

การพัฒนาวิธีการประเมินเชิงวิจจัยโดยประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะและ  
การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

นางอมรรัตน์ สร้อยสังวาลย์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรดุษฎีบัณฑิต  
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2551

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

THE DEVELOPMENT OF A DIAGNOSTIC ASSESSMENT METHOD BY APPLYING  
THE ATTRIBUTE HIERARCHY MODEL AND COMPUTERIZED ADAPTIVE TESTING

Mrs. Amornrat Soisangwarn

A Dissertation Submitted in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy Program in Educational Measurement and Evaluation

Department of Educational Research and Psychology

Faculty of Education

Chulalongkorn University

Academic year 2008

Copyright of Chulalongkorn University







# # 4784640227 : MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION

KEYWORDS: DIAGNOSTIC ASSESSMENT / THE ATTRIBUTE HIERARCHY MODEL /

COMPUTERIZED ADAPTIVE TESTING

AMORN RAT SOISANGWARN: THE DEVELOPMENT OF A DIAGNOSTIC ASSESSMENT METHOD BY APPLYING THE ATTRIBUTE HIERARCHY MODEL AND COMPUTERIZED ADAPTIVE TESTING THESIS ADVISOR: ASSOC. PROF. SIRIDEJ SUJIVA, PH.D., THESIS CO-ADVISOR : PROF. SUWIMON WONGWANICH, PH.D., 317 pp.

The purposes of this research were to develop a diagnostic method for detecting knowledge state and incorrect rules on addition and subtraction of fractions by applying the attribute hierarchy model and computerized adaptive testing, to verify quality of the developed method from validity, reliability, time spent and number of items in the testing, to compare quality of the developed method when criteria for judging knowledge state were different and to study attitude of participates in the difficulty of using, the feasibility to apply, utility and satisfaction. The design of the research was divided into 4 phases 1) developed test let and paper and pencil test 2) developed computerized adaptive testing program 3) diagnostic assessed and 4) verified quality of the developed method and studied attitude of the participates. The primary students, teacher and expert in mathematic curriculum, mathematic instruction and educational measurement were samples.

This research found that the development method has quality to applied cognitive theory and modern test theory for educational diagnostic assessment. It has quality both aspect of validity, reliability, time spent and number of items in the testing. Participants were satisfied with this development method and it was not difficult of use, feasible to apply and utilized. The study of quality when criteria for knowledge state were different found no significance at .05 levels in validity, reliability and number of items in the testing, but it was found the significance at .05 levels of different of time spent.

Department: Educational Research and Psychology Student's signature.....

Field of study: Educational Measurement and Evaluation Advisor's signature.....

Academic year: 2008 Co-advisor's .....



## กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความสามารถและความช่วยเหลืออย่างดียิ่งจากรองศาสตราจารย์ ดร.ศิริเดช สุชีวะและศาสตราจารย์ ดร.สุวิมล ว่องวาณิช อาจารย์ที่ปรึกษาผู้ซึ่งให้คำปรึกษา แนะนำและตรวจสอบแก้ไขข้อบกพร่องต่าง ๆ รวมทั้งเป็นกำลังใจให้กับผู้วิจัยด้วยดีเสมอมา ผู้วิจัยจึงขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณคณาจารย์ภาควิชาวิจัยและจิตวิทยาการศึกษาทุกท่านที่ได้ประสิทธิประสาทวิชาความรู้ให้กับผู้วิจัย โดยเฉพาะ ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี รองศาสตราจารย์ ดร.เอมอร จังศิริพรปกรณ์ ผู้ช่วยศาสตราจารย์ ดร. ญัฐสุภรณ์ หลาวทองและอาจารย์ ดร. กมลวรรณ ตังธนากานนท์ ที่ให้ความห่วงใย ให้ข้อเสนอแนะและคำปรึกษาที่เป็นประโยชน์จากการเรียนวิชาสัมมนาวิทยานิพนธ์ระดับดุษฎีบัณฑิต

ขอกราบขอบพระคุณ ดร.รังสรรค์ มณีเล็ก ที่ให้คำปรึกษาและความช่วยเหลือในการดำเนินการจัดทำวิทยานิพนธ์และขอกราบขอบพระคุณอาจารย์สมาน สำเลิศรัมย์ ที่ให้สื่อและเอกสารที่เป็นประโยชน์ในการกำหนดกรอบความคิดเรื่องการบวกลบเศษส่วน รวมทั้งขอกราบขอบพระคุณคุณไทรรงค์ เจนการ คุณวรรณณ ชุนศรี และอาจารย์คณิต เพ็ชรปัญญา ที่ได้กรุณาสละเวลาเป็นผู้เชี่ยวชาญและให้ข้อคิดเห็นซึ่งช่วยให้การจัดทำแผนภาพลำดับขั้นของคุณลักษณะมีความถูกต้อง ชัดเจนมากขึ้น ตลอดจนขอกราบขอบพระคุณโรงเรียน คณาจารย์และนักเรียนที่เป็นกลุ่มตัวอย่างซึ่งให้ความร่วมมือในการเก็บรวบรวมข้อมูลเป็นอย่างดี

ขอขอบพระคุณจุฬาลงกรณ์มหาวิทยาลัยที่ได้ให้ “ทุน 90 ปี จุฬาลงกรณ์มหาวิทยาลัย” และทุนในการนำเสนองานทางวิชาการในต่างประเทศ ตลอดจนขอขอบพระคุณพี่ ๆ เพื่อน ๆ และน้อง ๆ ภาควิชาวิจัยและจิตวิทยาการศึกษา รวมทั้งกัลยาณมิตรทุกท่านที่ไม่ได้เอ่ยนาม ณ ที่นี้ในความปรารถนาดีและกำลังใจที่มีให้ โดยเฉพาะ ดร.พัชรี จันทรพิง ที่ให้ความช่วยเหลือในการเปรียบเทียบคะแนน

วิทยานิพนธ์ฉบับนี้จะสำเร็จลงไม่ได้หากไม่ได้รับความรัก ความห่วงใย ความช่วยเหลือ การสนับสนุน และเป็นกำลังใจให้จากทุกคนอันเป็นที่รักยิ่งในครอบครัว โดยเฉพาะคุณชวินพัฒน์ สร้อยสังวาลย์ เด็กชายปิยวัฒน์ สร้อยสังวาลย์ คุณไข่มุก เดชาวาสน์ คุณทิพมาศ คำแดง คุณฉ่ำ สร้อยสังวาลย์ และคุณประทีน สร้อยสังวาลย์ ผู้วิจัยจึงขอกราบขอบพระคุณบุคคลดังกล่าวด้วยความซาบซึ้งยิ่ง คุณค่าและประโยชน์ที่ได้รับจากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแต่คุณพ่อคุณแม่ คำแดง และคุณแม่มาลัย คำแดง









## สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฎ
สารบัญแผนภาพ.....	ฏ
สารบัญแผนภูมิ.....	ฒ
บทที่	
1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา.....	1
คำถามการวิจัย.....	8
วัตถุประสงค์ของการวิจัย.....	9
ขอบเขตของการวิจัย.....	9
ข้อจำกัดของการวิจัย.....	10
คำจำกัดความที่ใช้ในการวิจัย.....	10
ประโยชน์ที่คาดว่าจะได้รับ.....	12
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	13
ตอนที่ 1 มโนทัศน์เบื้องต้นเกี่ยวกับการวิจัยทางการศึกษา.....	13
ความหมายของการวิจัยทางการศึกษา.....	13
ประเภทของการวิจัยทางการศึกษา.....	15
รูปแบบและวิธีการที่ใช้ในการวิจัยทางการศึกษา.....	16
งานวิจัยเกี่ยวกับการวิจัยทางการศึกษาในประเทศไทย.....	16
ตอนที่ 2 การประเมินเชิงวิจจัยโดยใช้โมเดลลำดับชั้นของคุณลักษณะ.....	23
แนวคิดที่เกี่ยวข้องกับโมเดลลำดับชั้นของคุณลักษณะ: การพัฒนา Rule Space Model.	23
มโนทัศน์เบื้องต้นเกี่ยวกับโมเดลลำดับชั้นของคุณลักษณะ.....	33
ตอนที่ 3 การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์.....	59
มโนทัศน์เบื้องต้นกับการทดสอบแบบปรับเหมาะ.....	59
ประเภทของการทดสอบแบบปรับเหมาะ.....	62

บทที่	หน้า
ยุทธวิธีที่ใช้ในการทดสอบแบบปรับเหมาะ.....	62
การให้คะแนนการทดสอบแบบปรับเหมาะ.....	67
การประยุกต์ใช้คอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ.....	69
องค์ประกอบของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์.....	71
ประโยชน์ของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์.....	73
ข้อจำกัดของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์.....	73
งานวิจัยที่เกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ในประเทศไทย.....	74
ตอนที่ 4 งานวิจัยที่เกี่ยวข้อง.....	86
งานวิจัยเกี่ยวกับลำดับชั้นของคุณลักษณะ.....	86
งานวิจัยเกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เพื่อการประเมินเชิงวินิจฉัย.....	88
3 วิธีดำเนินการวิจัย.....	93
ประชากรและกลุ่มตัวอย่าง.....	93
เครื่องมือที่ใช้ในการวิจัย.....	96
ขั้นตอนและวิธีดำเนินการวิจัย.....	98
ระยะที่ 1 การพัฒนาชุดของข้อสอบ แบบทดสอบแบบเขียนตอบบนกระดาษคำตอบ และคลังข้อสอบ.....	99
ขั้นตอนที่ 1 การกำหนดลำดับชั้นของคุณลักษณะ.....	99
ขั้นตอนที่ 2 การจัดทำชุดของข้อสอบ.....	100
ขั้นตอนที่ 3 การจัดทำแบบทดสอบแบบเขียนตอบบนกระดาษคำตอบ.....	102
ขั้นตอนที่ 4 การจัดทำคลังข้อสอบ.....	103
ระยะที่ 2 การพัฒนาโปรแกรมคอมพิวเตอร์เพื่อใช้ในการประเมินเชิงวินิจฉัย.....	104
ขั้นตอนที่ 1 การออกแบบโปรแกรมคอมพิวเตอร์.....	104
ขั้นตอนที่ 2 การสร้างและพัฒนาโปรแกรมคอมพิวเตอร์.....	107
ขั้นตอนที่ 3 การทดลองใช้โปรแกรมคอมพิวเตอร์.....	107
ระยะที่ 3 การประเมินเชิงวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ ในการบวกลบเศษส่วน.....	108

บทที่	หน้า
ชั้นตอนที่ 1 การประเมินเชิงวิจิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ ในการบวกลบเศษส่วน โดยการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์..	108
ชั้นตอนที่ 2 การประเมินเชิงวิจิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ ในการบวกลบเศษส่วน โดยใช้เทคนิคการคิดออกเสียง.....	109
ระยะที่ 4 การศึกษาคุณภาพและความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจิจฉัยที่พัฒนาขึ้น	110
ชั้นตอนที่ 1 การศึกษาคุณภาพของวิธีการประเมินเชิงวิจิจฉัยที่พัฒนาขึ้น.....	110
ชั้นตอนที่ 2 การเปรียบเทียบคุณภาพของวิธีการประเมินเชิงวิจิจฉัยที่พัฒนาขึ้น..	109
ชั้นตอนที่ 3 การศึกษาความคิดเห็นของวิธีการประเมินเชิงวิจิจฉัยที่พัฒนาขึ้น..	111
4 ผลการวิเคราะห์ข้อมูล.....	112
ตอนที่ 1 ผลการพัฒนาวิธีการประเมินเชิงวิจิจฉัยโดยประยุกต์ใช้โมเดลลำดับขั้น ของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์.....	112
ตอนที่ 2 ผลการศึกษาคูณภาพของวิธีการประเมินเชิงวิจิจฉัยที่พัฒนาขึ้น.....	140
ตอนที่ 3 ผลการเปรียบเทียบคุณภาพของวิธีการประเมินเชิงวิจิจฉัยที่พัฒนาขึ้น.....	163
ตอนที่ 4 ผลการศึกษาคูณภาพความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจิจฉัยที่พัฒนาขึ้น.....	166
5 สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ.....	173
สรุปผลการวิจัย.....	177
อภิปรายผลการวิจัย.....	181
ข้อเสนอแนะเพื่อการจัดการเรียนรู้เรื่องการบวกลบเศษส่วน.....	186
ข้อเสนอแนะเพื่อนำวิธีการประเมินเชิงวิจิจฉัยที่พัฒนาขึ้นไปใช้.....	187
ข้อเสนอแนะเพื่อการวิจัยต่อไป.....	191
รายการอ้างอิง.....	192
ภาคผนวก.....	202
ภาคผนวก ก รายชื่องานวิจัยที่เกี่ยวข้องกับการวิจัยทางการศึกษาและแหล่งที่มา.....	202
ภาคผนวก ข เครื่องมือที่ใช้ในการวิจัย.....	214
ภาคผนวก ค ตัวอย่างรายงานผลการประเมิน.....	233
ภาคผนวก ง คู่มือการประเมินเชิงวิจิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) โดยใช้เทคนิคการคิดออกเสียง.....	240

