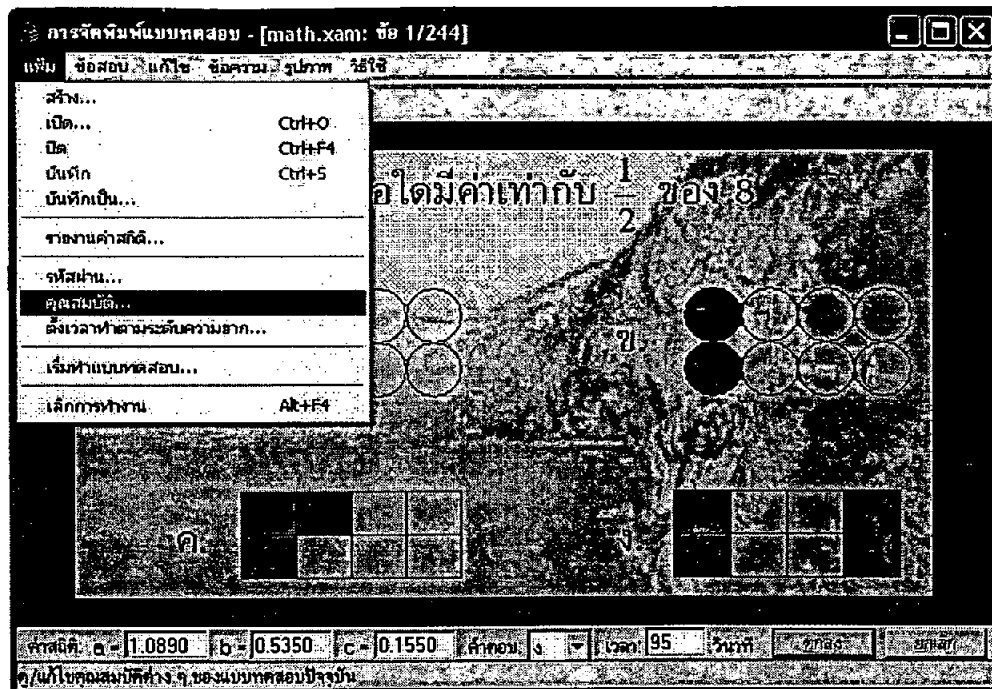
รายงานค่าสถิติของแต่ละข้อในแบบทดสอบ
=====

ชื่อแบบทดสอบ : การทดสอบเรื่อง เศษส่วน
ชนิดตัวเลือก : ก, ข, ค, ...
จำนวนตัวเลือก : 4 ตัวเลือก
จำนวนข้อสูงสุดในการทำ : 36 ข้อ
เวลาเฉลี่ยในการทำแต่ละข้อ: 90 วินาที
จำนวนข้อสอบ : 244 ข้อ

Item	A	B	C	Answer	Time	Image File
1	1.0890	0.5350	0.1550	ง	95	Item001.bmp
2	1.0640	0.6670	0.1520	ค	95	Item002.bmp
3	0.8560	2.2140	0.2680	ข	107	Item003.bmp
4	0.9790	2.2110	0.1320	ค	107	Item004.bmp
5	1.1980	1.3500	0.0920	ข	99	Item005.bmp
6	0.7070	1.1370	0.2400	ก	99	Item006.bmp
7	0.9610	0.2410	0.2490	ง	90	Item007.bmp
8	0.8100	0.1910	0.2390	ค	90	Item008.bmp
9	1.2980	1.4610	0.2240	ง	99	Item009.bmp
10	1.5130	2.0550	0.1920	ข	107	Item010.bmp
11	0.7530	-0.0750	0.2900	ข	84	Item011.bmp
12	0.9440	2.0020	0.1870	ค	103	Item012.bmp
13	0.9320	0.6420	0.2180	ก	95	Item013.bmp
14	1.0820	-0.0300	0.1970	ข	84	Item014.bmp
15	1.1850	-0.3000	0.2100	ก	84	Item015.bmp
16	1.6130	2.7200	0.1800	ง	112	Item016.bmp

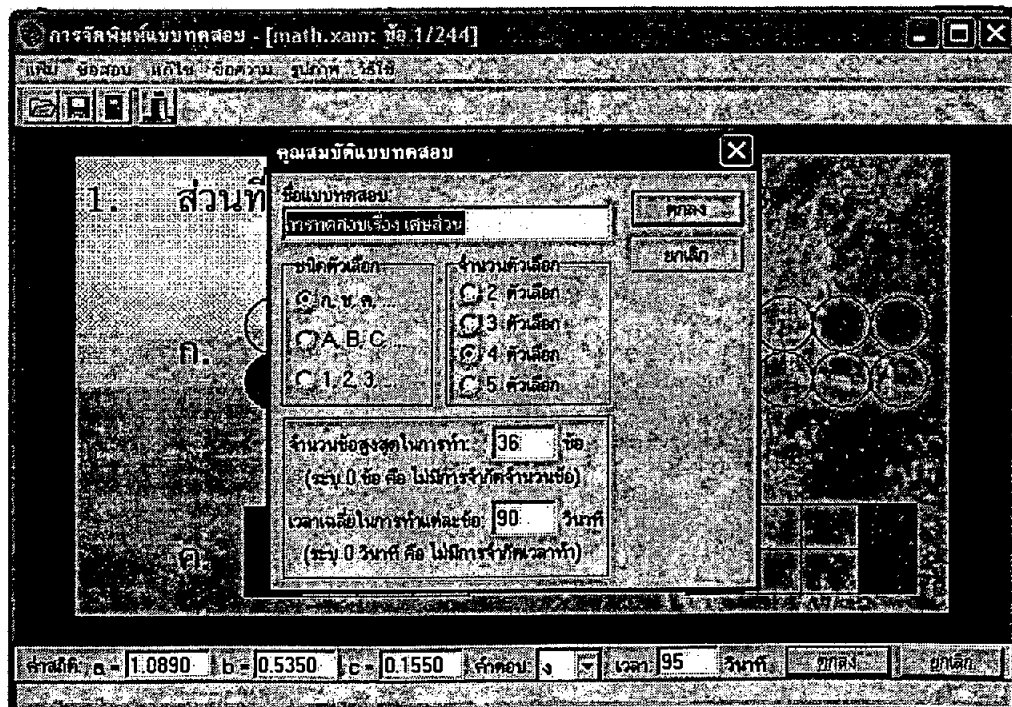
ภาพประกอบ 14 ค่าสถิติของข้อสอบแต่ละข้อของแบบทดสอบ

ค่าสถิติของข้อสอบ และแบบทดสอบซึ่งประกอบด้วย ชื่อของแบบทดสอบ ชนิดของตัวเลือก จำนวนตัวเลือก จำนวนข้อสูงสุดของแบบทดสอบ เวลาเฉลี่ยในการตอบแต่ละข้อ จำนวนข้อสอบของแบบทดสอบ ลำดับข้อ ค่าพารามิเตอร์ข้อสอบ เฉลยข้อสอบ เวลาในการตอบแต่ละข้อ และไฟล์รูปภาพที่ใช้กับข้อสอบแต่ละข้อ

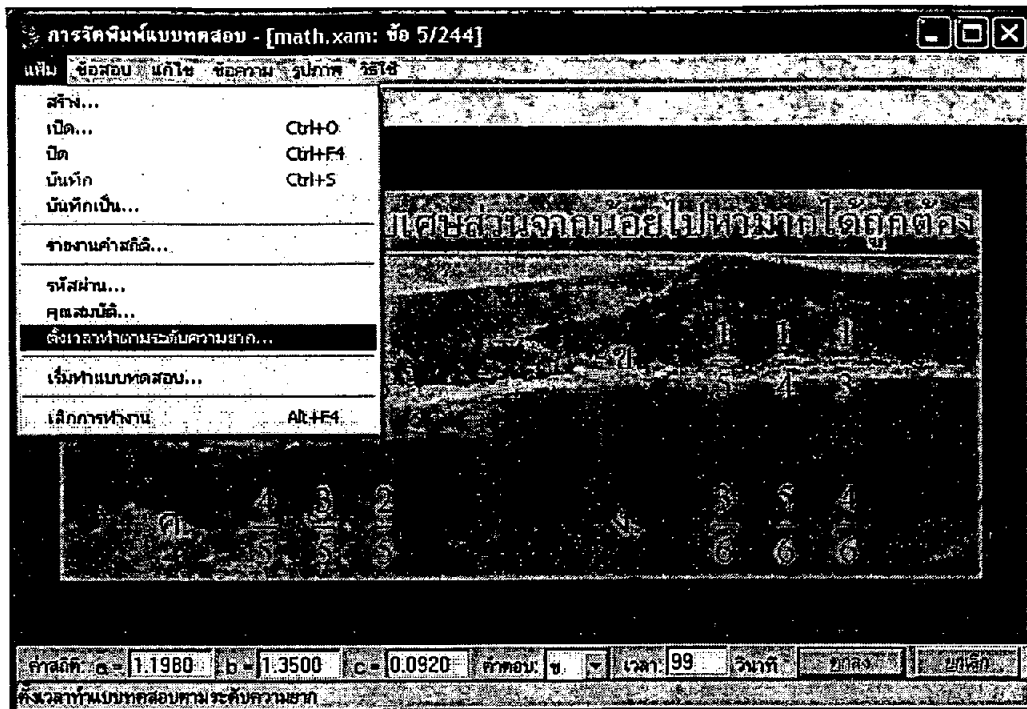


ภาพประกอบ 15 หน้าจอโปรแกรมคำสั่งแฟ้ม - รหัสผ่าน - คุณสมบัติ

เมื่อเลือกคำสั่งแฟ้ม และคลิกเมาส์คำสั่งรหัสผ่าน โปรแกรมจะให้กำหนดรหัสผ่านเข้าสู่แบบทดสอบตามที่ต้องการ และเมื่อคลิกคำสั่งคุณสมบัติ โปรแกรมจะให้เลือกคุณสมบัติของแบบทดสอบ ดังภาพประกอบ 16

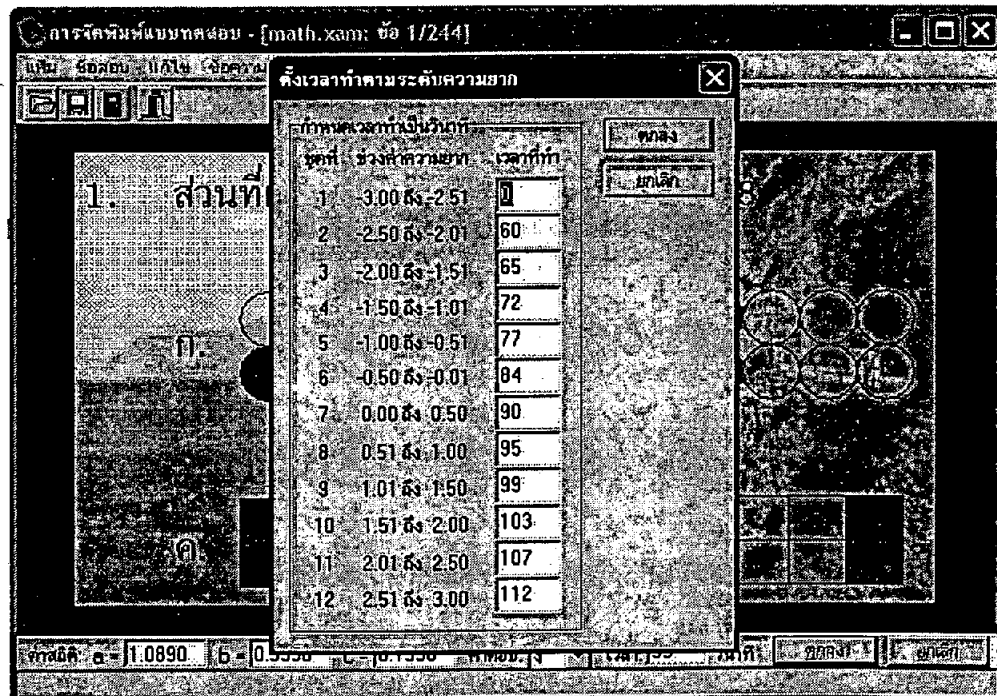


ภาพประกอบ 16 หน้าจอโปรแกรมสำหรับกำหนดคุณสมบัติของแบบทดสอบ

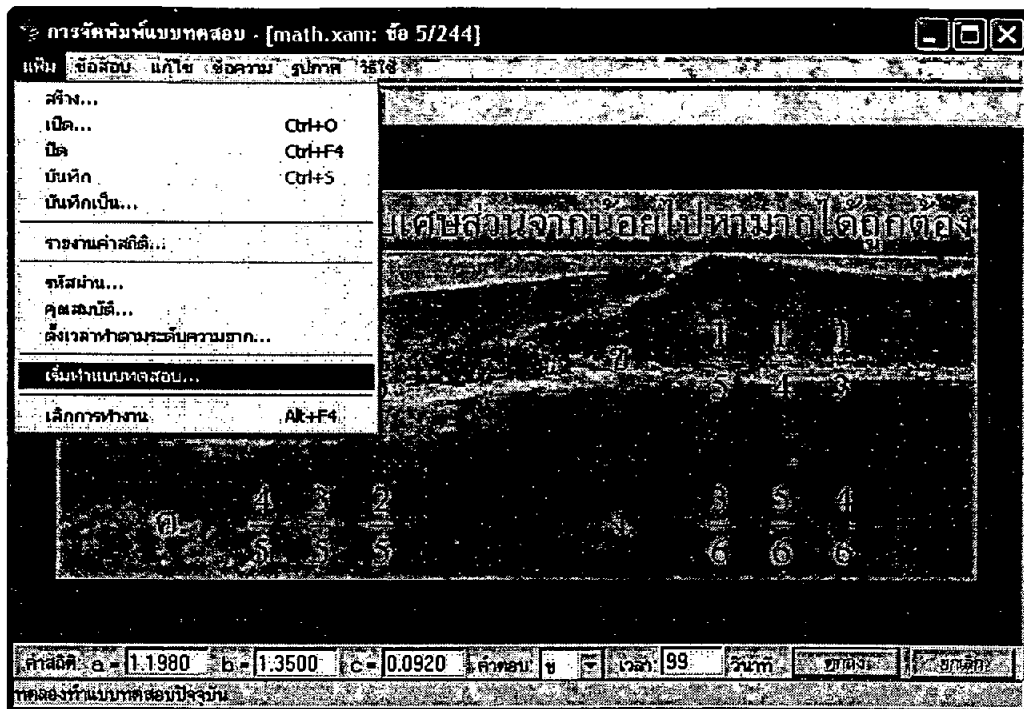


ภาพประกอบ 17 หน้าจอโปรแกรม คำสั่งแฟ้ม - ตั้งเวลาตามระดับความยาก

เมื่อเลือกคำสั่งแฟ้ม และคลิกเมาส์คำสั่งกำหนดเวลาตามระดับความยาก โปรแกรมจะให้กำหนดเวลาให้ตอบข้อสอบแต่ละข้อ ตามระดับความยากง่ายของข้อสอบ ซึ่งสามารถกำหนดเวลาเป็นวินาทีได้ 12 ค่า ตามช่วงความยากง่ายของแบบทดสอบ 12 ช่วง ดังภาพประกอบ 18

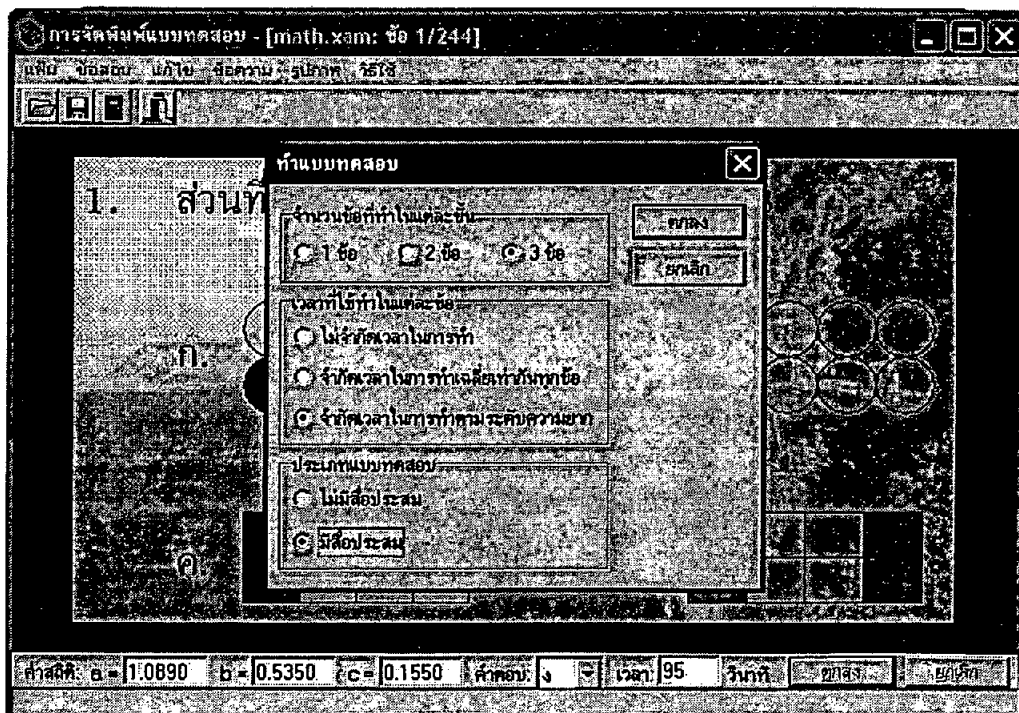


ภาพประกอบ 18 หน้าจอโปรแกรมสำหรับกำหนดเวลาให้ตอบในแต่ละข้อตามระดับความยาก



ภาพประกอบ 19 หน้าจอโปรแกรมคำสั่งแฟ้ม - เริ่มทำแบบทดสอบ

เมื่อเลือกคำสั่งแฟ้ม และคลิกเมาส์ที่คำสั่งเริ่มทำแบบทดสอบ โปรแกรมจะให้กำหนดรูปแบบชนิดของแบบทดสอบว่าจำ จะกำหนดจำนวนข้อในแต่ละชั้น เวลาที่ใช้ในการตอบแต่ละข้อ และรูปแบบของข้อสอบอย่างไร ดังภาพประกอบ 20