บทที่	หน้า
ภาคผนวก จ	
คู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ในการประเมินเชิงวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6).....	273
ภาคผนวก ฉ	
สภาพทั่วไปของกลุ่มตัวอย่างในการสำรวจและจัดเรียงลำดับชั้นของ คุณลักษณะที่ใช้ในการบวกลบเศษส่วน สภาพทั่วไปของกลุ่มตัวอย่าง ในการประเมินความถูกต้องของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน และประเมินแผนภาพลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน และสภาพทั่วไปของผู้ตอบแบบประเมินความครบถ้วนและความครอบคลุม ของแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน.....	307
ภาคผนวก ช	
ผลการประเมินเชิงวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ ในการบวกลบเศษส่วนจากการทดลองใช้โปรแกรมคอมพิวเตอร์.....	312
ประวัติผู้เขียนวิทยานิพนธ์.....	317

บทที่

หน้า



## สารบัญตาราง

ตารางที่	หน้า
1 ความรู้ทักษะและกระบวนการที่ใช้ในการทำแบบทดสอบ TIMSS-R สำหรับนักเรียน เกรด 8.....	29
2 ประเทศที่มีค่าเฉลี่ยของผลสัมฤทธิ์ในแต่ละคุณลักษณะสูงสุด.....	32
3 เมทริกซ์การตอบข้อสอบที่คาดหวัง คะแนนรวมและคุณลักษณะของผู้สอบ สำหรับชุดของผู้สอบจำนวน 15 คน จากพื้นฐานโครงสร้างของลำดับชั้น ในแผนภาพที่ 2ค .....	40
4 เมทริกซ์การตอบข้อสอบที่คาดหวัง คะแนนรวมและคุณลักษณะของผู้สอบ สำหรับชุดของผู้สอบจำนวน 15 คน จากลำดับชั้นของคุณลักษณะในแผนภาพที่ 3	47
5 ค่าพารามิเตอร์ของข้อสอบจำนวน 15 ข้อ ซึ่งประมาณค่าจากการใช้แบบแผนการตอบ ที่คาดหวังในตารางที่ 4.....	48
6 การจำแนกแบบแผนการตอบข้อสอบ (111000000000000) โดยใช้วิธีการจำแนก เบื้องต้น.....	49
7 การจำแนกแบบแผนการตอบข้อสอบ (101000000000000) โดยใช้วิธีการจำแนก เบื้องต้น.....	50
8 การจำแนกแบบแผนการตอบข้อสอบ (111000000000000) โดยใช้ในการตรวจสอบ การจำแนกเบื้องต้น.....	51
9 รายละเอียดโปรแกรมย่อยที่ใช้ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์..	105
10 คุณลักษณะย่อยที่ใช้ในการบวกลบเศษส่วนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) จำแนกตามคุณลักษณะหลัก.....	114
11 ผลการเปรียบเทียบลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน โดยใช้ กระบวนการให้น้ำหนักรายคู่.....	117
12 เมทริกซ์แสดงแบบแผนการตอบข้อสอบที่คาดหวัง จำนวนคุณลักษณะที่ทำข้อสอบ ได้ถูกต้องและสถานะความรู้ของผู้สอบ.....	123
13 ค่าพารามิเตอร์ของข้อสอบ จำแนกตามฉบับ.....	127
14 ค่าพารามิเตอร์ของข้อสอบในคลัง จำแนกตามคุณลักษณะ.....	128
15 ข้อมูลทั่วไปของนักเรียนที่ตอบแบบสอบถามปัญหาที่เกิดจากการทดลองใช้ โปรแกรมคอมพิวเตอร์.....	135

ตารางที่	หน้า
16	ปัญหาที่เกิดจากการทดลองใช้โปรแกรมคอมพิวเตอร์ตามความคิดเห็นของนักเรียน 136
17	ข้อมูลทั่วไปของนักเรียนที่ตอบแบบสอบถามปัญหาที่เกิดจากการใช้ โปรแกรมคอมพิวเตอร์..... 139
18	ปัญหาที่เกิดจากการใช้โปรแกรมตามความคิดเห็นของนักเรียน..... 140
19	จำนวนนักเรียนจำแนกตามแบบแผนการตอบข้อสอบที่คาดหวังซึ่งวินิจฉัยได้ จากการทดสอบครั้งที่ 1 และการทดสอบครั้งที่ 2..... 141
20	จำนวนนักเรียนจำแนกตามแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนซึ่งวินิจฉัยได้ จากการทดสอบครั้งที่ 1 และการทดสอบครั้งที่ 2..... 149
21	จำนวนนักเรียนที่มีแบบการคิดที่ผิด จำแนกตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้.. 155
22	จำนวนนักเรียนที่มีแบบการคิดที่ผิด จำแนกตามระดับความสามารถของนักเรียน..... 155
23	จำนวนนักเรียนจำแนกตามแบบแผนการตอบข้อสอบที่คาดหวังซึ่งวินิจฉัยได้ จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ครั้งที่ 1 และจากการใช้เทคนิค การคิดออกเสียง..... 156
24	จำนวนนักเรียนจำแนกตามแบบแผนการตอบข้อสอบที่คาดหวังซึ่งวินิจฉัยได้ จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ครั้งที่ 2 และจากการใช้เทคนิค การคิดออกเสียง..... 157
25	จำนวนนักเรียนจำแนกตามแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนซึ่งวินิจฉัยได้ จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 1 และจากการใช้เทคนิค การคิดออกเสียง..... 161
26	จำนวนนักเรียนจำแนกตามแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนซึ่งวินิจฉัยได้ จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 2 และจากการใช้เทคนิค การคิดออกเสียง..... 161
27	จำนวนเวลาที่ใช้ในการทดสอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำแนกตามครั้งที่ทดสอบ..... 162
28	จำนวนข้อสอบที่ใช้ในการทดสอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำแนกตามครั้งที่ทดสอบ..... 163

ตารางที่	ล ม หน้า
29 ผลการเปรียบเทียบความเที่ยงของผลการประเมินเชิงวินิจัยจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์จำนวน 2 ครั้ง จำแนกตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้.....	164
30 ผลการเปรียบเทียบความตรงของผลการประเมินเชิงวินิจัยจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์และจากการใช้เทคนิคการคิดออกเสียงจำแนกตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้.....	165
31 ผลการเปรียบเทียบจำนวนข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์.....	165
32 ผลการเปรียบเทียบเวลาที่ใช้ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์....	166
33 ความคิดเห็นของครูที่มีต่อวิธีการประเมินเชิงวินิจัยที่พัฒนาขึ้น.....	168
34 ข้อมูลทั่วไปของนักเรียนที่ตอบแบบสอบถามความคิดเห็นที่มีต่อการประเมินเชิงวินิจัยที่พัฒนาขึ้น.....	169
35 ความคิดเห็นของนักเรียนที่มีต่อวิธีการประเมินเชิงวินิจัยที่พัฒนาขึ้น.....	172