ภาพประกอบ 20 หน้าจอโปรแกรมสำหรับกำหนดรูปแบบชนิดของแบบทดสอบ

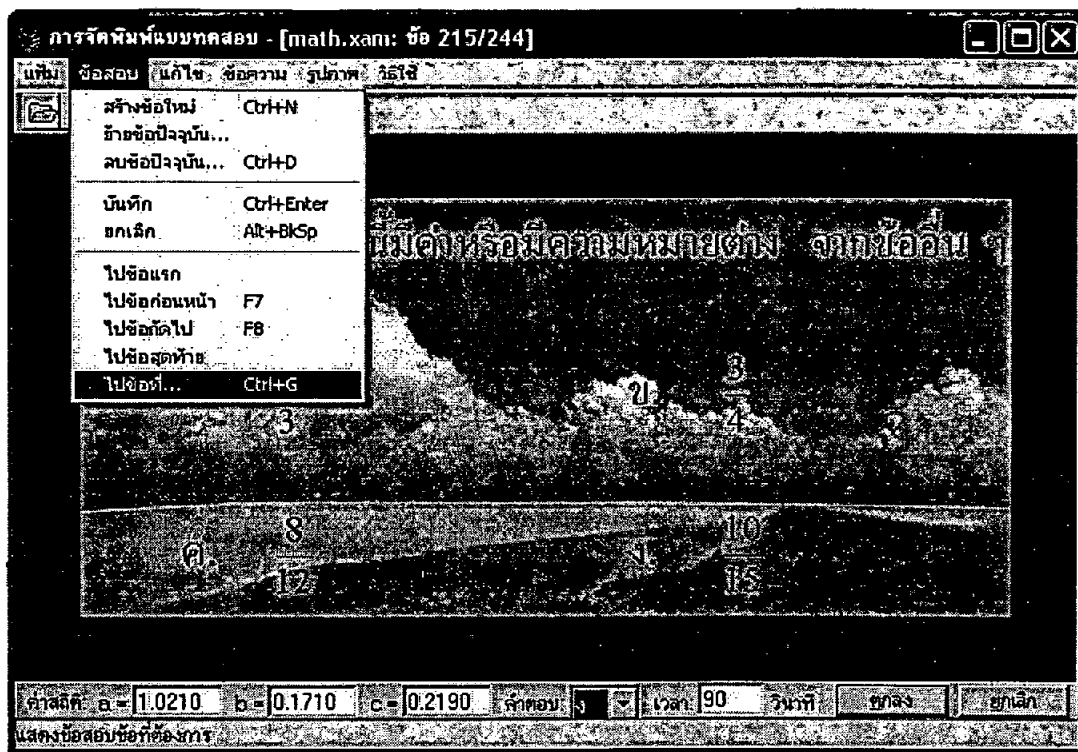


ภาพประกอบ 21 หน้าจอโปรแกรมคำสั่งข้อสอบ

เมื่อเลือกคำสั่งข้อสอบ จะปรากฏกลุ่มคำสั่งสร้างข้อสอบข้อใหม่ เพื่อสร้างข้อสอบเพิ่มเติม คำสั่งลบข้อปัจจุบันเพื่อลบข้อที่ไม่ต้องการออก และคำสั่งเพื่อทำการบันทึก ยกเลิก ไปข้อแรก ไปข้อก่อนหน้า ไปข้อถัดไป ไปข้อสุดท้าย และไปข้อที่ ... ตามที่ต้องการ และบางครั้งโปรแกรมจะให้ยืนยัน คำสั่งเดิมก่อนที่จะไปทำงานในคำสั่งต่อไป ดังภาพประกอบ 22

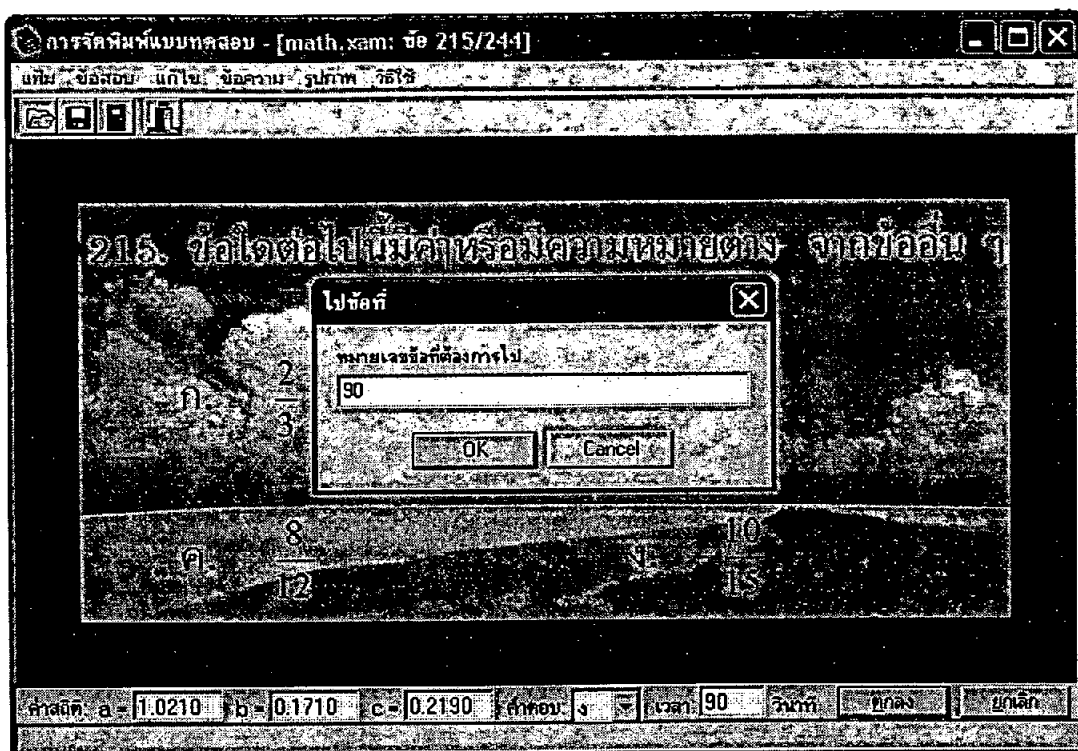


ภาพประกอบ 22 หน้าจอโปรแกรมสำหรับการยืนยันการเปลี่ยนแปลงแก้ไข

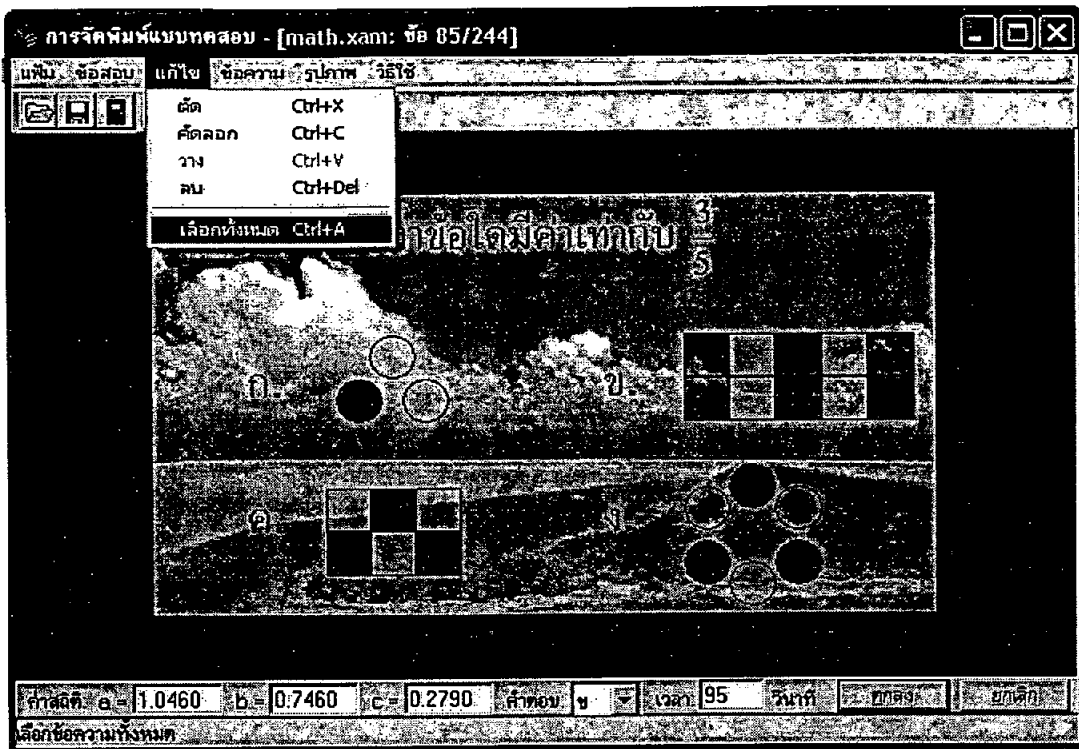


ภาพประกอบ 23 หน้าจอโปรแกรมคำสั่งข้อสอบ - ไปข้อที่

เมื่อเลือกคำสั่งข้อสอบ และคลิกเมาส์คำสั่งไปข้อที่ โปรแกรมจะให้เลือกข้อสอบในแบบทดสอบที่ต้องการค้นหา ดังภาพประกอบ 24 เมื่อต้องการค้นหาข้อสอบข้อที่ 90

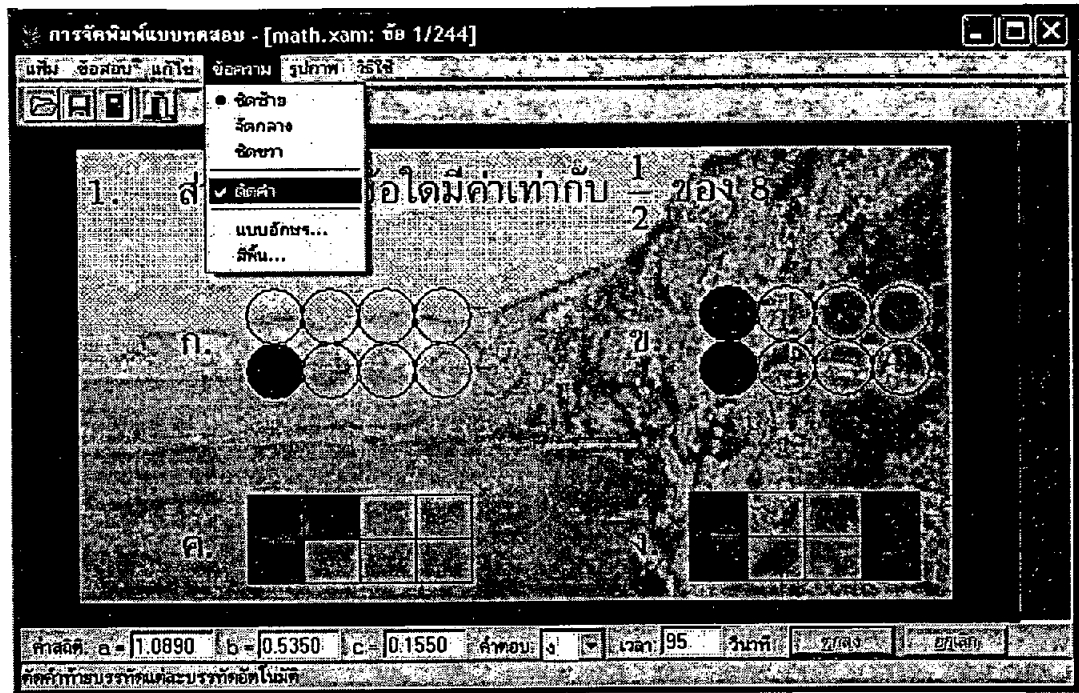


ภาพประกอบ 24 หน้าจอโปรแกรมสำหรับค้นหาข้อสอบ



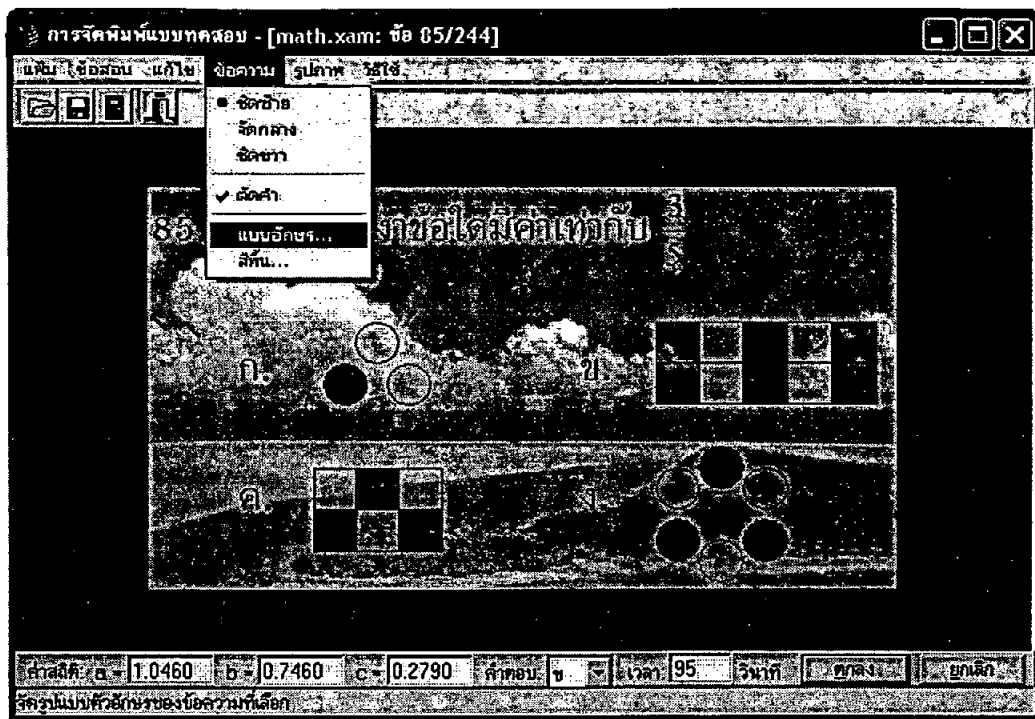
ภาพที่ 25 หน้าจอโปรแกรมคำสั่งแก้ไข

เมื่อเลือกคำสั่งแก้ไข จะประกอบด้วยกลุ่มของคำสั่งตัด คัดลอก วาง ลบ และเลือกทั้งหมด ซึ่งกลุ่มขอ คำสั่งเหล่านี้จะใช้กับการสร้างข้อสอบใหม่ หรือเปลี่ยนแปลงแก้ไขข้อสอบ



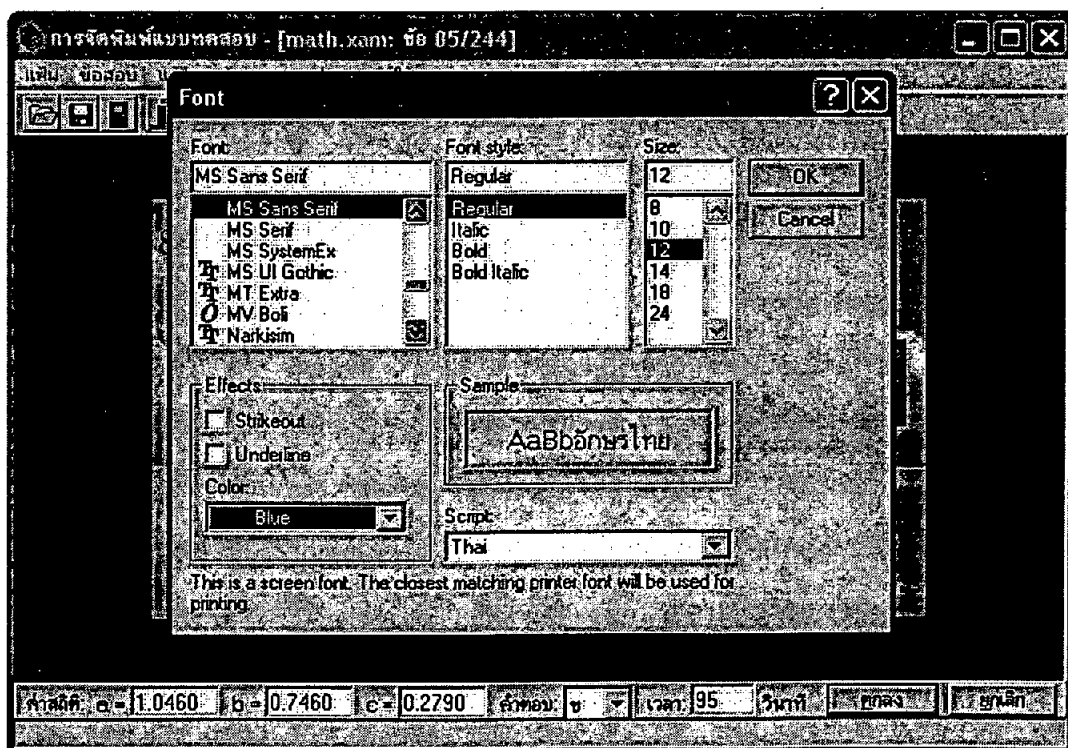
ภาพประกอบ 26 หน้าจอโปรแกรมคำสั่งข้อความ

เมื่อเลือกคำสั่งข้อความ คำสั่งจะประกอบด้วยกลุ่มของคำสั่งชิดซ้าย จัดกลาง ชิดขวา คัดคำ สำหรับจัดข้อความ หรือรูปภาพของข้อสอบ และมีคำสั่งแบบอักษร และสีพื้น ดังภาพประกอบ 27-30

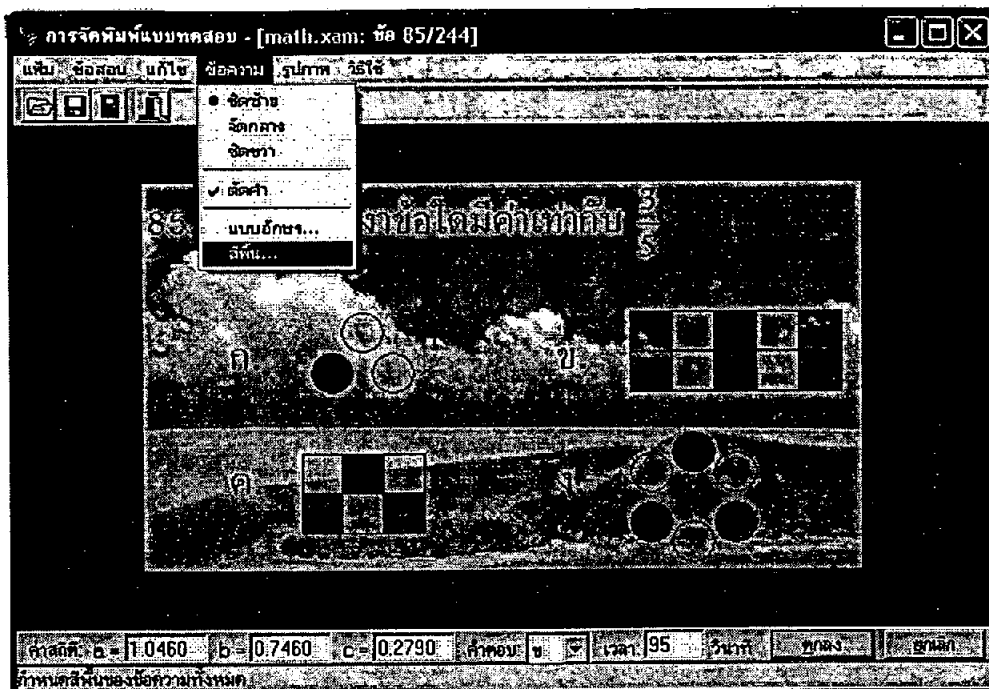


ภาพประกอบ 27 หน้าจอโปรแกรมคำสั่ง ข้อความ - แบบอักษร.....

เมื่อเลือกคำสั่งข้อความ และคลิกเมาส์ที่คำสั่งแบบอักษร... โปรแกรมจะให้เลือกรูปแบบของตัวอักษร สีของตัวอักษร และขนาดของตัวอักษรได้ตามต้องการ ดังภาพประกอบ 28

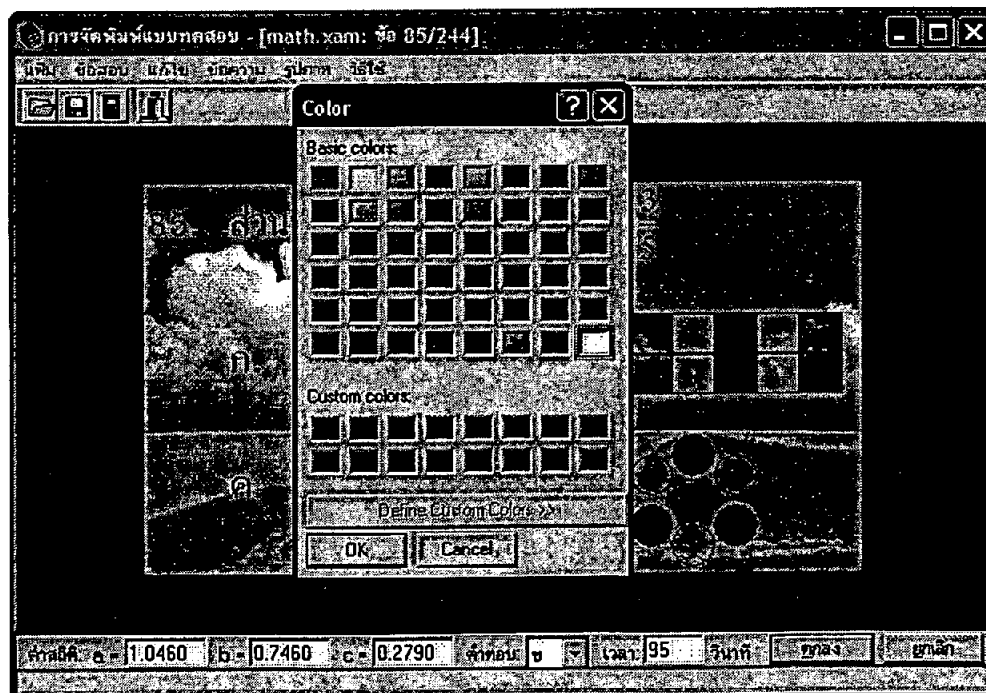


ภาพประกอบ 28 หน้าจอโปรแกรมรูปแบบอักษร

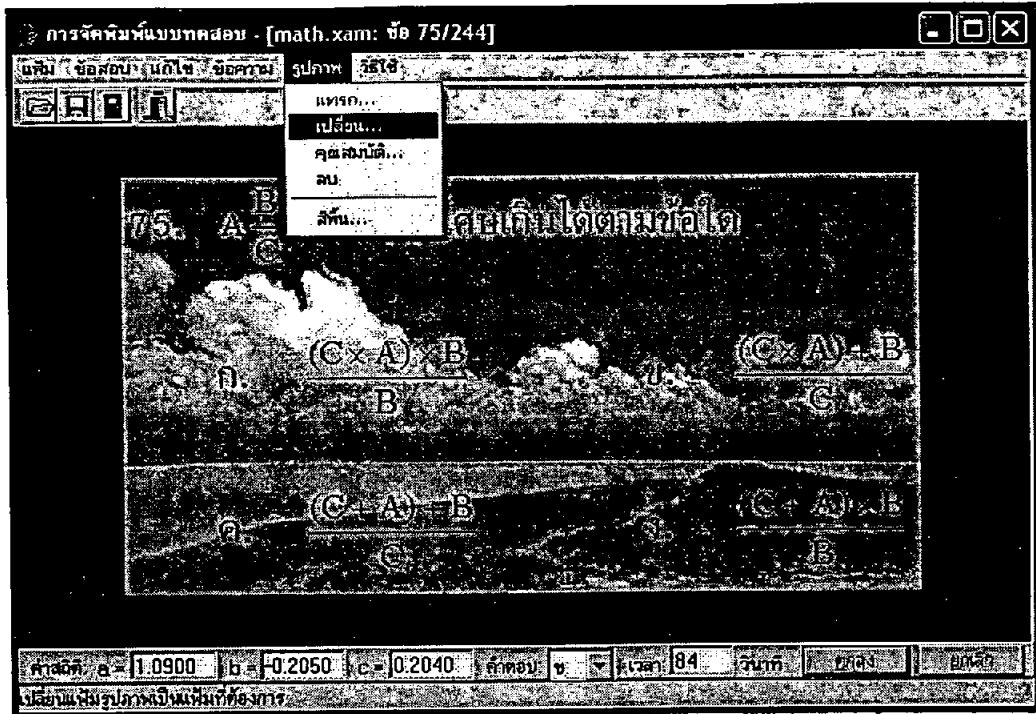


ภาพที่ 29 หน้าจอโปรแกรมคำสั่งสีพื้น

เมื่อเลือกคำสั่งข้อความ และคลิกเมาส์ที่คำสั่งสีพื้นโปรแกรมจะให้เลือกรูปแบบของข้อสอบตามที่
ต้องการ ดังภาพประกอบ 30

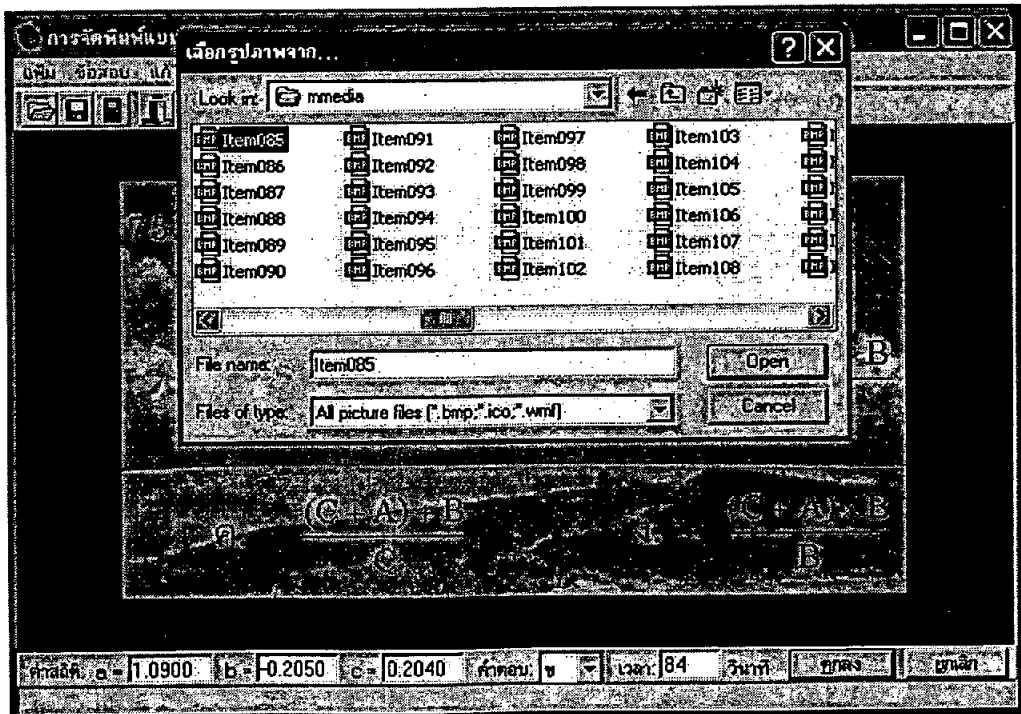


ภาพประกอบ 30 หน้าจอโปรแกรมคำสั่งสีพื้น

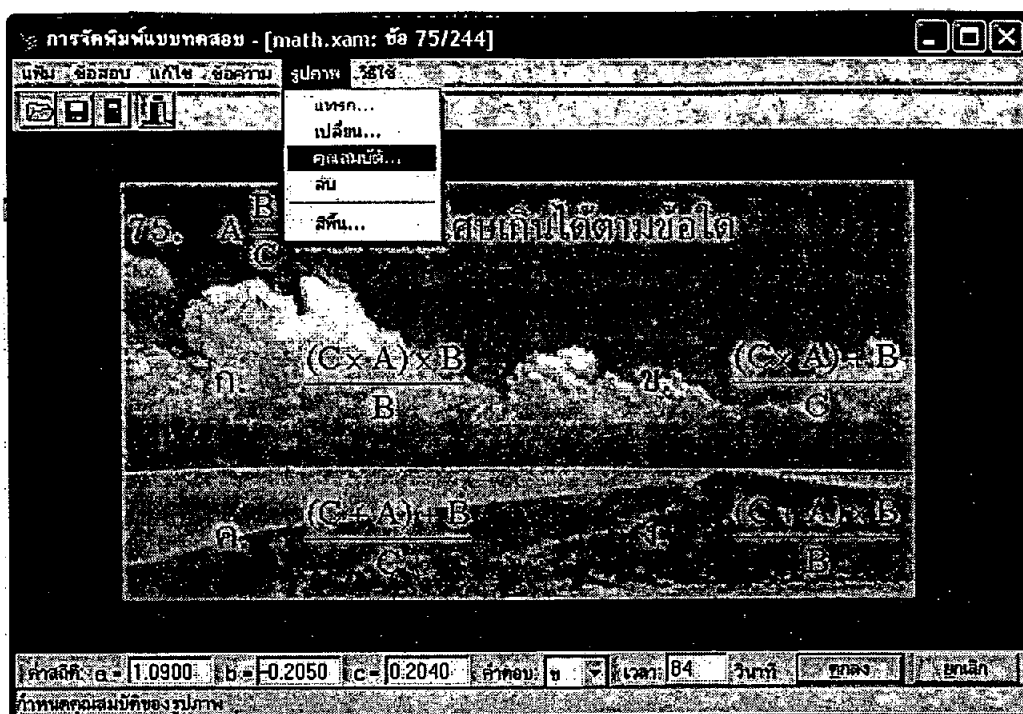


ภาพประกอบ 31 หน้าจอโปรแกรมคำสั่ง รูปภาพ

เมื่อเลือกคำสั่งรูปภาพ จะประกอบด้วยกลุ่มคำสั่ง แทรก.... เปลี่ยน..... คุณสมบัติ..... ลบ ลี ฟัน.... ซึ่งจากภาพประกอบ 31 เมื่อเลือกคำสั่งรูปภาพ และคลิกเมาส์ที่คำสั่งเปลี่ยน..... โปรแกรมจะ ให้เลือกรูปภาพ หรือข้อสอบที่อยู่ในไฟล์รูปภาพ สร้างข้อสอบใหม่ได้ ดังภาพประกอบ 32

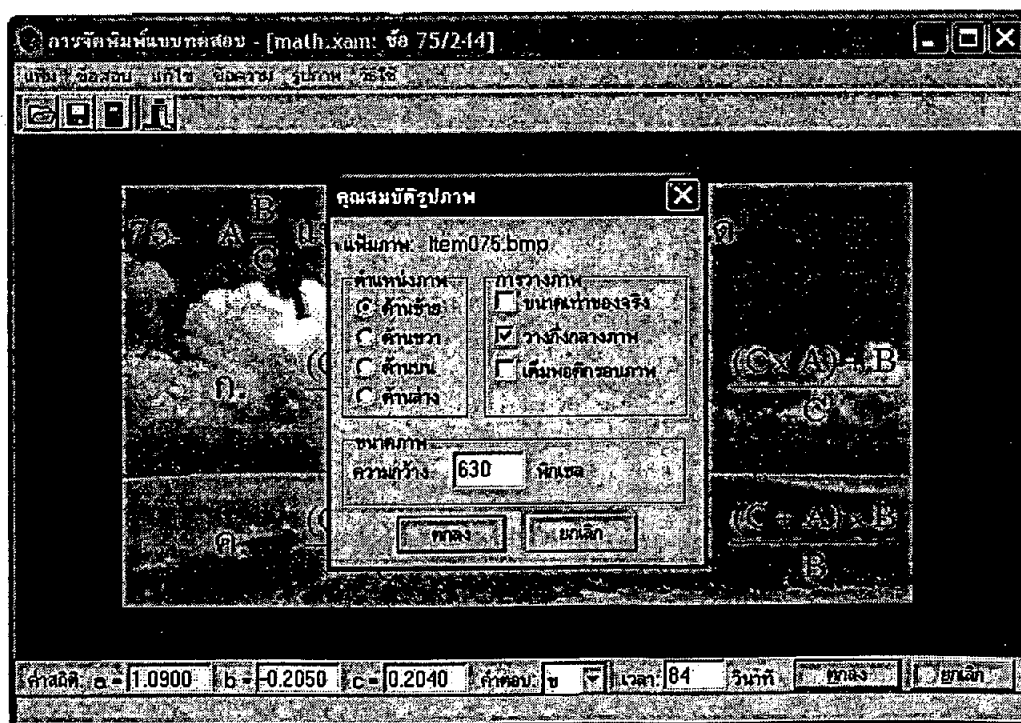


ภาพประกอบ 32 หน้าจอโปรแกรมคำสั่งเลือกจากรูปภาพ

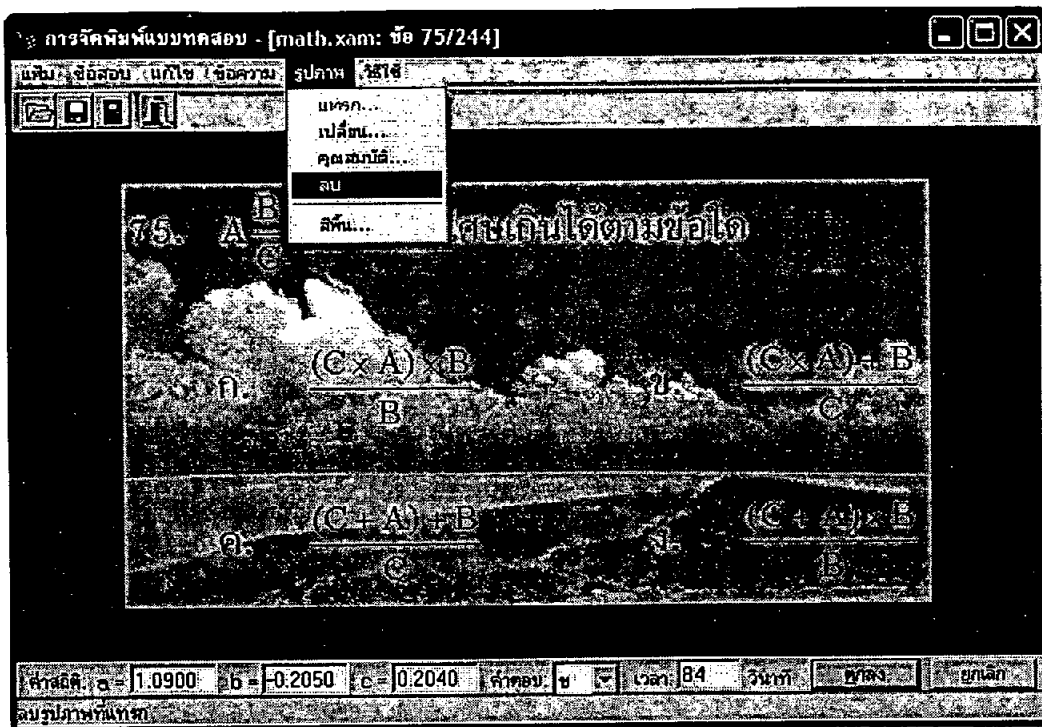


ภาพประกอบ 33 หน้าจอโปรแกรมคำสั่งรูปภาพ - คุณสมบัติ

เมื่อเลือกคำสั่งรูปภาพ และคลิกเมาส์ที่คำสั่งคุณสมบัติโปรแกรมจะให้เลือกคุณสมบัติของรูปภาพให้ได้ขนาด และตำแหน่งตามที่ต้องการ ดังภาพประกอบ 34