ตาราง

หน้า

แผนภาพที่	หน้า
1 ตัวอย่างการจำแนกคุณลักษณะที่ใช้ในการตอบคำถาม.....	30
2 โครงสร้างของลำดับชั้น.....	36
3 ลำดับชั้นของคุณลักษณะในการใช้เหตุผลเชิงตรรกะตามทฤษฎีความสามารถ ทางสมอง.....	45
4 ลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6).....	117
5 แผนผังแนวคิดในการจำแนกแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน.....	124
6 ลักษณะการกระจายค่าความยากของข้อสอบในคลังข้อสอบในภาพรวม.....	129
7 ลักษณะการกระจายของค่าความยากของข้อสอบสำหรับคุณลักษณะที่ 1.....	130
8 ลักษณะการกระจายของค่าความยากของข้อสอบสำหรับคุณลักษณะที่ 2.....	130
9 ลักษณะการกระจายของค่าความยากของข้อสอบสำหรับคุณลักษณะที่ 3.....	131
10 ลักษณะการกระจายของค่าความยากของข้อสอบสำหรับคุณลักษณะที่ 4.....	131
11 ค่าความน่าจะเป็นโดยเฉลี่ยในการตอบข้อสอบในคลังข้อสอบในภาพรวม.....	132
12 ค่าความน่าจะเป็นโดยเฉลี่ยในการตอบข้อสอบในคลังข้อสอบ จำแนกตาม คุณลักษณะ.....	132
13 ค่าฟังก์ชันสารสนเทศของคลังข้อสอบ ในภาพรวม.....	133
14 ค่าฟังก์ชันสารสนเทศของคลังข้อสอบ จำแนกตามคุณลักษณะ.....	133
15 จำนวนนักเรียนจำแนกตามแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน สำหรับ คุณลักษณะที่ 1 .....	150
16 จำนวนนักเรียนจำแนกตามแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน สำหรับ คุณลักษณะที่ 2 .....	151
17 จำนวนนักเรียนจำแนกตามแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน สำหรับ คุณลักษณะที่ 3 .....	152
18 จำนวนนักเรียนจำแนกตามแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน สำหรับ คุณลักษณะที่ 4 .....	153

ภาพประกอบ

หน้า

## สารบัญแผนภูมิ

ต

แผนภูมิที่		หน้า
1	ขั้นตอนการทดสอบแบบปรับเหมาะ.....	61
2	ขั้นตอนการทดสอบแบบปรับเหมาะ โดยใช้ยุทธวิธีสองขั้นตอน.....	64
3	ขั้นตอนการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์.....	70
4	สถานะความรู้ที่ใช้ในการบวกลบเศษส่วนของนักเรียน จำแนกตามแบบแผนการตอบข้อสอบที่คาดหวังซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะครั้งที่ 1 และครั้งที่ 2.....	142
5	สถานะความรู้ที่ใช้ในการบวกลบเศษส่วนของนักเรียน จำแนกตามคุณลักษณะซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะครั้งที่ 1 และครั้งที่ 2.....	143
6	สถานะความรู้ที่ใช้ในการบวกลบเศษส่วนของนักเรียน จำแนกตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้.....	145
7	สถานะความรู้ที่ใช้ในการบวกลบเศษส่วนของนักเรียน จำแนกตามระดับความสามารถ.....	145
8	ผลการประเมินเชิงวินิจฉัยของนักเรียนกลุ่มที่ไม่มีสถานะความรู้ที่ใช้ในการบวกลบเศษส่วนจำแนกตามคุณลักษณะซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะครั้งที่ 1 และครั้งที่ 2.....	147
9	สถานะความรู้ที่ใช้ในการบวกลบเศษส่วนของนักเรียน จำแนกตามแบบแผนการตอบข้อสอบที่คาดหวังซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้งและจากการใช้เทคนิคการคิดออกเสียง.....	157
10	สถานะความรู้ที่ใช้ในการบวกลบเศษส่วนของนักเรียน จำแนกตามคุณลักษณะซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้งและจากการใช้เทคนิคการคิดออกเสียง.....	158
11	ผลการประเมินเชิงวินิจฉัยของนักเรียนกลุ่มที่ไม่มีสถานะความรู้ที่ใช้ในการบวกลบเศษส่วนจำแนกตามคุณลักษณะซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้งและจากการใช้เทคนิคการคิดออกเสียง.....	159

# สารบัญแผนภูมิ

ณ

แผนภูมิที่

หน้า



ภาพประกอบ

หน้า

# บทที่ 1

## บทนำ

### ความเป็นมาและความสำคัญของปัญหา

การประเมินเชิงวินิจฉัย (Diagnostic Assessment) เป็นการรวบรวมข้อมูลเพื่อตัดสินจุดเด่น จุดด้อย ปัญหาอุปสรรค รวมทั้งสาเหตุที่ทำให้เกิดปัญหาอุปสรรคเหล่านั้น โดยมีเป้าหมายเพื่อการปรับปรุงแก้ไขและการพัฒนา ส่วนใหญ่ใช้ในทางการแพทย์เพื่อศึกษาอาการหรือพฤติกรรมเบี่ยงเบนที่เป็นปัญหาและค้นหาสาเหตุเพื่อรักษาอาการหรือแก้ไขปรับปรุงพฤติกรรมเบี่ยงเบนนั้น ๆ ส่วนทางการศึกษามีการใช้การประเมินเชิงวินิจฉัยอย่างแพร่หลายมากขึ้น ทั้งในบริบทของการศึกษาสำหรับบุคคลที่มีความต้องการพิเศษ เช่น ความบกพร่องทางการเรียนรู้ (learning disability) และบริบทของการศึกษาสำหรับบุคคลทั่วไปซึ่งเกี่ยวข้องกับจัดการเรียนรู้ โดยพยายามบ่งชี้ถึงจุดเด่น จุดด้อย ข้อบกพร่อง ปัญหาหรืออุปสรรคที่ทำให้นักเรียนไม่สามารถบรรลุวัตถุประสงค์การเรียนรู้ที่กำหนดไว้ได้ (Gipps, 1995; Hopkins and Antes, 1990) สารสนเทศที่ได้จากการประเมินจะเป็นแนวทางให้ครูหาวิธีในการจัดการเรียนรู้ หรือมาตรการในการจัดส่งเสริมให้กับนักเรียน เพื่อให้ นักเรียนประสบความสำเร็จในการเรียนรู้และสามารถพัฒนาตนเองได้อย่างเต็มตามศักยภาพ การประเมินเชิงวินิจฉัยจึงเป็นกระบวนการสำคัญในการเชื่อมโยงและสร้างความสอดคล้องระหว่างองค์ประกอบหลักของการจัดการศึกษา ได้แก่ หลักสูตร การจัดการเรียนรู้และการวัดและประเมินผล อันเป็นแนวทางหนึ่งในการพัฒนาการจัดการศึกษาให้เกิดประสิทธิภาพสูงสุด

แนวทางการประเมินเชิงวินิจฉัย จำแนกได้ 2 ลักษณะ ลักษณะแรก คือ การวัดสิ่งที่ขาดหายไป (deficit measurement) มีจุดเน้นที่จุดด้อยหรือความบกพร่อง (weaknesses) ได้รับแนวคิดมาจากการวินิจฉัยทางการแพทย์ซึ่งพยายามจำแนกผู้ป่วยตามอาการหรือความบกพร่อง แนวทางนี้จะใช้มากในเรื่องของความบกพร่องทางการเรียนรู้ ลักษณะที่สอง คือ การวิเคราะห์ความคลาดเคลื่อน (error analysis) มีจุดเน้นที่ประเภทความผิดพลาดของนักเรียน โดยวินิจฉัยจากแบบแผนความผิดพลาดของนักเรียน ในอดีตผู้ที่ใช้การพิจารณาวินิจฉัยลักษณะนี้จะเป็นผู้เชี่ยวชาญทางด้านเนื้อหา มากกว่านักวัดทางจิต แต่ในระยะหลังมีการใช้พื้นฐานการวัดทางจิตช่วยในการวินิจฉัยมากขึ้น ซึ่งได้รับความนิยมนำไปใช้กันอย่างแพร่หลายมาเป็นลำดับ (Bejar, 1984)

การวินิจฉัยที่มีประสิทธิภาพจะต้องให้สารสนเทศที่ถูกต้องเป็นประโยชน์ต่อการพัฒนา นักเรียนและการจัดการเรียนรู้ วิธีการหนึ่งซึ่งมีการดำเนินการที่ชัดเจนและให้ผลการวินิจฉัยที่ประโยชน์

ก็คือ การใช้แบบทดสอบ โดยเฉพาะแบบทดสอบที่สร้างขึ้นโดยมีเป้าหมายเพื่อการวินิจฉัย ที่เรียกว่า แบบทดสอบเชิงวินิจฉัย การวินิจฉัยผลที่ได้จากแบบทดสอบส่วนใหญ่ใช้วิธีการวิเคราะห์ความคลาดเคลื่อนทั้งจากคะแนนสอบ (test score) และจากแบบแผนการตอบข้อสอบ (item response patterns) แต่จากการดำเนินการที่ผ่านมา พบว่าคะแนนสอบเพียงอย่างเดียวไม่เพียงพอสำหรับการวินิจฉัยที่มีประสิทธิภาพ เนื่องจากไม่สามารถระบุจุดเด่น จุดด้อย ปัญหา อุปสรรคและสาเหตุที่มีต่อการเรียนรู้ของนักเรียนแต่ละคนได้อย่างชัดเจน ด้วยเหตุนี้การวินิจฉัยโดยใช้การวิเคราะห์จากแบบแผนการตอบข้อสอบจึงถูกนำไปใช้มากกว่า แต่การดำเนินการในระยะแรกพบว่ามีข้อจำกัดในเรื่องของแบบแผนการตอบข้อสอบสำหรับแต่ละคะแนนสอบมีจำนวนมากเกินไป การดำเนินการในระยะต่อมาจึงใช้การวิเคราะห์จากแบบแผนการตอบข้อสอบที่ผิดปกติ (unusual response pattern) แทนการใช้การวิเคราะห์จากแบบแผนการตอบข้อสอบทั้งหมด (Hamish and Linn, 1981)

การวินิจฉัยโดยใช้การวิเคราะห์จากแบบแผนการตอบข้อสอบจำแนกแนวคิดได้เป็นสองกลุ่ม กลุ่มแรกเป็นการพัฒนาดัชนีรายบุคคล (personal-index approach) เพื่อบ่งชี้ระดับความผิดปกติของแบบแผนการตอบข้อสอบของแต่ละบุคคล ซึ่งจำแนกเป็น การพัฒนาที่ใช้แบบแผนการตอบถูกและผิดเพื่อวิเคราะห์เกี่ยวกับผลการสอบ เช่น การใช้ทฤษฎีการตอบสนองข้อสอบ (item response theory: IRT) หรือการใช้การทดสอบไคสแควร์ ( $\chi^2$ ) ในการทดสอบความเหมาะสมกับบุคคล (person fit) และการพัฒนาโดยใช้แบบแผนการตอบถูกและผิดรายบุคคล โดยพิจารณาจากความสัมพันธ์ของแบบแผนการตอบข้อสอบกับคะแนนรวม (Hamish and Linn, 1981; สำราญ มีแจ้ง, 2533) กลุ่มที่สองเป็นการพัฒนาวิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนจากแนวคิดการประเมินกฎ (rule-assessment method) ซึ่งเอกสารและงานวิจัยเกี่ยวกับจิตวิทยาด้านพุทธิปัญญา (cognitive psychology) เซอร์ปัญญา (artificial intelligence) และการวัดทางจิต (psychometric) แสดงให้เห็นถึงการพัฒนาวิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนของนักเรียน จำแนกเป็น 3 วิธี วิธีแรกใช้การตัดสินใจแบบต้นไม้ (binary decision tree method) ในบริบทของปัญหาความสมดุล พัฒนาโดย Siegler ปี ค.ศ. 1976 และ ปี ค.ศ. 1978 ซึ่งมีความเชื่อว่าการเจริญงอกงามทางความคิดและการมีมโนทัศน์ที่ถูกต้องสามารถแสดงออกมาในรูปของการเพิ่มขึ้นของกฎหรือแบบการคิดที่มีศักยภาพในการแก้ปัญหา การพิจารณามโนทัศน์ที่คลาดเคลื่อนจึงสามารถพิจารณาได้จากแบบการคิดที่ใช้ในการแก้ปัญหาทั้งหมด (ศิริเดช สุชีวะ, 2538) วิธีที่สอง วิธีการใช้โครงสร้างหน้าที่ของงาน (function method) พิจารณาวินิจฉัยจากความผิดพลาดที่เกิดขึ้นระหว่างการปฏิบัติงานซึ่งมีผลต่อความสมบูรณ์ของงาน พัฒนาโดย Anderson ปี ค.ศ. 1974 และ ปี ค.ศ. 1981 และวิธีสุดท้ายมีแนวคิดพื้นฐานมาจากความสนใจโครงสร้างความรู้และพัฒนาการทาง