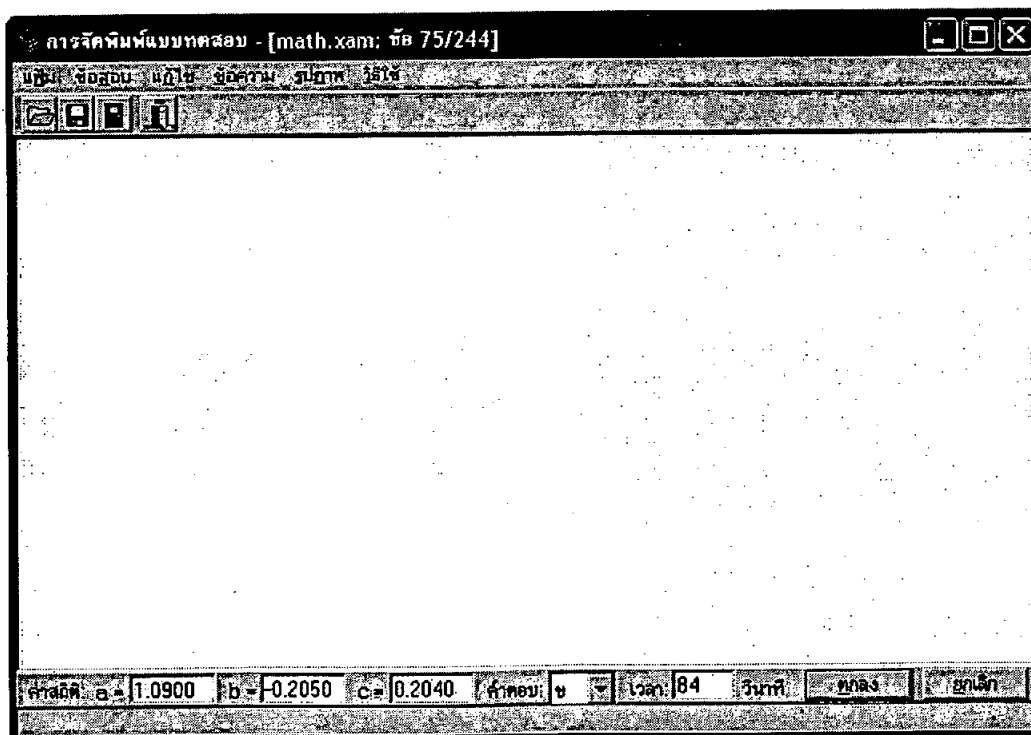


ภาพประกอบ 34 หน้าจอโปรแกรมคำสั่งคุณสมบัติภาพ

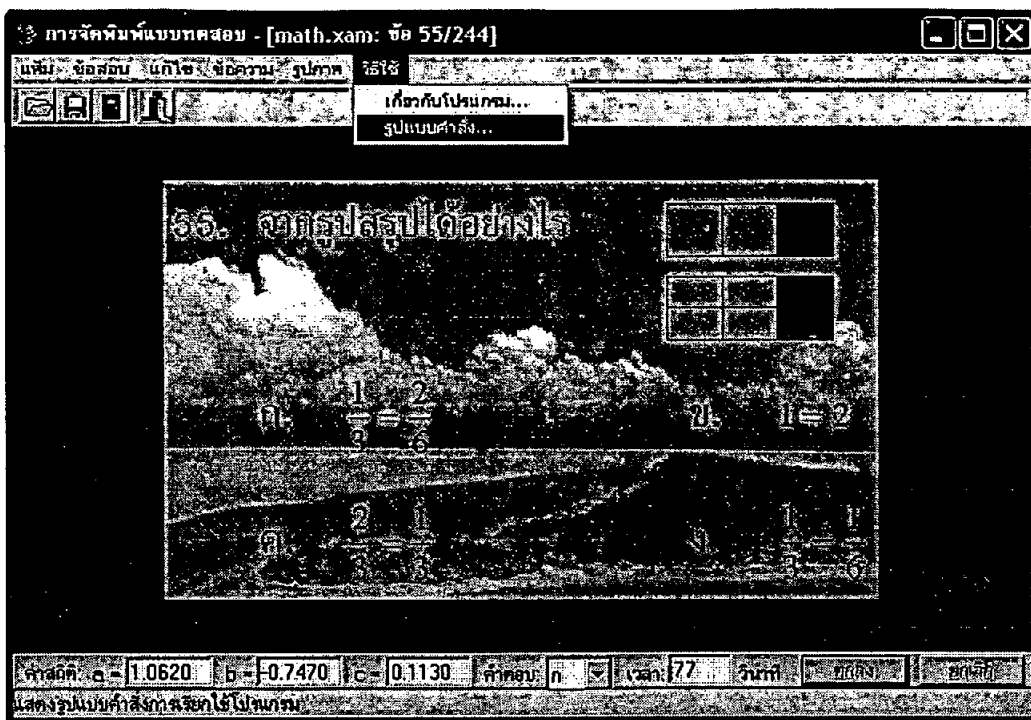


ภาพประกอบ 35 หน้าจอโปรแกรมคำสั่งรูปภาพ ลบ

เมื่อเลือกคำสั่งรูปภาพ และคลิกเมาส์ที่คำสั่งลบ โปรแกรมจะลบรูปภาพ หรือข้อสอบที่เป็นไฟล์รูปภาพที่ปรากฏอยู่ในขณะนั้นออกไปทั้งหมด ดังภาพประกอบ 36

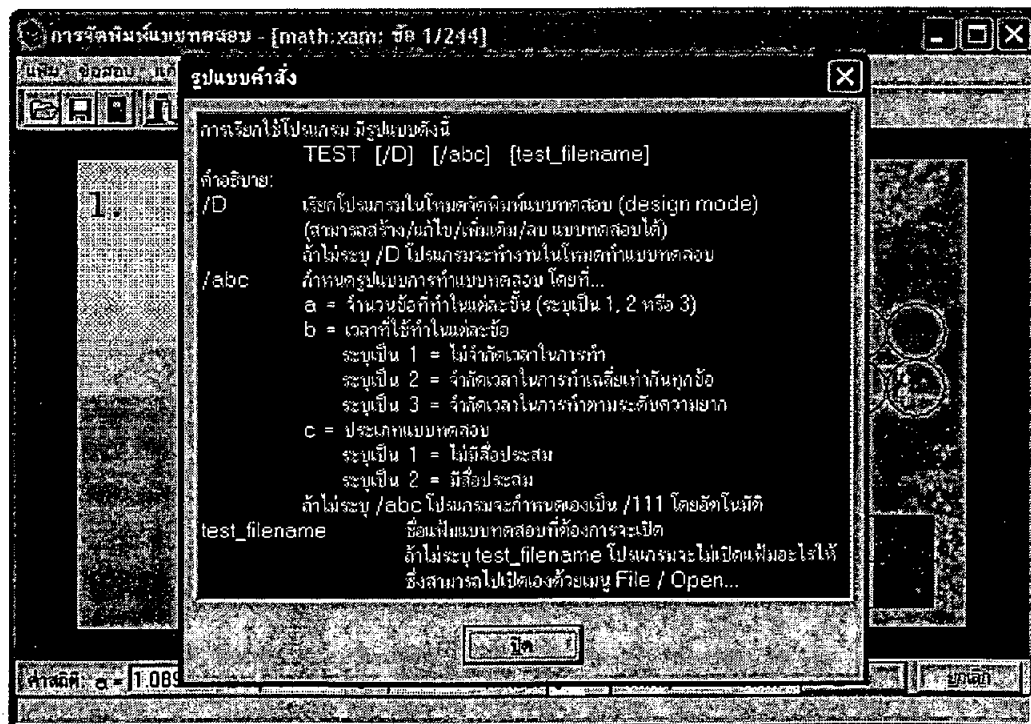


ภาพประกอบ 36 หน้าจอโปรแกรมเมื่อใช้คำสั่งรูปภาพ - ลบ



ภาพประกอบ 37 หน้าจอโปรแกรมวิธีใช้ - รูปแบบคำสั่ง

เมื่อเลือกคำสั่งวิธีใช้ และคลิกเมาส์ที่คำสั่งรูปแบบคำสั่ง โปรแกรมจะแสดงวิธีการเรียกใช้โปรแกรม และคำอธิบายในโปรแกรม ดังภาพประกอบ 38



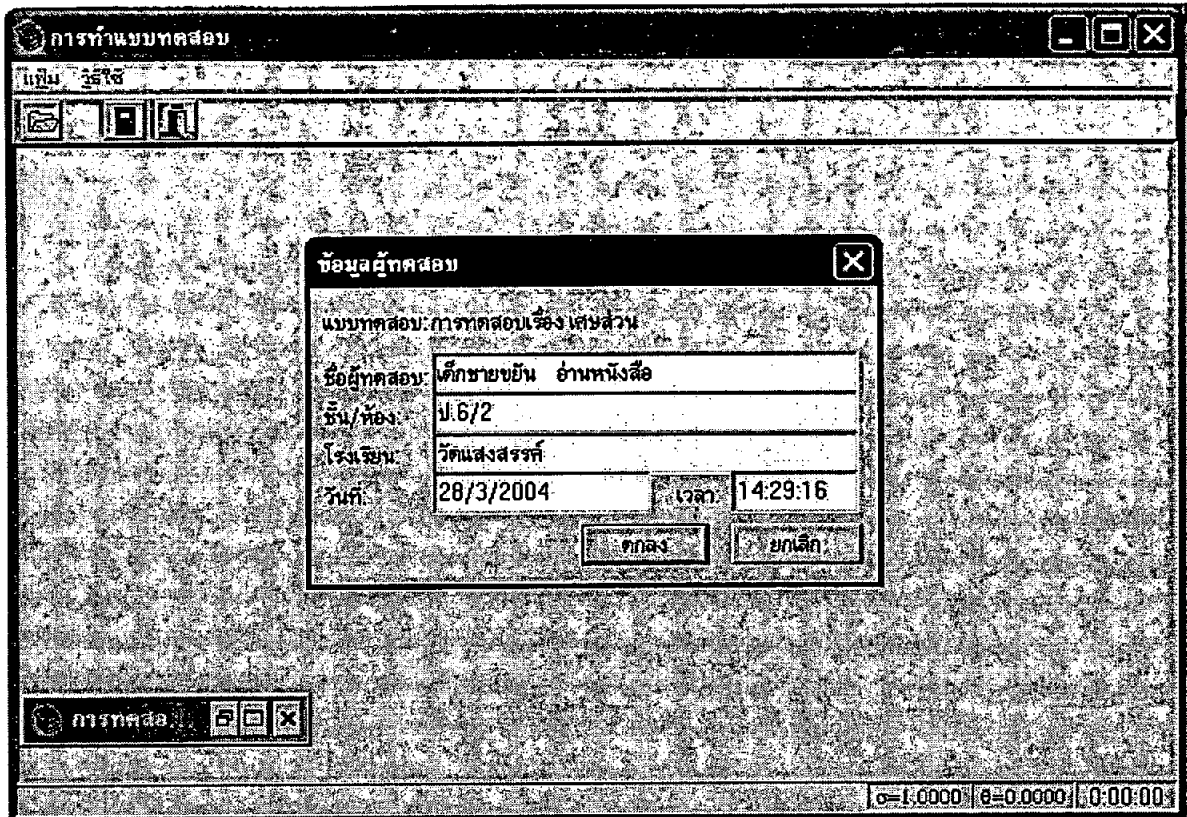
ภาพประกอบ 38 หน้าจอโปรแกรมรูปแบบคำสั่ง

**คู่มือโปรแกรม
สำหรับการทดสอบ**



ภาพประกอบ 39 โลโก้โปรแกรมสำหรับการทำแบบทดสอบ

เมื่อคลิกเมาส์ที่ปุ่มปิด หรือกด Enter ที่แป้นพิมพ์ โปรแกรมจะแสดงหน้าจอถัดไป เพื่อให้ข้อสอบป้อนข้อมูลส่วนตัว ดังภาพประกอบ 40



ภาพประกอบ 40 หน้าจอโปรแกรมสำหรับการทำแบบทดสอบ

หน้าจอโปรแกรมสำหรับการทำแบบทดสอบ ชื่อผู้สอบ ชั้น/ห้อง โรงเรียน วันที่ และเวลาเริ่มสอบ เมื่อผู้สอบป้อนข้อมูลส่วนตัวครบสมบูรณ์ แล้วคลิกเมาส์ที่ปุ่มตกลง โปรแกรมจะเริ่มให้ทำการทำแบบทดสอบด้วยข้อสอบข้อแรกที่มีความยากง่ายปานกลาง ดังภาพประกอบ 41



ภาพประกอบ 41 หน้าจอโปรแกรมการทำแบบทดสอบ

หน้าจอโปรแกรมการทำแบบทดสอบ จะบอกชื่อและชุดรูปแบบของแบบทดสอบ จำนวนข้อสอบที่กำลังทำขณะนั้น ซึ่งขณะที่ปรากฏนี้ เป็นการทดสอบวัดความเข้าใจเรื่องเศษส่วน ในชุดที่ 332 หมายความว่าจัดชุดข้อสอบให้ตอบชั้นละ 3 ข้อ กำหนดเวลาให้ตอบตามหน้านักความยากของข้อสอบและเป็นข้อสอบที่มีสื่อประสม ข้อที่ 1 ให้เลือกตอบตามข้อยลที่ปรากฏบนขอบล่างของหน้าต่าง และจะมีช่องบอกเวลา ซึ่งขณะนี้เหลือเวลาสำหรับตอบข้อที่ 1 เท่ากับ 79 วินาที ซึ่งยังมีเวลามากพอในการคิดหาคำตอบช่องบอกเวลานี้จึงแสดงด้วยสีเขียว



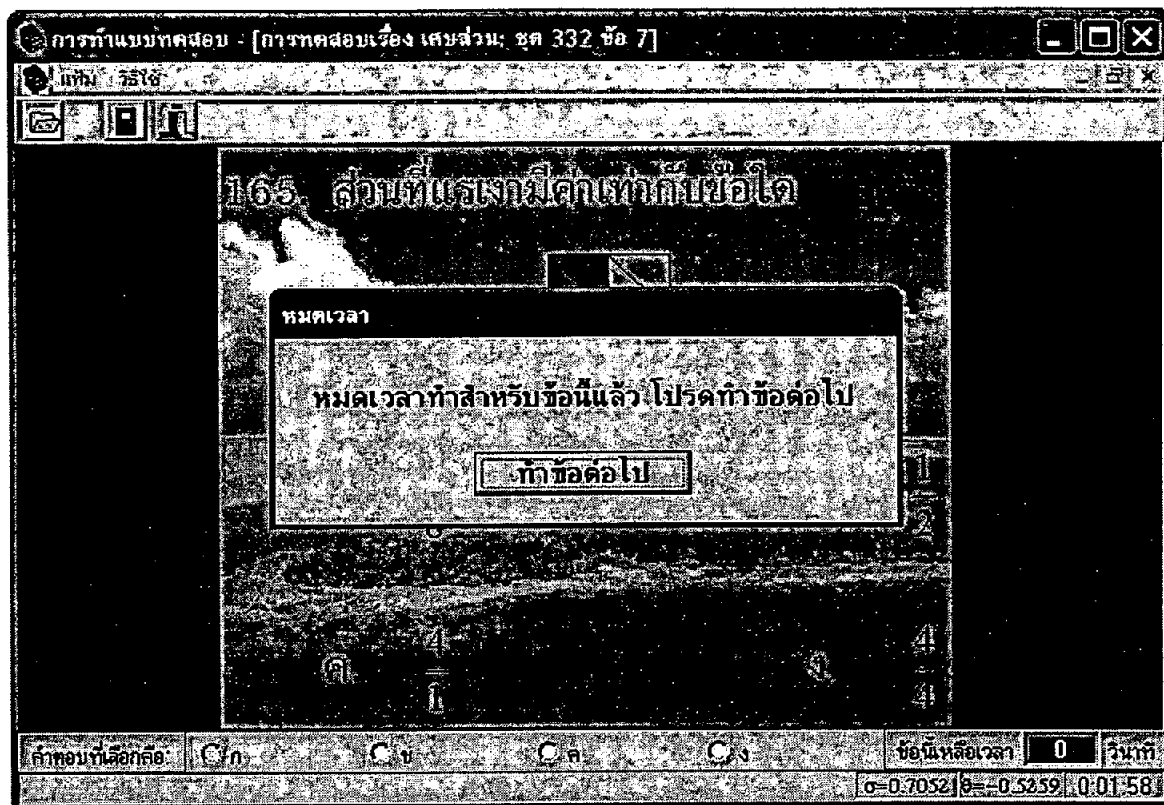
ภาพประกอบ 42 หน้าจอโปรแกรมการทำแบบทดสอบข้อที่ 6

ขณะที่ผู้สอบทำการทดสอบในข้อที่ 6 ซึ่งโปรแกรมคัดเลือกข้อสอบ ข้อที่ 57 ในคลังข้อสอบให้ทำการทดสอบ และขณะนี้เหลือเวลาสำหรับการทำข้อสอบนี้ 22 วินาที ซึ่งน้อยกว่า 30 วินาที ช่องบอกเวลานี้จึงแสดงด้วยสีเหลือง