ปัญญาของมนุษย์ วินิจฉัยมโนทัศน์ที่คลาดเคลื่อนจากจำนวนความผิดพลาดของกฎหรือแบบการคิดที่ใช้ในการแก้ปัญหา ซึ่งได้พัฒนาโปรแกรมคอมพิวเตอร์เพื่อช่วยในการวิเคราะห์ด้วย วิธีการนี้พัฒนาโดยกลุ่มนักวิจัยด้านเซา์ปัญญา (artificial intelligence researchers) มีการพัฒนาแนวคิดตั้งแต่ปี ค.ศ. 1978 ถึง ปี ค.ศ. 1983 (Tatsuoka, 1983)

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการประเมินเชิงวินิจฉัยทางการศึกษาที่ผ่านมาพบว่า ปัจจุบันนักการศึกษาให้ความสนใจแนวคิดทฤษฎีพุทธิปัญญา (cognitive theory) มาใช้ในการวินิจฉัยมากขึ้น ความเกี่ยวข้องระหว่างแนวคิดทฤษฎีพุทธิปัญญา โดยเฉพาะจิตวิทยาด้านพุทธิปัญญา (cognitive psychology) และการวัดทางจิต (psychometrics) แสดงให้เห็นอย่างชัดเจนตั้งแต่ ปี ค.ศ. 1989 โดย Richard Snow และ David Lohman ได้เสนอแนวคิดเรื่อง "Implications of cognitive psychology for educational measurement" ในหนังสือ Educational Measurement พิมพ์ครั้งที่ 3 ซึ่งชี้ให้เห็นว่าการทำให้การวัดทางจิตเป็นที่ยอมรับในสังคมจะต้องใช้จิตวิทยาในการสร้างแบบทดสอบ (psychology of test performance) เพื่อเข้าใจลำดับโครงสร้างคะแนนและให้ความหมายผลการทดสอบที่ถูกต้อง แนวคิดดังกล่าวได้รับการยอมรับจากนักวิจัยและนักปฏิบัติจำนวนมาก เนื่องจากแบบทดสอบทางการศึกษาส่วนใหญ่มีพื้นฐานมาจากการแก้ปัญหาทางพุทธิปัญญา (cognitive problem-solve task) จิตวิทยาทางด้านพุทธิปัญญาจึงกลายมาเป็นพื้นฐานของงานวิจัยเกี่ยวกับการวัดทางจิต โดยมีการพัฒนาโมเดลวินิจฉัยทางพุทธิปัญญา (cognitive diagnostic model :CDMs) จำนวนมาก ดังจะเห็นได้จากเอกสารทางการศึกษาและการวัดทางจิตต่าง ๆ เช่น หนังสือ บทความในวารสารหรือสิ่งตีพิมพ์ที่เน้นการนำจิตวิทยาด้านพุทธิปัญญาไปใช้ในการวัดผลทางการศึกษา (Gierl, Leighton and Hunka, 2000)

โมเดลวินิจฉัยทางพุทธิปัญญาที่ได้รับความนิยมนำไปใช้ในการวินิจฉัยทางการศึกษามากโมเดลหนึ่ง ก็คือ rule space model ซึ่งพัฒนาขึ้นโดย KiKumi K. Tatsuoka เนื่องจากเห็นว่าวิธีการวินิจฉัยที่ผ่านมาไม่สามารถวินิจฉัยแหล่งของมโนทัศน์ที่คลาดเคลื่อนได้ ทำให้สารสนเทศที่ได้จากการวินิจฉัยไม่สามารถใช้ประโยชน์เพื่อการปรับปรุงแก้ไขนักเรียนและการจัดการเรียนรู้ของครูได้อย่างแท้จริง rule space model เป็นโมเดลที่พัฒนาขึ้นจากแนวคิดวิธีการประเมินกฎ (rule-assessment method) และวิธีดัชนีบุคคล (personal-index approach) โดยประยุกต์ใช้ rule space model ซึ่งเป็นโมเดลที่พัฒนาเพื่อใช้ในบริบทของการแก้ปัญหาทางวิศวกรรมศาสตร์และวิทยาศาสตร์กายภาพในการเปรียบเทียบความใกล้เคียงของแบบแผนการตอบข้อสอบกับแบบการคิดแบบต่าง ๆ โดยใช้จุดในเวกเตอร์สเปซของมิติระหว่างดัชนีบ่งชี้ความผิดปกติของแบบแผนการตอบข้อสอบกับความสามารถของผู้สอบ ซึ่งต่อมาได้นำทฤษฎีทางพุทธิปัญญาเข้ามาใช้ในการ

วินิจฉัยโดยปรับเปลี่ยนแนวคิดในการวินิจฉัยจากแบบการคิดที่ผิดมาเป็นการวินิจฉัยที่ให้สารสนเทศเกี่ยวกับจุดเด่นจุดด้อย ซึ่งเป็นสถานะความรู้จากการพิจารณาว่าการทำข้อสอบได้ถูกต้องต้องใช้ความรู้และกระบวนการทางพุทธิปัญญาใดบ้าง rule space model ที่พัฒนาขึ้นตามแนวคิดทฤษฎีพุทธิปัญญาจึงมีคุณลักษณะเป็นโมเดลทางคณิตศาสตร์เพื่อจำแนกแบบแผนการตอบข้อสอบของผู้สอบไปสู่แบบแผนความรู้ (attribute-mastery patterns) จากการใช้ทักษะทางพุทธิปัญญาที่แตกต่างกัน

แม้ว่า rule space model จะเป็นวิธีการประเมินเชิงวินิจฉัยจากแบบแผนการตอบข้อสอบที่ได้รับความนิยมนำไปใช้อย่างแพร่หลาย โดยให้สารสนเทศจากการวินิจฉัยที่เป็นประโยชน์ทั้งในส่วนของการประเมินระดับชั้นเรียนและการประเมินระดับกว้าง (large scale assessment) ในรูปแบบของสถานะความรู้และความไม่รู้อย่างชัดเจนของนักเรียน แต่ rule space model มีข้อจำกัดที่ทำให้ผลการวินิจฉัยได้ไม่สมบูรณ์ เนื่องจากการวินิจฉัยดำเนินการหลังจากสร้างแบบทดสอบเรียบร้อยแล้ว การจำแนกคุณลักษณะที่ต้องใช้ในการตอบข้อสอบแต่ละข้อจึงไม่ครอบคลุมคุณลักษณะหรือความรู้แต่ละเรื่องของผู้สอบต้องมี ด้วยเหตุนี้จึงได้มีการพัฒนาโมเดลลำดับชั้นของคุณลักษณะ (attribute hierarchy model) ขึ้น โมเดลลำดับชั้นของคุณลักษณะเป็นโมเดลทฤษฎีการตอบสนองข้อสอบทางพุทธิปัญญา (Cognitive item response theory model) พัฒนาขึ้นโดย Jacqueline P. Leighton, Mark J. Gierl และ Stephen M. Hunka ในปี ค.ศ. 2000