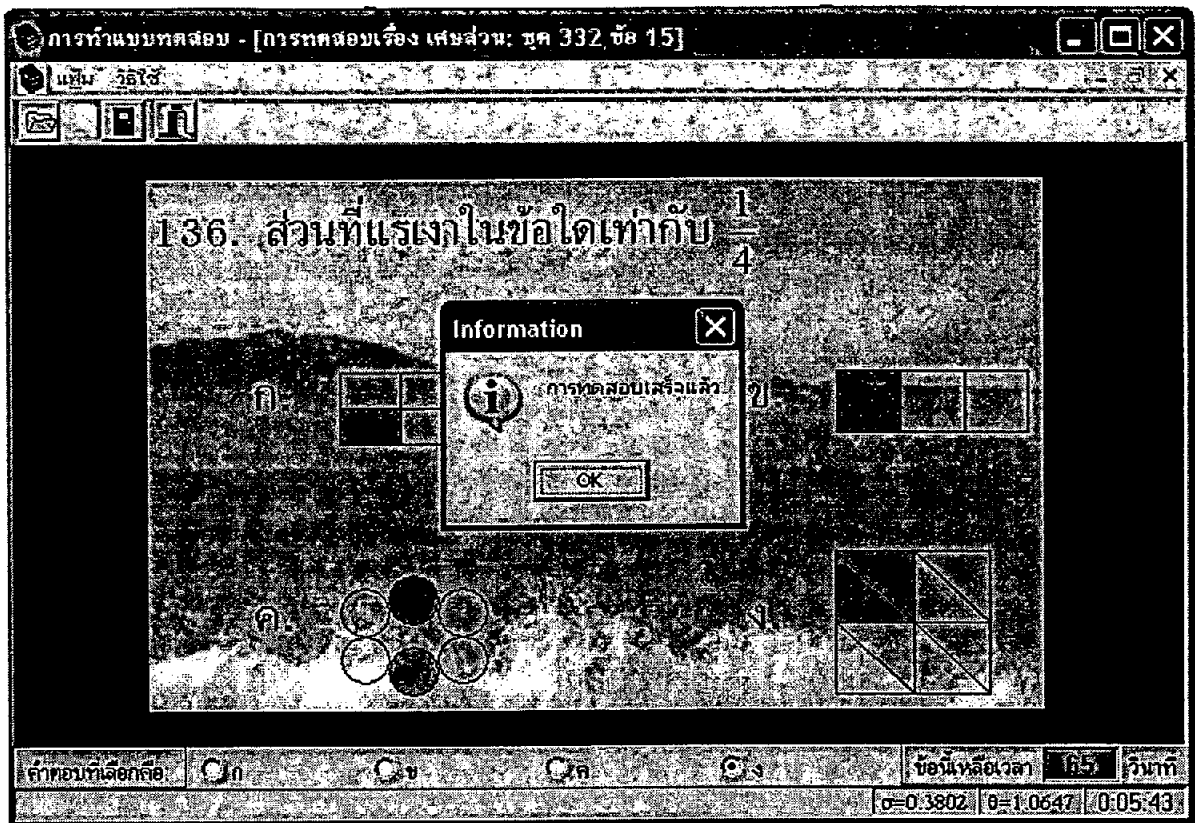
ภาพประกอบ 43 หน้าจอโปรแกรมทำแบบทดสอบข้อที่ 7

ขณะที่ผู้สอบทำการทดสอบในข้อที่ 7 ซึ่งโปรแกรมคัดเลือกข้อสอบข้อที่ 165 ในคลังข้อสอบ ที่ทำการทดสอบ และขณะนี้เหลือเวลาสำหรับการทำข้อสอบข้อนี้เพียง 7 วินาที ซึ่งน้อยกว่า 10 วินาที ช่องบอกเวลานี้จึงแสดงด้วยสีแดง ซึ่งเป็นการเตือนบอกเวลาช่วงสุดท้าย



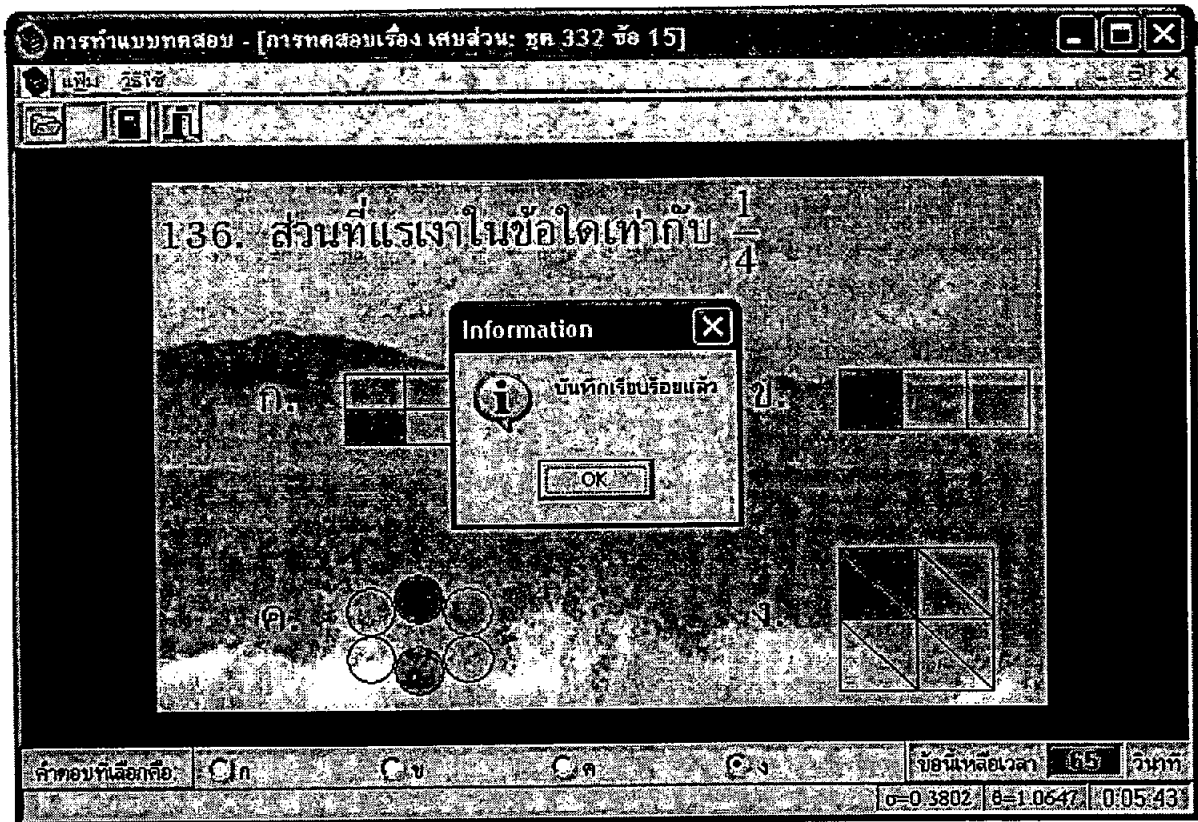
ภาพประกอบ 44 หน้าจอโปรแกรมการทำแบบทดสอบ ข้อที่ 7

จากหน้าจอโปรแกรมที่ปรากฏนี้ แสดงว่าผู้สอบไม่สามารถตอบข้อสอบเสร็จทันตามเวลาที่กำหนด ผู้สอบจึงไม่สามารถตอบคำถามข้อนี้ได้อีก ในขณะเดียวกันโปรแกรมก็จะหยุดเวลาในการทำแบบทดสอบไว้ด้วย ผู้สอบจะต้องคลิกเมาส์ที่ปุ่ม “ทำข้อต่อไป” โปรแกรมจึงจะนำเสนอข้อสอบข้อถัดไป พร้อมทั้งเริ่มจับเวลาทำแบบทดสอบเพิ่มอีกครั้ง



ภาพประกอบ 45 หน้าจอโปรแกรมเมื่อทำการทดสอบเสร็จสิ้นลง

เมื่อผู้สอบทำการทดสอบไปถึงเกณฑ์ยุติการสอบลง และหน้าจอของโปรแกรมจะแสดงข้อความ "การทดสอบเสร็จแล้ว" ผู้สอบจะต้องใช้เมาส์คลิกที่ปุ่ม OK เพื่อทำการบันทึกผลการสอบ
 ดังภาพประกอบ 46



ภาพประกอบ 46 หน้าจอโปรแกรมเมื่อบันทึกผลการสอบเรียบร้อยแล้ว

เมื่อทำการบันทึกผลการทดสอบเสร็จสิ้นลง และใช้เมาส์คลิกปุ่ม OK โปรแกรมจะประมวลผลการทดสอบ และรายงานผลการสอบทันที ดังภาพประกอบ 47

ผลการทดสอบ

แบบทดสอบ: ภาพทดสอบเรื่อง เคนส์ว้า

ชื่อผู้ทดสอบ: เด็กชายชน อานทองสี

ชั้น/ห้อง: ป.6/2 โรงเรียน: วัดแสงสวรรค์

วันที่: 28/3/2004 เวลา: 14:29:16

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 332

เวลาที่ใช้ในการทดสอบ: 0:05:43

จำนวนข้อที่ตอบถูก: 10 จาก 15

ตำแหน่งเปอร์เซ็นต์ไทล์: 85

ความสามารถในการสอบ: 1.0647 (สูง)

สารสนเทศของแบบทดสอบ: 11.3295

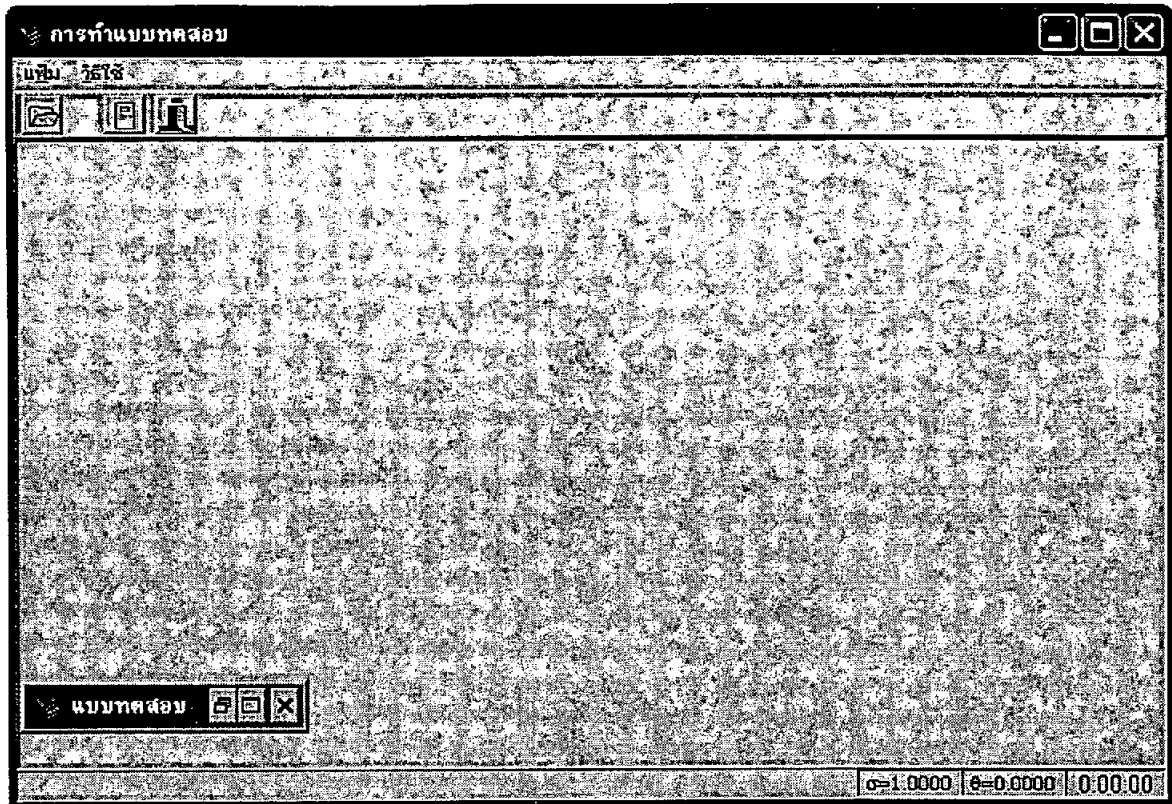
ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2971

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I-Zeta	SEE
7	165	ข	ถูก	0.2738	3.7288	0.5179
8	173	ค	ถูก	0.6596	5.0912	0.4432
9	58	ง	ถูก	0.9420	6.3715	0.3962
10	157	ค	ผิด	0.7206	6.6591	0.3875
11	131	ค	ผิด	0.5365	7.2203	0.3722
12	94	ง	ถูก	0.7417	7.8521	0.3569
13	83	ก	ถูก	0.8547	8.9710	0.3339
14	194	ข	ถูก	0.9766	10.3037	0.3115
15	136	ง	ถูก	1.0647	11.3295	0.2971

ค่ารวมที่: 0=0.3802 0=1.0647 0:05:43

ภาพประกอบ 47 หน้าจอโปรแกรม แสดงผลการทดสอบ

ผลการทดสอบที่ปรากฏ ประกอบด้วย ชื่อแบบทดสอบ ชื่อ-สกุล ผู้สอบ ชั้นเรียน โรงเรียน วันที่ และเวลาเริ่มต้นสอบ ส่วนผลการทดสอบประกอบด้วย ชุดของแบบทดสอบที่ทำ เวลาที่ใช้ในการสอบ จำนวนข้อสอบที่ตอบถูกและจำนวนข้อที่ทำทั้งหมด ตำแหน่งเปอร์เซ็นต์ไทล์ ค่าความสามารถของผู้สอบ ค่าฟังก์ชันสารสนเทศของแบบทดสอบ ค่าความคลาดเคลื่อนในการประมาณค่าความสามารถ ลำดับข้อที่ทำ ลำดับข้อสอบในคลัง ตัวเลือกที่เลือกตอบ ตรวจคำตอบ ค่าความสามารถ ค่าฟังก์ชันสารสนเทศของข้อสอบ และค่าฟังก์ชันสารสนเทศของแบบทดสอบ ณ ข้อที่ทำเสร็จขณะนั้น ๆ



ภาพประกอบ 48 หน้าจอโปรแกรมหลังจากสิ้นสุดการทดสอบ

เมื่อต้องการเริ่มทำแบบทดสอบ ผู้สอบจะต้องเลือกคำสั่งแฟ้มเพื่อเปิดไฟล์แบบทดสอบตามลำดับที่กล่าวมา

ภาคผนวก ค
ตัวอย่างกระดาษคำตอบของผู้สอบ
ค่าพารามิเตอร์ของข้อสอบในคลัง

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: อาทิตย์ บุษบิน
 ชั้น/ห้อง: ป6/1 โรงเรียน: บ.ก.ว.
 วันที่: 25/12/2003 เวลา: 11:25:17

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 121
 เวลาที่ใช้ในการทดสอบ: 0:03:14
 จำนวนข้อที่ตอบถูก: 6 จาก 13
 ตำแหน่งเปอร์เซ็นต์: 65
 ความสามารถในการสอบ: 0.3959 (ปานกลาง)
 สารสนเทศของแบบทดสอบ: 11.1618
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2993

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
			-			
1	53	ก	ถูก	0.3647	0.2829	1.8801
2	164	ข	ถูก	0.7111	0.9527	1.0245
3	165	ข	ถูก	1.1574	2.4726	0.6359
4	157	ค	ผิด	0.7926	2.9103	0.5862
5	58	ง	ถูก	0.9684	4.1498	0.4909
6	131	ข	ผิด	0.7278	5.1565	0.4404
7	173	ก	ผิด	0.5649	6.2463	0.4001
8	132	ค	ถูก	0.7085	7.0024	0.3779
9	194	ค	ผิด	0.5544	8.2325	0.3485
10	94	ข	ผิด	0.5116	8.5221	0.3426
11	79	ข	ถูก	0.5549	9.3279	0.3274
12	86	ข	ผิด	0.4751	10.2025	0.3131
13	136	ข	ผิด	0.3959	11.1618	0.2993

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: สุนทรศิลป์ พึ่งจิตตน
 ชั้น/ห้อง: 6/1 โรงเรียน: ชุมชนวัดบัวแก้วเกษร
 วันที่: 25/12/03 เวลา: 10:48:59
 ผลการทดสอบ:

ผลการทดสอบ:						
ชุดแบบทดสอบที่ทำ:		231				
เวลาที่ใช้ในการทดสอบ:		0:19:41				
จำนวนข้อที่ตอบถูก:		14 จาก 28				
ตำแหน่งเปอร์เซ็นต์ไทล์:		51				
ความสามารถในการสอบ:		0.0382 (ปานกลาง)				
สารสนเทศของแบบทดสอบ:		12.1851				
ความคลาดเคลื่อนมาตรฐานในการวัด:		0.2865				
ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	Zeta	SEE
1	53	ก	ถูก	0.3647	0.2829	1.8801
2	35	ง	ถูก	0.5623	0.4126	1.5569
3	164	ก	ผิด	-0.2510	1.2584	0.8914
4	79	ข	ถูก	0.0013	2.2925	0.6605
5	165	ง	ผิด	-0.1559	2.4235	0.6424
6	173	ก	ผิด	-0.2221	2.4547	0.6383
7	212	ค	ถูก	-0.0495	3.7105	0.5191
8	76	ข	ผิด	-0.3432	4.8764	0.4528
9	157	ง	ผิด	-0.3446	4.8765	0.4528
10	131	ง	ผิด	-0.3572	4.8822	0.4526
11	126	ก	ผิด	-0.5128	5.7441	0.4172
12	125	ก	ผิด	-0.5758	6.0962	0.4050
13	58	ก	ผิด	-0.5942	6.1348	0.4037
14	194	ข	ถูก	-0.5410	6.1594	0.4029
15	211	ข	ถูก	-0.4345	6.4978	0.3923
16	109	ก	ถูก	-0.3263	7.3452	0.3690
17	83	ข	ผิด	-0.3644	7.4444	0.3665
18	136	ง	ถูก	-0.2767	7.5849	0.3631
19	57	ข	ผิด	-0.3335	7.9324	0.3551
20	209	ค	ถูก	-0.2334	8.4158	0.3447
21	218	ก	ถูก	-0.1551	8.7437	0.3382
22	167	ก	ถูก	-0.0341	9.2759	0.3283
23	133	ง	ถูก	0.0806	9.7807	0.3198
24	139	ก	ผิด	-0.0114	10.4058	0.3100
25	214	ค	ผิด	-0.0684	10.6493	0.3064
26	86	ข	ผิด	-0.0998	10.8040	0.3042
27	156	ง	ถูก	-0.0339	11.5114	0.2947
28	124	ค	ถูก	0.0382	12.1851	0.2865