โมเดลลำดับชั้นของคุณลักษณะพัฒนาขึ้นโดยมีเป้าหมายเพื่อแสดงให้เห็นความเชื่อมโยงระหว่างทฤษฎีทางพุทธิปัญญา (cognitive theory) และการวัดทางจิต (psychometric practice) โดยมีความเชื่อว่าการวัดทางจิตในปัจจุบันไม่ได้นำทฤษฎีทางพุทธิปัญญามาใช้อย่างแท้จริง ทำให้การพัฒนาแบบทดสอบและการนำผลที่ได้จากแบบทดสอบไปใช้ ขาดความชัดเจนและความน่าเชื่อถือซึ่งไม่เกิดประโยชน์ต่อการพัฒนาการจัดการเรียนรู้และการพัฒนานักเรียนเท่าที่ควร แนวคิดพื้นฐานในการพัฒนานอกจากจะมาจากประโยชน์ของการใช้ทฤษฎีทางพุทธิปัญญาในการวัดทางจิตแล้ว ยังมาจากข้อจำกัดในการใช้ rule space model ของ Tatsuoka ในการประเมินเชิงวินิจฉัยด้วย จึงอาจกล่าวได้ว่า โมเดลลำดับชั้นของคุณลักษณะมีการพัฒนาแนวคิดมาจาก rule space model ของ Tatsuoka โดยมีจุดต่างจาก rule space model ที่สำคัญ คือ การให้คำจำกัดความของคำว่าคุณลักษณะ (attribute) ที่ชัดเจนขึ้น โดยได้กล่าวถึงคุณลักษณะไว้ว่า “คุณลักษณะเป็นการบรรยายวิธีการหรือความรู้ที่ต้องใช้ในการทำงานที่มีขอบเขตเฉพาะเพื่อให้ทำงานนั้นบรรลุตามเป้าหมาย คุณลักษณะจะช่วยสร้างยุทธวิธีที่ใช้ในการแก้ปัญหาและมีลักษณะเป็นพลวัต ชุดของคุณลักษณะของนักเรียนที่มีความสามารถในเวลาหนึ่งอาจจะไม่เหมือนชุดของ

คุณลักษณะของนักเรียนที่มีความสามารถในอีกเวลาหนึ่ง ทั้งนี้อาจเกิดจากพัฒนาการที่เพิ่มขึ้นของนักเรียน ไม่ว่าจะเพิ่มอายุที่เพิ่มขึ้นหรือการเรียนรู้ที่เกิดจากการจัดการเรียนรู้” (Gierl, Leighton and Hunka, 2000) นอกจากนี้จะมีการให้คำจำกัดความที่ชัดเจนขึ้นดังกล่าวแล้ว การนำแนวคิดโมเดลลำดับขั้นของคุณลักษณะไปใช้ในการพัฒนาแบบทดสอบยังกำหนดให้มีการจำแนกคุณลักษณะก่อนการพัฒนาแบบทดสอบและมีการกำหนดให้คุณลักษณะมีความสัมพันธ์เป็นลำดับขั้น ซึ่งเป็นขั้นตอนสำคัญในการช่วยกำหนดลักษณะของข้อสอบแต่ละข้อทำให้การวินิจฉัยจากแบบแผนการตอบข้อสอบให้สารสนเทศที่เป็นประโยชน์และมีความชัดเจนขึ้น แต่การนำโมเดลลำดับขั้นของคุณลักษณะไปใช้ ยังคงมีข้อจำกัดเช่นเดียวกับการประเมินเชิงวินิจฉัยโดยใช้แบบแผนการตอบข้อสอบด้วยวิธีอื่น ๆ โดยเฉพาะประเด็นการจำแนกแบบแผนการตอบข้อสอบที่สังเกตได้ กรณีไม่เป็นไปตามแบบแผนการตอบข้อสอบที่คาดหวัง เช่น จากการทำข้อสอบ 15 ข้อ ผู้สอบมีแบบแผนการตอบข้อสอบที่สังเกตได้เป็น 101000000000000 แสดงให้เห็นว่าผู้สอบตอบข้อสอบข้อที่ 1 และข้อที่ 3 ได้ถูกต้อง ซึ่งถ้าลำดับขั้นของคุณลักษณะกำหนดให้ผู้สอบต้องทำข้อสอบข้อที่ 1 และข้อที่ 2 ได้ถูกต้อง จึงจะสามารถทำข้อสอบข้อที่ 3 ได้ถูกต้องแล้ว จะเห็นได้ว่าแบบแผนการตอบข้อสอบที่สังเกตได้ดังกล่าวไม่สอดคล้องกับแบบแผนการตอบข้อสอบที่คาดหวังได้เลย การไม่สามารถจำแนกแบบแผนการตอบข้อสอบได้หรือการจำแนกได้อย่างไม่ชัดเจนดังกล่าวจึงส่งผลต่อความถูกต้องในการจำแนกแบบแผนการรับรู้จากการประเมินสถานะความรู้ (knowledge state) หรือการพิจารณาจากชุดของคุณลักษณะ (attribute vector) ซึ่งเป็นหัวใจสำคัญของการประเมินเชิงวินิจฉัยโดยใช้การพิจารณาจากแบบแผนการตอบข้อสอบ

สาเหตุสำคัญที่ทำให้เกิดข้อจำกัดข้างต้นอาจจำแนกได้ 2 ประการ ประการแรกเป็นสาเหตุที่เกิดจากความไม่ถูกต้องของลำดับขั้นของคุณลักษณะ กล่าวคือ ลำดับขั้นของคุณลักษณะที่กำหนดไม่ได้สะท้อนถึงทักษะหรือกระบวนการทางพุทธิปัญญาที่ใช้ในการตอบข้อสอบที่เป็นลำดับขั้นอย่างแท้จริง เช่น เป็นทักษะหรือกระบวนการทางพุทธิปัญญาที่อยู่ในลำดับขั้นเดียวกัน แต่นำมาจัดอยู่ในลำดับขั้นที่สูงกว่าหรือต่ำกว่า เป็นต้น ประการที่สองเป็นสาเหตุจากความไม่น่าเชื่อถือของแบบแผนการตอบข้อสอบ กล่าวคือ แบบแผนการตอบข้อสอบไม่ได้แสดงถึงความสามารถของผู้สอบอย่างแท้จริง การตอบถูกไม่ได้แสดงถึงการมีทักษะหรือกระบวนการทางพุทธิปัญญานั้น เช่นเดียวกับการตอบไม่ถูกไม่ได้แสดงถึงการไม่มีทักษะหรือกระบวนการทางพุทธิปัญญานั้น เกิดขึ้นได้ในกรณีที่ผู้สอบมีการเดาข้อสอบหรือคัดลอกข้อสอบจากผู้อื่น