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: ด.ญ.ชมพูนุช ปิยะธรรมชัย
 ชั้น/ห้อง: ป.6/2 โรงเรียน: ร.ร.บ.ก.ว.
 วันที่: 25/12/03 เวลา: 1:37:11

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 331
 เวลาที่ใช้ในการทดสอบ: 0:05:37
 จำนวนข้อที่ตอบถูก: 12 จาก 18
 ตำแหน่งเปอร์เซ็นต์ไทล์: 74
 ความสามารถในการสอบ: 0.6550 (ปานกลาง)
 สारสนเทศของแบบทดสอบ: 14.9199
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2589

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53		ผิด	-0.4623	0.2480	2.0080
2	35	ง	ถูก	-0.2407	0.3706	1.6426
3	122	ง	ผิด	-0.6303	0.6216	1.2683
4	164	ข	ถูก	-0.2565	1.4599	0.8276
5	79	ข	ถูก	0.0491	2.5238	0.6295
6	211	ข	ถูก	0.3563	3.5007	0.5345
7	165	ข	ถูก	0.7760	5.6312	0.4214
8	173	ค	ถูก	1.0731	7.4777	0.3657
9	131	ง	ผิด	0.7389	8.5123	0.3427
10	157	ข	ผิด	0.6496	8.6971	0.3391
11	187	ข	ผิด	0.6094	8.7589	0.3379
12	41	ข	ถูก	0.7405	8.9508	0.3342
13	212	ค	ถูก	0.7693	9.2117	0.3295
14	76	ข	ผิด	0.3092	9.9573	0.3169
15	57	ก	ถูก	0.3959	11.0199	0.3012
16	58	ง	ถูก	0.4928	12.4232	0.2837
17	194	ข	ถูก	0.5811	13.6927	0.2702
18	136	ง	ถูก	0.6550	14.9199	0.2589

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: ด.ญ. พนมภรณ์ นิลเทพ
 ชั้น/ห้อง: ป 6/1 โรงเรียน: วัดแสงสวรรค์
 วันที่: 27/01/04 เวลา: 9:21:00

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 212
 เวลาที่ใช้ในการทดสอบ: 0:10:09
 จำนวนข้อที่ตอบถูก: 9 จาก 20
 ตำแหน่งเปอร์เซ็นต์ไทล์: 52
 ความสามารถในการสอบ: 0.0510 (ปานกลาง)
 สารสนเทศของแบบทดสอบ: 12.2357
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2859

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53	ข	ผิด	-0.4623	0.2480	2.0080
2	35	ง	ถูก	-0.2407	0.3706	1.6426
3	164	ข	ถูก	0.1868	1.5267	0.8093
4	79	ข	ถูก	0.4684	2.4277	0.6418
5	173	ข	ผิด	0.1589	2.6648	0.6126
6	165	ค	ผิด	-0.0180	2.9263	0.5846
7	57	ก	ถูก	0.2135	3.8734	0.5081
8	58	ก	ผิด	0.0481	4.4461	0.4743
9	194	ข	ถูก	0.2232	5.0626	0.4444
10	86	ง	ถูก	0.4124	5.8327	0.4141
11	136	ง	ถูก	0.5824	7.0131	0.3776
12	83	ค	ผิด	0.3741	7.9985	0.3536
13	157	ข	ผิด	0.3522	8.0219	0.3531
14	131	ข	ผิด	0.2945	8.2273	0.3486
15	212	ข	ผิด	0.0394	9.3675	0.3267
16	76	ก	ถูก	0.0914	10.3961	0.3101
17	132	ก	ผิด	0.0738	10.4778	0.3089
18	94	ข	ผิด	0.0634	10.5169	0.3084
19	125	ค	ถูก	0.1267	11.3943	0.2962
20	211	ค	ผิด	0.0510	12.2357	0.2859

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: ด.ช จิรวัดน์ ทิพย์อิน
 ชั้นห้อง: ป. 6/1 โรงเรียน: วัดแสงสวรรค์
 วันที่: 27/01/04 เวลา: 9:34:20

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 212
 เวลาที่ใช้ในการทดสอบ: 0:13:39
 จำนวนข้อที่ตอบถูก: 10 จาก 22
 ตำแหน่งเปอร์เซ็นต์: 49
 ความสามารถในการสอบ: -0.0206 (ปานกลาง)
 สารสนเทศของแบบทดสอบ: 11.5299
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2945

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	Zeta	SEE
1	53	ก	ถูก	0.3647	0.2829	1.8801
2	35	ง	ถูก	0.5623	0.4126	1.5569
3	164	ก	ผิด	-0.2510	1.2584	0.8914
4	79	ข	ถูก	0.0013	2.2925	0.6605
5	165	ง	ผิด	-0.1559	2.4235	0.6424
6	173	ค	ถูก	0.0736	2.5799	0.6226
7	212	ค	ถูก	0.2501	3.4049	0.5419
8	76	ก	ถูก	0.3668	4.0760	0.4953
9	58	ก	ผิด	0.1518	4.8456	0.4543
10	194	ก	ผิด	0.0461	5.1913	0.4389
11	83	ก	ถูก	0.1773	5.8634	0.4130
12	136	ง	ถูก	0.3346	6.7298	0.3855
13	86	ข	ผิด	0.2343	7.2205	0.3721
14	70	ก	ผิด	0.1665	7.5690	0.3635
15	57	ง	ผิด	0.0660	8.3635	0.3458
16	218	ข	ผิด	-0.0185	8.8593	0.3360
17	133	ง	ถูก	0.0750	9.3573	0.3269
18	214	ค	ผิด	0.0169	9.6846	0.3213
19	167	ข	ผิด	-0.0372	10.2134	0.3129
20	211	ข	ถูก	0.0351	11.0407	0.3010
21	234	ก	ผิด	0.0000	11.3673	0.2966
22	110	ง	ผิด	-0.0206	11.5299	0.2945

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: ด.ช. ศุภกร หุ่นดี
 ชั้นห้อง: ป6/4 โรงเรียน: วัดหงส์ปทุมมาวาส
 วันที่: 26/1/2004 เวลา: 10:58:49

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 222
 เวลาที่ใช้ในการทดสอบ: 0:07:12
 จำนวนข้อที่ตอบถูก: 8 จาก 18
 ตำแหน่งเปอร์เซ็นต์: 53
 ความสามารถในการสอบ: 0.1000 (ปานกลาง)
 สารสนเทศของแบบทดสอบ: 11.8681
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2903

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53	ก	ถูก	0.3647	0.2829	1.8801
2	35	ค	ผิด	0.0035	0.4117	1.5586
3	164	ข	ถูก	0.3912	1.4421	0.8327
4	79	ข	ถูก	0.6286	2.1631	0.6799
5	165	ข	ถูก	1.0231	4.0278	0.4983
6	173	ก	ผิด	0.6030	5.2281	0.4373
7	157	ข	ผิด	0.5267	5.3097	0.4340
8	187	ข	ผิด	0.4913	5.3390	0.4328
9	212	ค	ถูก	0.5388	5.7929	0.4155
10	76	ค	ผิด	0.1722	6.7202	0.3858
11	58	ก	ผิด	0.0916	7.3723	0.3683
12	194	ข	ถูก	0.1956	7.9406	0.3549
13	83	ข	ผิด	0.1163	8.5152	0.3427
14	136	ง	ถูก	0.2111	9.1871	0.3299
15	86	ค	ผิด	0.1630	9.5829	0.3230
16	133	ง	ถูก	0.2466	10.2989	0.3116
17	57	ค	ผิด	0.1698	11.2049	0.2987
18	218	ข	ผิด	0.1000	11.8681	0.2903

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: ด.ช. ปวพล ดันรัมย์
 ชั้น/ห้อง: 6/3 โรงเรียน: วัดหงส์ปทุมมาวาส
 วันที่: 26/1/2004 เวลา: 12:48:13

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 222
 เวลาที่ใช้ในการทดสอบ: 0:10:52
 จำนวนข้อที่ตอบถูก: 9 จาก 18
 ตำแหน่งเปอร์เซ็นต์: 61
 ความสามารถในการสอบ: 0.3041 (ปานกลาง)
 สารสนเทศของแบบทดสอบ: 11.5945
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2937

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53	ก	ถูก	0.3647	0.2829	1.8801
2	35	ง	ถูก	0.5623	0.4126	1.5569
3	164	ข	ถูก	0.8495	0.9322	1.0357
4	79	ข	ถูก	1.0118	1.2735	0.8861
5	165	ข	ถูก	1.3135	2.3722	0.6493
6	58	ก	ผิด	0.6845	3.8868	0.5072
7	157	ข	ผิด	0.5985	4.0192	0.4988
8	94	ค	ผิด	0.5208	4.3191	0.4812
9	211	ค	ผิด	0.3208	5.2953	0.4346
10	57	ข	ผิด	0.2006	6.2307	0.4006
11	173	ค	ถูก	0.3248	6.7197	0.3858
12	131	ง	ผิด	0.2753	6.9073	0.3805
13	218	ง	ผิด	0.1662	7.6651	0.3612
14	83	ก	ถูก	0.2484	8.4542	0.3439
15	194	ข	ถูก	0.3485	9.3074	0.3278
16	86	ค	ผิด	0.2882	9.8775	0.3182
17	136	ง	ถูก	0.3747	10.8053	0.3042
18	133	ข	ผิด	0.3041	11.5945	0.2937

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: ด.ญ. อรวรรณ วงสงค์
 ชั้นห้อง: ป 6/3 โรงเรียน: วัดหงส์ปทุมมาวาส
 วันที่: 26/1/2004 เวลา: 13:03:19

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 222
 เวลาที่ใช้ในการทดสอบ: 0:17:00
 จำนวนข้อที่ตอบถูก: 9 จาก 22
 ตำแหน่งเปอร์เซ็นต์ไทล์: 44
 ความสามารถในการสอบ: -0.1393 (ปานกลาง)
 สารสนเทศของแบบทดสอบ: 11.6585
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2929

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53	ก	ถูก	0.3647	0.2829	1.8801
2	35	ง	ถูก	0.5623	0.4126	1.5569
3	164	ข	ถูก	0.8495	0.9322	1.0357
4	79	ค	ผิด	0.0047	1.9688	0.7127
5	165	ข	ถูก	0.3627	3.1014	0.5678
6	173	ค	ถูก	0.6860	4.5368	0.4695
7	157	ข	ผิด	0.5460	4.6298	0.4647
8	94	ค	ผิด	0.4433	4.8511	0.4540
9	212	ง	ผิด	-0.0006	6.0459	0.4067
10	76	ก	ถูก	0.0807	7.0871	0.3756
11	131	ง	ผิด	0.0489	7.1464	0.3741
12	132	ง	ผิด	0.0255	7.2121	0.3724
13	125	ข	ผิด	-0.0788	7.9925	0.3537
14	211	ง	ผิด	-0.1565	8.6214	0.3406
15	58	ก	ผิด	-0.1932	8.8634	0.3359
16	194	ก	ผิด	-0.2173	8.9815	0.3337
17	57	ก	ถูก	-0.1276	9.5473	0.3236
18	209	ง	ผิด	-0.1818	10.0815	0.3149
19	83	ค	ผิด	-0.2124	10.2699	0.3120
20	136	ง	ถูก	-0.1479	10.4990	0.3086
21	156	ง	ถูก	-0.0936	11.1593	0.2994
22	139	ค	ผิด	-0.1393	11.6585	0.2929

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: ด.ญ. ชุติพันธ์ เป่าเลี้ยง
 ชั้น/ห้อง: ป. 6/2 โรงเรียน: วัดนาบุญ
 วันที่: 15/1/04 เวลา: 12:22:41

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 222
 เวลาที่ใช้ในการทดสอบ: 0:07:17
 จำนวนข้อที่ตอบถูก: 11 จาก 18
 ตำแหน่งเปอร์เซ็นต์: 73
 ความสามารถในการสอบ: 0.6134 (ปานกลาง)
 สารสนเทศของแบบทดสอบ: 12.1041
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2874

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53	ก	ถูก	0.3647	0.2829	1.8801
2	35	ง	ถูก	0.5623	0.4126	1.5569
3	164	ข	ถูก	0.8495	0.9322	1.0357
4	79	ข	ถูก	1.0118	1.2735	0.8861
5	165	ง	ผิด	0.4251	2.6091	0.6191
6	58	ก	ผิด	0.1858	3.4475	0.5386
7	173	ค	ถูก	0.3932	4.0819	0.4950
8	157	ข	ผิด	0.3492	4.1047	0.4936
9	57	ค	ผิด	0.1691	5.0100	0.4468
10	211	ข	ถูก	0.2860	5.9816	0.4089
11	131	ง	ผิด	0.2288	6.1315	0.4038
12	194	ข	ถูก	0.3631	7.0134	0.3776
13	83	ก	ถูก	0.4689	8.1183	0.3510
14	136	ง	ถูก	0.5828	9.2990	0.3279
15	94	ก	ผิด	0.5325	9.6123	0.3225
16	132	ข	ผิด	0.4790	10.0080	0.3161
17	218	ก	ถูก	0.5413	11.0731	0.3005
18	167	ก	ถูก	0.6134	12.1041	0.2874

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: รมณียา เจริญศักดิ์
 ชั้นห้อง: 6/4 โรงเรียน: วัดหงส์ปทุมมาวาส
 วันที่: 26/1/2004 เวลา: 11:20:04

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 231
 เวลาที่ใช้ในการทดสอบ: 0:04:25
 จำนวนข้อที่ตอบถูก: 10 จาก 14
 ตำแหน่งเปอร์เซ็นต์: 82
 ความสามารถในการสอบ: 0.9518 (ปานกลาง)
 สারণเทศของแบบทดสอบ: 11.7502
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2917

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53	ก	ถูก	0.3647	0.2829	1.8801
2	35	ง	ถูก	0.5623	0.4126	1.5569
3	164	ข	ถูก	0.8495	0.9322	1.0357
4	79	ข	ถูก	1.0118	1.2735	0.8861
5	165	ข	ถูก	1.3135	2.3722	0.6493
6	58	ง	ถูก	1.4406	2.8769	0.5896
7	173	ง	ผิด	0.9125	4.7428	0.4592
8	194	ข	ถูก	1.0332	6.0091	0.4079
9	157	ข	ผิด	0.9070	6.8068	0.3833
10	187	ข	ผิด	0.8536	7.0588	0.3764
11	83	ก	ถูก	0.9145	8.1170	0.3510
12	136	ง	ถูก	0.9790	9.2370	0.3290
13	131	ง	ผิด	0.8637	10.5714	0.3076
14	132	ค	ถูก	0.9518	11.7502	0.2917

แบบทดสอบ: การทดสอบเรื่องเศษส่วน
 ชื่อผู้ทดสอบ: ด.ช.ชลทิตย์ วิจิตรประชา
 ชั้นห้อง: ป 6/1 โรงเรียน: วัดบางเตือ
 วันที่: 26/12/03 เวลา: 13:22:07
 ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 311
 เวลาที่ใช้ในการทดสอบ: 0:31:53
 จำนวนข้อที่ตอบถูก: 10 จาก 27
 ตำแหน่งเปอร์เซ็นต์: 40
 ความสามารถในการสอบ: -0.2413 (ปานกลาง)
 สারণนเทศของแบบทดสอบ: 11.3859
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2964

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53	ข	ผิด	-0.4623	0.2480	2.0080
2	35	ง	ถูก	-0.2407	0.3706	1.6426
3	122	ค	ถูก	0.0744	0.8491	1.0852
4	164	ข	ถูก	0.4261	1.8455	0.7361
5	79	ก	ผิด	-0.1794	2.6883	0.6099
6	58	ก	ผิด	-0.3051	2.8396	0.5934
7	165	ง	ผิด	-0.3609	2.8815	0.5891
8	173	ง	ผิด	-0.3894	2.8932	0.5879
9	194	ข	ถูก	-0.2477	2.9961	0.5777
10	212	ค	ถูก	-0.0759	4.2814	0.4833
11	76	ก	ถูก	0.0446	5.3632	0.4318
12	125	ง	ผิด	-0.1184	6.1152	0.4044
13	131	ข	ผิด	-0.1424	6.1355	0.4037
14	157	ง	ผิด	-0.1455	6.1360	0.4037
15	132	ข	ผิด	-0.1637	6.1630	0.4028
16	126	ค	ถูก	-0.0805	7.0522	0.3766
17	211	ก	ผิด	-0.1753	7.6601	0.3613
18	109	ก	ถูก	-0.0969	8.5012	0.3430
19	86	ค	ผิด	-0.1282	8.6391	0.3402
20	136	ง	ถูก	-0.0342	8.9776	0.3337
21	83	ข	ผิด	-0.0900	9.2771	0.3283
22	57	ค	ผิด	-0.1466	9.8209	0.3191
23	209	ง	ผิด	-0.2033	10.3338	0.3111
24	218	ก	ถูก	-0.1457	10.6721	0.3061
25	167	ค	ผิด	-0.1862	11.0466	0.3009
26	133	ข	ผิด	-0.2154	11.2619	0.2980
27	214	ค	ผิด	-0.2413	11.3859	0.2964

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: ต.ญ.ลลิตา พุกธรีวา
 ชั้นห้อง: ป.6/1 โรงเรียน: วัดนาบุญ
 วันที่: 15/01/2004 เวลา: 10:03:57

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 322
 เวลาที่ใช้ในการทดสอบ: 0:18:01
 จำนวนข้อที่ตอบถูก: 13 จาก 24
 ตำแหน่งเปอร์เซ็นต์: 64
 ความสามารถในการสอบ: 0.3676 (ปานกลาง)
 สারণเทศของแบบทดสอบ: 12.9120
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2783

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53	ง	ผิด	-0.4623	0.2480	2.0080
2	35	ง	ถูก	-0.2407	0.3706	1.6426
3	122	ค	ถูก	0.0744	0.8491	1.0852
4	164	ข	ถูก	0.4261	1.8455	0.7361
5	79	ค	ผิด	-0.1794	2.6883	0.6099
6	58	ก	ผิด	-0.3051	2.8396	0.5934
7	165	ง	ผิด	-0.3609	2.8815	0.5891
8	173	ก	ผิด	-0.3894	2.8932	0.5879
9	194	ก	ผิด	-0.4325	2.9358	0.5836
10	212	ค	ถูก	-0.2726	4.3366	0.4802
11	76	ก	ถูก	-0.1549	5.5575	0.4242
12	125	ง	ผิด	-0.2677	6.1845	0.4021
13	157	ข	ผิด	-0.2690	6.1847	0.4021
14	131	ค	ผิด	-0.2815	6.1938	0.4018
15	132	ก	ผิด	-0.2923	6.2081	0.4013
16	126	ค	ถูก	-0.2080	7.1440	0.3741
17	109	ก	ถูก	-0.1255	7.9916	0.3537
18	211	ข	ถูก	-0.0304	8.7561	0.3379
19	86	ง	ถูก	0.0709	9.0481	0.3324
20	136	ง	ถูก	0.1854	9.6802	0.3214
21	70	ค	ถูก	0.3055	10.1957	0.3132
22	83	ข	ผิด	0.2007	10.9061	0.3028
23	57	ก	ถูก	0.2981	11.9192	0.2897
24	218	ก	ถูก	0.3676	12.9120	0.2783

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: วัลลภ ศรีเมือง
 ชั้น/ห้อง: ป.6/1 โรงเรียน: วัดมูลจินดาราม
 วันที่: 26/4/00 เวลา: 14:35:33

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 322
 เวลาที่ใช้ในการทดสอบ: 0:13:01
 จำนวนข้อที่ตอบถูก: 10 จาก 27
 ตำแหน่งเปอร์เซ็นต์: 40
 ความสามารถในการสอบ: -0.2362 (ปานกลาง)
 สารสนเทศของแบบทดสอบ: 11.8590
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2904

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53	ก	ถูก	0.3647	0.2829	1.8801
2	35	ง	ถูก	0.5623	0.4126	1.5569
3	122	ข	ผิด	-0.0930	0.8614	1.0774
4	164	ข	ถูก	0.2308	2.0036	0.7065
5	79	ข	ถูก	0.4447	2.9290	0.5843
6	58	ก	ผิด	0.1253	3.6459	0.5237
7	165	ง	ผิด	-0.0025	3.9272	0.5046
8	173	ก	ผิด	-0.0609	4.0046	0.4997
9	131	ง	ผิด	-0.0938	4.0314	0.4980
10	212	ค	ถูก	0.0121	5.2094	0.4381
11	76	ค	ผิด	-0.2158	6.4344	0.3942
12	211	ข	ถูก	-0.1141	7.1100	0.3750
13	157	ข	ผิด	-0.1161	7.1107	0.3750
14	132	ก	ผิด	-0.1303	7.1424	0.3742
15	194	ก	ผิด	-0.1656	7.2907	0.3704
16	125	ข	ผิด	-0.2408	7.9417	0.3548
17	126	ค	ถูก	-0.1797	8.8708	0.3358
18	57	ค	ผิด	-0.2255	9.3266	0.3274
19	136	ข	ผิด	-0.2508	9.4822	0.3247
20	83	ก	ถูก	-0.1945	9.6844	0.3213
21	86	ข	ผิด	-0.2135	9.7812	0.3197
22	209	ข	ผิด	-0.2600	10.2390	0.3125
23	156	ค	ผิด	-0.3112	10.6995	0.3057
24	139	ง	ถูก	-0.2470	11.0973	0.3002
25	218	ข	ผิด	-0.2794	11.3085	0.2974
26	167	ก	ถูก	-0.2129	11.6582	0.2929
27	133	ก	ผิด	-0.2362	11.8590	0.2904

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: ด.ญ.อรุมา อาน้อย
 ชั้น/ห้อง: ป.6/3 โรงเรียน: วัดหงส์ปทุมมาวาส
 วันที่: 26/1/2004 เวลา: 14:53:46

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 322
 เวลาที่ใช้ในการทดสอบ: 0:10:29
 จำนวนข้อที่ตอบถูก: 11 จาก 18
 ตำแหน่งเปอร์เซ็นต์: 68
 ความสามารถในการสอบ: 0.4761 (ปานกลาง)
 สารสนเทศของแบบทดสอบ: 11.2451
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2982

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	Zeta	SEE
1	53	ก	ถูก	0.3647	0.2829	1.8801
2	35	ง	ถูก	0.5623	0.4126	1.5569
3	122	ค	ถูก	0.7977	0.7710	1.1388
4	164	ข	ถูก	1.0036	1.1501	0.9325
5	79	ข	ถูก	1.1247	1.4139	0.8410
6	58	ก	ผิด	0.4912	2.8151	0.5960
7	165	ง	ผิด	0.2732	3.6742	0.5217
8	173	ข	ผิด	0.1827	3.9391	0.5038
9	194	ข	ถูก	0.3385	4.7729	0.4577
10	136	ง	ถูก	0.4849	5.8526	0.4134
11	83	ข	ผิด	0.3252	6.7651	0.3845
12	86	ก	ผิด	0.2495	7.2777	0.3707
13	131	ค	ผิด	0.2118	7.4156	0.3672
14	70	ก	ผิด	0.1651	7.7626	0.3589
15	110	ก	ถูก	0.2699	8.1748	0.3498
16	57	ก	ถูก	0.3571	9.2215	0.3293
17	211	ข	ถูก	0.4182	10.1899	0.3133
18	218	ก	ถูก	0.4761	11.2451	0.2982

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: ด.ญ.วรางค์รัตน์ เมฆอรุณ
 ชั้นห้อง: 6/3 โรงเรียน: วัดหงส์ปทุมมาวาส
 วันที่: 26/1/2004 เวลา: 15:08:19

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 322
 เวลาที่ใช้ในการทดสอบ: 0:23:30
 จำนวนข้อที่ตอบถูก: 13 จาก 24
 ตำแหน่งเปอร์เซ็นต์: 64
 ความสามารถในการสอบ: 0.3690 (ปานกลาง)
 สารสนเทศของแบบทดสอบ: 13.5822
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2713

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53	ก	ถูก	0.3647	0.2829	1.8801
2	35	ง	ถูก	0.5623	0.4126	1.5569
3	122	ข	ผิด	-0.0930	0.8614	1.0774
4	164	ข	ถูก	0.2308	2.0036	0.7065
5	79	ง	ผิด	-0.2328	2.7740	0.6004
6	58	ก	ผิด	-0.3364	2.9057	0.5866
7	165	ง	ผิด	-0.3833	2.9426	0.5830
8	173	ก	ผิด	-0.4077	2.9530	0.5819
9	194	ก	ผิด	-0.4465	2.9928	0.5780
10	212	ค	ถูก	-0.2910	4.3929	0.4771
11	76	ก	ถูก	-0.1757	5.6176	0.4219
12	125	ค	ถูก	-0.0525	6.4154	0.3948
13	157	ข	ผิด	-0.0566	6.4164	0.3948
14	131	ค	ผิด	-0.0803	6.4454	0.3939
15	94	ค	ผิด	-0.0921	6.4632	0.3933
16	126	ค	ถูก	-0.0213	7.3182	0.3697
17	211	ข	ถูก	0.0760	8.1806	0.3496
18	109	ก	ถูก	0.1375	8.9194	0.3348
19	86	ง	ถูก	0.2463	9.4273	0.3257
20	136	ง	ถูก	0.3551	10.3254	0.3112
21	70	ก	ผิด	0.2925	10.8238	0.3040
22	83	ก	ถูก	0.3756	11.8114	0.2910
23	133	ง	ถูก	0.4619	12.7775	0.2798
24	214	ง	ผิด	0.3690	13.5822	0.2713

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: ด.ญ.อภิัญญา หุ่นดี
 ชั้นห้อง: ป.6/3 โรงเรียน: วัดหงส์ปทุมาวาส
 วันที่: 26/1/2004 เวลา: 15:42:11

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 322
 เวลาที่ใช้ในการทดสอบ: 0:03:53
 จำนวนข้อที่ตอบถูก: 10 จาก 15
 ตำแหน่งเปอร์เซ็นต์ไทล์: 83
 ความสามารถในการสอบ: 0.9893 (ปานกลาง)
 สารสนเทศของแบบทดสอบ: 12.1994
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2863

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53	ก	ถูก	0.3647	0.2829	1.8801
2	35	ง	ถูก	0.5623	0.4126	1.5569
3	122	ก	ผิด	-0.0930	0.8614	1.0774
4	164	ข	ถูก	0.2308	2.0036	0.7065
5	79	ข	ถูก	0.4447	2.9290	0.5843
6	58	ง	ถูก	0.7367	4.4276	0.4752
7	157	ข	ผิด	0.5900	4.5529	0.4687
8	173	ข	ผิด	0.4042	5.2128	0.4380
9	131	ก	ถูก	0.5683	5.8401	0.4138
10	165	ข	ถูก	0.7821	7.9729	0.3542
11	83	ก	ถูก	0.8777	9.0698	0.3320
12	194	ข	ถูก	0.9832	10.3954	0.3102
13	187	ก	ผิด	0.9144	10.7399	0.3051
14	41	ก	ผิด	0.8614	11.0587	0.3007
15	94	ง	ถูก	0.9893	12.1994	0.2863

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: ด.ช.สิทธิโชค โตตะคุช
 ชั้นห้อง: ป.6/1 โรงเรียน: วัดนาบุญ
 วันที่: 15/1/04 เวลา: 10:04:16

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 332
 เวลาที่ใช้ในการทดสอบ: 0:06:14
 จำนวนข้อที่ตอบถูก: 12 จาก 15
 ตำแหน่งเปอร์เซ็นต์ไทล์: 90
 ความสามารถในการสอบ: 1.3059 (สูง)
 สารสนเทศของแบบทดสอบ: 12.0239
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2884

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53	ก	ถูก	0.3647	0.2829	1.8801
2	35	ค	ผิด	0.0035	0.4117	1.5586
3	122	ค	ถูก	0.2954	0.8913	1.0592
4	164	ข	ถูก	0.5972	1.6950	0.7681
5	58	ง	ถูก	0.9271	2.9973	0.5776
6	79	ข	ถูก	1.0274	3.3268	0.5483
7	165	ข	ถูก	1.2545	4.5803	0.4673
8	173	ค	ถูก	1.4118	5.7472	0.4171
9	194	ข	ถูก	1.4892	6.3309	0.3974
10	131	ก	ถูก	1.5846	7.3097	0.3699
11	86	ง	ถูก	1.6389	7.8374	0.3572
12	132	ง	ผิด	1.3646	9.1939	0.3298
13	110	ก	ถูก	1.4333	10.1515	0.3139
14	70	ค	ถูก	1.4826	10.9238	0.3026
15	219	ง	ผิด	1.3059	12.0239	0.2884

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: ด.ญ. ศรียรรณ ดีแจ้ง
 ชั้น/ห้อง: ป.6/4 โรงเรียน โรงเรียนวัดหงส์
 วันที่: 26/1/2004 เวลา: 13:43:28

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 332
 เวลาที่ใช้ในการทดสอบ: 0:12:11
 จำนวนข้อที่ตอบถูก: 11 จาก 24
 ตำแหน่งเปอร์เซ็นต์ไทล์: 50
 ความสามารถในการสอบ: 0.0189 (ปานกลาง)
 สารสนเทศของแบบทดสอบ: 12.8490
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2790

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53	ข	ผิด	-0.4623	0.2480	2.0080
2	35	ง	ถูก	-0.2407	0.3706	1.6426
3	122	ข	ผิด	-0.6303	0.6216	1.2683
4	164	ข	ถูก	-0.2565	1.4599	0.8276
5	79	ข	ถูก	0.0491	2.5238	0.6295
6	211	ข	ถูก	0.3563	3.5007	0.5345
7	165	ง	ผิด	0.0746	3.8994	0.5064
8	173	ก	ผิด	-0.0269	3.9923	0.5005
9	131	ข	ผิด	-0.0786	4.0216	0.4987
10	212	ค	ถูก	0.0511	5.1453	0.4409
11	76	ก	ถูก	0.1453	6.1073	0.4046
12	57	ค	ผิด	0.0198	6.8482	0.3821
13	194	ข	ถูก	0.1563	7.3514	0.3688
14	86	ค	ผิด	0.0859	7.6587	0.3613
15	58	ก	ผิด	0.0055	8.1592	0.3501
16	83	ข	ผิด	-0.0594	8.4927	0.3431
17	218	ก	ถูก	0.0163	9.0364	0.3327
18	136	ง	ถูก	0.1207	9.5724	0.3232
19	167	ข	ผิด	0.0470	10.1962	0.3132
20	133	ก	ผิด	-0.0057	10.6014	0.3071
21	214	ง	ผิด	-0.0519	10.8597	0.3035
22	139	ง	ถูก	0.0240	11.5192	0.2946
23	209	ข	ผิด	-0.0406	12.1918	0.2864
24	124	ค	ถูก	0.0189	12.8490	0.2790

แบบทดสอบ: การทดสอบเรื่อง เศษส่วน
 ชื่อผู้ทดสอบ: พิมพ์พิสุทธิ์ ปัจจัย
 ชั้นห้อง: ป 6/3 โรงเรียน: วัดหงส์ปทุมมาวาส
 วันที่: 26/1/2004 เวลา: 15:46:50

ผลการทดสอบ:

ชุดแบบทดสอบที่ทำ: 332
 เวลาที่ใช้ในการทดสอบ: 0:07:44
 จำนวนข้อที่ตอบถูก: 12 จาก 21
 ตำแหน่งเปอร์เซ็นต์ไทล์: 67
 ความสามารถในการสอบ: 0.4566 (ปานกลาง)
 สारสนเทศของแบบทดสอบ: 12.9448
 ความคลาดเคลื่อนมาตรฐานในการวัด: 0.2779

ข้อที่	ข้อในข้อสอบ	เลือกตัวเลือก	ผล	Zeta	I Zeta	SEE
1	53	ก	ถูก	0.3647	0.2829	1.8801
2	35	ค	ผิด	0.0035	0.4117	1.5586
3	122	ค	ถูก	0.2954	0.8913	1.0592
4	164	ข	ถูก	0.5972	1.6950	0.7681
5	58	ก	ผิด	0.1397	2.4402	0.6402
6	79	ข	ถูก	0.3136	3.4774	0.5363
7	165	ง	ผิด	0.1314	3.9843	0.5010
8	173	ง	ผิด	0.0541	4.1260	0.4923
9	194	ข	ถูก	0.2221	4.7406	0.4593
10	57	ก	ถูก	0.3823	5.7982	0.4153
11	83	ก	ถูก	0.5115	6.9441	0.3795
12	218	ค	ผิด	0.3133	7.8860	0.3561
13	131	ง	ผิด	0.2621	8.0622	0.3522
14	157	ง	ผิด	0.2518	8.0732	0.3519
15	132	ก	ผิด	0.2196	8.2252	0.3487
16	212	ค	ถูก	0.2642	9.0293	0.3328
17	76	ก	ถูก	0.3026	9.7836	0.3197
18	211	ก	ผิด	0.2002	10.7266	0.3053
19	86	ง	ถูก	0.2864	11.2940	0.2976
20	136	ง	ถูก	0.3699	12.2146	0.2861
21	70	ค	ถูก	0.4566	12.9448	0.2779

ค่าพารามิเตอร์ของข้อสอบในคลัง

ค่าพารามิเตอร์ของข้อสอบในคลัง

=====

ชื่อแบบทดสอบ : การทดสอบความเข้าใจเรื่องเศษส่วน
 ชนิดตัวเลือก : ก, ข, ค, ง
 จำนวนตัวเลือก : 4 ตัวเลือก
 จำนวนข้อสูงสุดในการทำ : 36 ข้อ
 เวลาเฉลี่ยในการทำแต่ละข้อ : 90 วินาที
 จำนวนข้อสอบ : 244 ข้อ

Item	A	B	C	Answer	Time	Image File
1	1.0890	0.5350	0.1550	ง	95	Item001.bmp
2	1.0640	0.6670	0.1520	ค	95	Item002.bmp
3	0.8560	2.2140	0.2680	ข	107	Item003.bmp
4	0.9790	2.2110	0.1320	ค	107	Item004.bmp
5	1.1980	1.3500	0.0920	ข	99	Item005.bmp
6	0.7070	1.1370	0.2400	ก	99	Item006.bmp
7	0.9610	0.2410	0.2490	ง	90	Item007.bmp
8	0.8100	0.1910	0.2390	ค	90	Item008.bmp
9	1.2980	1.4610	0.2240	ง	99	Item009.bmp
10	1.5130	2.0550	0.1920	ข	107	Item010.bmp
11	0.7530	-0.0750	0.2900	ข	84	Item011.bmp
12	0.9440	2.0020	0.1870	ค	103	Item012.bmp
13	0.9320	0.6420	0.2180	ก	95	Item013.bmp
14	1.0820	-0.0300	0.1970	ข	84	Item014.bmp
15	1.1850	-0.3000	0.2100	ก	84	Item015.bmp
16	1.6130	2.7200	0.1800	ง	112	Item016.bmp
17	0.9460	1.2230	0.1030	ก	99	Item017.bmp
18	0.9660	0.7270	0.2950	ค	95	Item018.bmp
19	0.7070	-0.5290	0.1800	ง	77	Item019.bmp
20	1.0250	1.6670	0.2260	ข	103	Item020.bmp
21	1.4390	1.5620	0.2070	ง	103	Item021.bmp
22	1.1100	1.0440	0.2570	ข	99	Item022.bmp
23	1.4730	0.8130	0.2840	ค	95	Item023.bmp
24	0.9380	0.9570	0.2010	ก	95	Item024.bmp
25	1.3640	0.8910	0.1490	ก	95	Item025.bmp
26	0.4860	-0.9310	0.2440	ง	77	Item026.bmp
27	0.9290	0.2540	0.2190	ข	90	Item027.bmp
28	0.4060	1.6990	0.1930	ค	103	Item028.bmp
29	1.3130	0.9870	0.1530	ง	95	Item029.bmp
30	1.1970	1.3550	0.1870	ก	99	Item030.bmp
31	1.1730	2.6330	0.2150	ข	112	Item031.bmp
32	0.6080	-1.5700	0.2130	ค	65	Item032.bmp
33	0.9180	2.1180	0.2300	ง	107	Item033.bmp
34	0.9380	1.2480	0.2020	ง	99	Item034.bmp
35	0.5110	0.0130	0.1880	ง	90	Item035.bmp
36	1.2030	1.2320	0.2050	ค	99	Item036.bmp
37	1.7100	1.7060	0.1760	ข	103	Item037.bmp
38	0.5980	-1.0800	0.1660	ก	72	Item038.bmp
39	0.4550	-1.2500	0.2120	ง	72	Item039.bmp
40	1.2970	0.8020	0.1360	ง	95	Item040.bmp

Item	A	B	C	Answer	Time	Image File
41	1.7440	1.5720	0.1530	ข	103	Item041.bmp
42	0.6220	-0.7360	0.2610	ง	77	Item042.bmp
43	0.9040	0.4550	0.1100	ง	90	Item043.bmp
44	1.5000	1.6860	0.1540	ก	103	Item044.bmp
45	1.0420	1.6260	0.1640	ง	103	Item045.bmp
46	0.8630	-0.0310	0.1920	ค	84	Item046.bmp
47	1.1700	1.2810	0.2580	ก	99	Item047.bmp
48	1.2990	1.8390	0.2780	ง	103	Item048.bmp
49	0.6980	-0.5210	0.1410	ค	77	Item049.bmp
50	0.5820	-0.0490	0.1780	ง	84	Item050.bmp
51	1.2070	0.2540	0.2480	ข	90	Item051.bmp
52	1.6310	1.2400	0.1980	ข	99	Item052.bmp
53	0.7060	-0.0100	0.1140	ก	84	Item053.bmp
54	1.1210	-0.3990	0.1450	ข	84	Item054.bmp
55	1.0620	-0.7470	0.1130	ก	77	Item055.bmp
56	0.6350	-0.3670	0.1460	ข	84	Item056.bmp
57	1.3840	0.4120	0.1290	ก	90	Item057.bmp
58	1.6840	0.5900	0.1570	ง	95	Item058.bmp
59	1.1800	1.2530	0.1330	ค	99	Item059.bmp
60	1.0880	0.2950	0.2040	ค	90	Item060.bmp
61	0.8040	0.8990	0.2310	ค	95	Item061.bmp
62	0.8610	-0.0640	0.0880	ง	84	Item062.bmp
63	1.0050	0.4050	0.2220	ก	90	Item063.bmp
64	1.1330	2.2340	0.2140	ข	107	Item064.bmp
65	0.6150	-1.5200	0.2470	ก	65	Item065.bmp
66	1.1690	1.6790	0.2130	ค	103	Item066.bmp
67	1.6510	1.4810	0.2670	ข	99	Item067.bmp
68	0.7660	-0.3270	0.1600	ค	84	Item068.bmp
69	1.2910	2.2950	0.2350	ง	107	Item069.bmp
70	1.4990	0.8750	0.1510	ค	95	Item070.bmp
71	0.4950	-2.2300	0.2150	ค	60	Item071.bmp
72	0.5870	-0.1960	0.2800	ง	84	Item072.bmp
73	1.1080	0.4010	0.1470	ง	90	Item073.bmp
74	0.7980	0.3520	0.1890	ข	90	Item074.bmp
75	1.0900	-0.2050	0.2040	ข	84	Item075.bmp
76	1.6100	-0.3020	0.2210	ก	84	Item076.bmp
77	1.3010	-0.3540	0.2340	ค	84	Item077.bmp
78	1.1970	-0.1940	0.2480	ก	84	Item078.bmp
79	1.5580	0.0400	0.2470	ข	90	Item079.bmp
80	0.9840	1.7800	0.2870	ค	103	Item080.bmp
81	0.8750	-0.9180	0.1390	ก	77	Item081.bmp
82	0.8820	-0.2660	0.1720	ข	84	Item082.bmp
83	1.6230	0.5560	0.2350	ก	95	Item083.bmp
84	0.4390	-0.5950	0.2340	ง	77	Item084.bmp
85	1.0460	0.7460	0.2790	ข	95	Item085.bmp
86	1.5690	0.8150	0.1610	ง	95	Item086.bmp
87	0.8250	-0.4740	0.1740	ค	84	Item087.bmp
88	1.0400	0.3150	0.2240	ก	90	Item088.bmp

Item	A	B	C	Answer	Time	Image File
89	1.3750	1.1790	0.2960	Ⓝ	99	Item089.bmp
90	0.8320	2.2920	0.2820	Ⓝ	107	Item090.bmp
91	0.8710	0.6420	0.1870	Ⓝ	95	Item091.bmp
92	1.1060	1.0990	0.2280	Ⓝ	99	Item092.bmp
93	0.8730	1.1230	0.1290	Ⓝ	99	Item093.bmp
94	1.6750	1.2930	0.1250	Ⓝ	99	Item094.bmp
95	1.5550	0.7410	0.2800	Ⓝ	95	Item095.bmp
96	0.6110	-1.0600	0.2330	Ⓝ	72	Item096.bmp
97	1.3640	2.5410	0.1480	Ⓝ	112	Item097.bmp
98	0.8130	0.6660	0.2050	Ⓝ	95	Item098.bmp
99	1.0720	0.1540	0.2450	Ⓝ	90	Item099.bmp
100	0.4620	-0.5140	0.1560	Ⓝ	77	Item100.bmp
101	0.8230	-0.5070	0.1800	Ⓝ	77	Item101.bmp
102	0.8670	2.3940	0.2170	Ⓝ	107	Item102.bmp
103	0.6300	-0.5710	0.2330	Ⓝ	77	Item103.bmp
104	0.6950	-0.7570	0.1770	Ⓝ	77	Item104.bmp
105	1.1340	0.5270	0.2400	Ⓝ	95	Item105.bmp
106	0.6990	-0.1330	0.1670	Ⓝ	84	Item106.bmp
107	1.5860	2.1150	0.2010	Ⓝ	107	Item107.bmp
108	0.7230	-0.8730	0.1350	Ⓝ	77	Item108.bmp
109	1.2060	-0.3060	0.1050	Ⓝ	84	Item109.bmp
110	1.4730	0.9650	0.1220	Ⓝ	95	Item110.bmp
111	0.9550	1.2360	0.1760	Ⓝ	99	Item111.bmp
112	1.2660	0.6460	0.2700	Ⓝ	95	Item112.bmp
113	0.8380	0.5780	0.1310	Ⓝ	95	Item113.bmp
114	0.5110	-1.3100	0.1960	Ⓝ	72	Item114.bmp
115	1.0380	1.3310	0.1460	Ⓝ	99	Item115.bmp
116	0.5910	-0.9860	0.1850	Ⓝ	77	Item116.bmp
117	1.5760	2.0880	0.2700	Ⓝ	107	Item117.bmp
118	1.1960	1.1030	0.2880	Ⓝ	99	Item118.bmp
119	1.5430	2.2530	0.0820	Ⓝ	107	Item119.bmp
120	0.9740	-1.0100	0.1740	Ⓝ	72	Item120.bmp
121	1.0560	1.3650	0.1270	Ⓝ	99	Item121.bmp
122	1.0190	0.0210	0.2280	Ⓝ	90	Item122.bmp
123	0.4700	-0.3100	0.1920	Ⓝ	84	Item123.bmp
124	1.3320	0.2890	0.2150	Ⓝ	90	Item124.bmp
125	1.2830	0.1190	0.1520	Ⓝ	90	Item125.bmp
126	1.3500	-0.3830	0.1740	Ⓝ	84	Item126.bmp
127	1.0010	0.7350	0.2220	Ⓝ	95	Item127.bmp
128	1.7010	2.4170	0.1560	Ⓝ	107	Item128.bmp
129	1.0000	1.2220	0.2220	Ⓝ	99	Item129.bmp
130	0.8340	0.6480	0.2240	Ⓝ	95	Item130.bmp
131	1.8350	1.0400	0.2020	Ⓝ	99	Item131.bmp
132	1.6000	1.1560	0.1420	Ⓝ	99	Item132.bmp
133	1.4490	0.6210	0.1710	Ⓝ	95	Item133.bmp
134	0.6170	-0.9510	0.1540	Ⓝ	77	Item134.bmp
135	1.2850	0.4710	0.1920	Ⓝ	90	Item135.bmp
136	1.5500	0.6530	0.1720	Ⓝ	95	Item136.bmp

Item	A	B	C	Answer	Time	Image File
137	0.6360	-0.9130	0.2420	↓	77	Item137.bmp
138	0.9430	0.0640	0.1880	↓	90	Item138.bmp
139	1.3190	0.3820	0.1630	↓	90	Item139.bmp
140	1.4810	1.7050	0.1980	∩	103	Item140.bmp
141	1.0130	1.6520	0.2130	∩	103	Item141.bmp
142	0.9510	1.6840	0.2410	∩	103	Item142.bmp
143	1.2570	1.6850	0.1510	∪	103	Item143.bmp
144	1.3730	1.1650	0.1500	∩	99	Item144.bmp
145	1.6370	1.9200	0.0610	∩	103	Item145.bmp
146	0.7880	-1.4200	0.1880	∩	72	Item146.bmp
147	0.7230	-1.6700	0.1860	∩	65	Item147.bmp
148	0.8800	0.5520	0.1760	∪	95	Item148.bmp
149	1.4990	1.4290	0.1410	∩	99	Item149.bmp
150	1.2550	1.0550	0.1380	∩	99	Item150.bmp
151	0.9380	0.0810	0.2320	∩	90	Item151.bmp
152	1.5010	1.3190	0.1570	∪	99	Item152.bmp
153	1.4900	1.0140	0.2110	∪	99	Item153.bmp
154	0.9980	-0.5510	0.2180	∩	77	Item154.bmp
155	1.0240	-0.1150	0.2230	↓	84	Item155.bmp
156	1.3650	0.1240	0.2560	↓	90	Item156.bmp
157	2.3380	1.4060	0.1180	∩	99	Item157.bmp
158	1.0760	2.6450	0.1920	↓	112	Item158.bmp
159	0.6620	-0.0600	0.2560	↓	84	Item159.bmp
160	0.8990	0.4630	0.1080	↓	90	Item160.bmp
161	1.2480	2.0150	0.1980	∩	107	Item161.bmp
162	1.2180	1.0520	0.2110	∩	99	Item162.bmp
163	0.7020	-0.5680	0.1860	∩	77	Item163.bmp
164	1.5140	0.0400	0.1850	∪	90	Item164.bmp
165	1.9290	0.7520	0.1190	∪	95	Item165.bmp
166	0.7680	-0.1090	0.2060	↓	84	Item166.bmp
167	1.3710	0.5070	0.1430	∩	90	Item167.bmp
168	1.0540	1.3230	0.1800	∩	99	Item168.bmp
169	1.0780	0.3920	0.2050	∩	90	Item169.bmp
170	1.1450	-0.1610	0.2940	∩	84	Item170.bmp
171	1.4030	1.3920	0.1070	↓	99	Item171.bmp
172	1.7040	1.6070	0.1680	∩	103	Item172.bmp
173	1.8840	0.9110	0.1580	∩	95	Item173.bmp
174	1.0260	-0.5380	0.1870	∪	77	Item174.bmp
175	1.1980	0.1400	0.2360	∩	90	Item175.bmp
176	1.2020	-0.7070	0.1430	∪	77	Item176.bmp
177	2.1500	2.1890	0.2050	∩	107	Item177.bmp
178	1.4200	1.4630	0.1110	∪	99	Item178.bmp
179	1.1330	0.9000	0.2890	∪	95	Item179.bmp
180	0.6460	-0.4650	0.2110	∩	84	Item180.bmp
181	1.5520	0.9760	0.2450	∩	95	Item181.bmp
182	0.9490	-0.5170	0.1850	∪	77	Item182.bmp
183	0.5750	-0.3780	0.2020	∩	84	Item183.bmp
184	0.5230	-0.2200	0.2220	↓	84	Item184.bmp

Item	A	B	C	Answer	Time	Image File
185	1.2660	1.5390	0.1990	↓	103	Item185.bmp
186	1.2830	1.1420	0.2290	↗	99	Item186.bmp
187	2.0700	1.5760	0.1450	↓	103	Item187.bmp
188	1.0450	0.7640	0.3000	↓	95	Item188.bmp
189	1.0640	1.3630	0.2670	↻	99	Item189.bmp
190	0.9850	0.1730	0.2090	↻	90	Item190.bmp
191	1.0350	0.3440	0.2040	↻	90	Item191.bmp
192	1.1690	0.1580	0.2390	↓	90	Item192.bmp
193	1.1660	1.0680	0.2120	↻	99	Item193.bmp
194	1.6680	0.7140	0.1770	↗	95	Item194.bmp
195	1.0290	-1.4300	0.2060	↗	72	Item195.bmp
196	0.8190	-0.1510	0.1870	↻	84	Item196.bmp
197	1.2080	1.0840	0.1170	↻	99	Item197.bmp
198	1.4940	2.1440	0.0940	↓	107	Item198.bmp
199	1.5810	1.6540	0.2560	↻	103	Item199.bmp
200	1.4550	1.5200	0.1120	↻	103	Item200.bmp
201	0.9030	-1.9320	0.2390	↻	65	Item201.BMP
202	0.6330	0.0350	0.2180	↻	90	Item202.bmp
203	1.2120	1.2110	0.1880	↻	99	Item203.bmp
204	0.8500	0.8620	0.3000	↗	95	Item204.bmp
205	0.9240	-0.2410	0.2470	↓	84	Item205.bmp
206	0.7590	0.2860	0.1860	↓	90	Item206.bmp
207	0.9640	0.3640	0.1940	↓	90	Item207.bmp
208	1.0560	-0.1960	0.1630	↻	84	Item208.bmp
209	1.3300	0.2850	0.1770	↻	90	Item209.bmp
210	0.9600	0.2280	0.2000	↓	90	Item210.bmp
211	1.3860	0.2380	0.1830	↗	90	Item211.bmp
212	1.6190	-0.3580	0.1570	↻	84	Item212.bmp
213	0.4510	-1.8010	0.2610	↻	65	Item213.bmp
214	1.5900	0.6020	0.2690	↗	95	Item214.bmp
215	1.0210	0.1710	0.2190	↗	90	Item215.bmp
216	1.2080	0.1550	0.3060	↻	90	Item216.bmp
217	0.7580	0.1650	0.1820	↻	90	Item217.bmp
218	1.5960	0.4140	0.2840	↻	90	Item218.bmp
219	1.4670	1.0050	0.1360	↻	95	Item219.bmp
220	1.0760	0.6730	0.1780	↗	95	Item220.bmp
221	0.9270	-0.0490	0.1110	↓	84	Item221.bmp
222	0.8980	-0.6360	0.2050	↗	77	Item222.bmp
223	0.7340	0.7740	0.1800	↻	95	Item223.bmp
224	0.9600	0.8750	0.2320	↓	95	Item224.bmp
225	1.1120	0.4860	0.2430	↻	90	Item225.bmp
226	0.8110	-0.3250	0.2490	↗	84	Item226.bmp
227	0.5180	0.0560	0.2960	↗	90	Item227.bmp
228	0.9640	0.3920	0.3010	↗	90	Item228.bmp
229	1.0760	-0.2700	0.2400	↗	84	Item229.bmp
230	1.2430	0.1120	0.2130	↓	90	Item230.bmp
231	0.1324	1.1630	0.3000	↻	99	Item231.bmp
232	0.8160	-0.4970	0.2290	↓	84	Item232.bmp

Item	A	B	C	Answer	Time	Image File
233	1.3680	-0.4470	0.2240	၅	84	Item233.bmp
234	1.4060	0.7500	0.1400	၂	95	Item234.bmp
235	1.0510	2.2600	0.2390	၅	107	Item235.bmp
236	1.3160	1.2890	0.1510	ဂ	99	Item236.bmp
237	1.2700	0.4050	0.1760	၅	90	Item237.bmp
238	1.0244	2.1980	0.2800	ဂ	107	Item238.bmp
239	0.9700	1.5150	0.2310	ဂ	103	Item239.bmp
240	1.3660	1.3620	0.1350	ဂ	99	Item240.bmp
241	1.0260	2.2880	0.2480	၂	107	Item241.bmp
242	1.3570	1.7060	0.2160	၂	103	Item242.bmp
243	1.1640	1.3890	0.2570	ဂ	99	Item243.bmp
244	0.9710	1.7110	0.2160	၅	103	Item244.bmp

ประวัติย่อผู้วิจัย

ประวัติย่อผู้วิจัย

ชื่อ-สกุล	นายเกียรติศักดิ์ ส่องแสง
วัน เดือน ปีเกิด	วันที่ 16 ธันวาคม พุทธศักราช 2508
สถานที่เกิด	จังหวัดอุบลราชธานี
สถานที่อยู่ปัจจุบัน	104/64 หมู่ที่ 5 แขวงฉิมพลี เขตตลิ่งชัน กรุงเทพมหานคร
ตำแหน่งหน้าที่ปัจจุบัน	อาจารย์คณะศึกษาศาสตร์ สถาบันเทคโนโลยีราชมงคล

ประวัติการศึกษา

พ.ศ.2532	กศ.บ. (การวัดผลการศึกษา) มหาวิทยาลัยศรีนครินทรวิโรฒ
พ.ศ.2538	กศ.ม. (การวัดผลการศึกษา) มหาวิทยาลัยศรีนครินทรวิโรฒ
พ.ศ.2547	กศ.ด. (การทดสอบและวัดผลการศึกษา) มหาวิทยาลัยศรีนครินทรวิโรฒ