การลดข้อจำกัดในการจำแนกแบบแผนการตอบข้อสอบจากการใช้โมเดลลำดับขั้นของคุณลักษณะจึงพิจารณาดำเนินการได้ใน 2 ลักษณะ ลักษณะแรก ก็คือ การให้ความสำคัญกับ

การกำหนดลำดับชั้นของคุณลักษณะให้มีความถูกต้อง ส่วนลักษณะที่สอง ก็คือ การใช้วิธีการทดสอบที่ลดข้อจำกัดในเรื่องของการเดาข้อสอบและการคัดลอกข้อสอบจากผู้อื่น ซึ่งวิธีการทดสอบที่น่าสนใจนำมาประยุกต์ใช้ ก็คือ การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เนื่องจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เป็นการทดสอบที่ผู้สอบแต่ละคนไม่จำเป็นต้องทำข้อสอบฉบับเดียวกันและไม่จำเป็นต้องทำข้อสอบทุกข้อ ทำให้ผู้สอบไม่เกิดความเบื่อหน่าย กรณีที่ข้อสอบง่ายเกินไปหรือยากเกินไปซึ่งเป็นแนวทางหนึ่งในการลดการเดาข้อสอบ ประกอบกับในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ผู้สอบจะได้ทำข้อสอบที่มีการคัดเลือกข้อสอบให้ตรงกับความสามารถของผู้สอบทำให้ผู้สอบแต่ละคนมีโอกาสทำข้อสอบในข้อที่แตกต่างกัน ผู้สอบจึงไม่สามารถคัดลอกข้อสอบจากผู้อื่นได้ การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์จึงช่วยลดข้อจำกัดของการทดสอบแบบดั้งเดิมโดยใช้การเขียนตอบบนกระดาษคำตอบ (paper and pencil) ตลอดจนช่วยเพิ่มประสิทธิภาพในการวัดและการประมาณค่าความสามารถของผู้สอบ ทั้งในส่วนของการลดภาระงานในการทดสอบและการเพิ่มความถูกต้องแม่นยำในการประมาณค่าความสามารถที่แท้จริงของผู้สอบ

จากประโยชน์ของการใช้โมเดลลำดับชั้นของคุณลักษณะในการพัฒนาแบบทดสอบซึ่งให้สารสนเทศที่สามารถอธิบายได้ด้วยทฤษฎีทางพุทธิปัญญา ประกอบกับประโยชน์ของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ในการเพิ่มประสิทธิภาพของการวัดและการประมาณค่าความสามารถของผู้สอบ ผู้วิจัยจึงสนใจที่จะพัฒนาวิธีการประเมินเชิงวินิจฉัย โดยประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ในการวินิจฉัยสถานะความรู้ (knowledge states or latent knowledge states) จากแบบแผนการตอบข้อสอบที่ตอบถูก และแบบการคิดที่ผิด (incorrect rules) จากแบบแผนการตอบข้อสอบที่ตอบผิด รวมทั้งสนใจที่จะศึกษาคุณภาพของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น โดยพิจารณาจากความเที่ยงและความตรง จำนวนเวลาที่ใช้ในการทดสอบและจำนวนข้อสอบที่ใช้ในการทดสอบ พร้อมทั้งเปรียบเทียบคุณภาพของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น เมื่อเกณฑ์ที่ใช้ในการตัดสินความรอบรู้ในแต่ละลำดับชั้นแตกต่างกัน ทั้งนี้เพื่อให้มั่นใจได้ถึงความถูกต้อง เที่ยงตรงของผลการประเมิน นอกจากนี้ยังสนใจที่จะศึกษาความคิดเห็นที่มีต่อวิธีการประเมินเชิงวินิจฉัยในด้านความยากง่ายในการใช้ ความเป็นไปได้ในการนำไปใช้ ประโยชน์ที่ได้รับจากการวินิจฉัยและความพึงพอใจที่มีต่อการใช้ด้วย ซึ่งจะเป็นข้อมูลเบื้องต้นที่เป็นประโยชน์ต่อการนำวิธีการที่พัฒนาขึ้นไปใช้ต่อไป

วิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น มีจุดเน้นสำคัญที่ขั้นตอนการพัฒนาเครื่องมือการประเมิน ซึ่งประกอบไปด้วยการสร้างแบบทดสอบและการพัฒนาโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะ ในขั้นตอนการสร้างแบบทดสอบใช้แนวคิดจากโมเดลลำดับชั้นของ

คุณลักษณะในการกำหนดลักษณะของข้อสอบและกำหนดแบบแผนการตอบข้อสอบที่คาดหวัง ซึ่งแสดงให้เห็นถึงสถานะความรู้ของผู้สอบ กรณีที่ผู้สอบทำข้อสอบในลำดับขั้นใด ๆ ได้ถูกต้อง แสดงว่าผู้สอบมีความรู้สำหรับคุณลักษณะนั้น ส่วนกรณีที่ผู้สอบทำข้อสอบในลำดับขั้นใด ๆ ได้ไม่ถูกต้องแสดงว่าผู้สอบไม่มีความรู้สำหรับคุณลักษณะนั้น วิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น สามารถวินิจฉัยถึงแบบการคิดที่ผิดซึ่งผู้สอบใช้ในการทำข้อสอบได้ โดยประยุกต์ใช้วิธีการย้อนรอย กระบวนการคิดซึ่งพัฒนาโดย คิริเดช สุชีวะ (2538) เนื่องจากเป็นวิธีที่สามารถวินิจฉัยมโนทัศน์ที่ คลาดเคลื่อนได้อย่างมีความตรงและความเที่ยงตามวัตถุประสงค์ที่ตั้งไว้ สามารถให้ผลการวินิจฉัย กับนักเรียนทุกคน ไม่ว่าจะทำข้อสอบผิดเนื่องจากไม่รอบคอบ เดาสุ่มหรือมีแบบการคิดที่ผิด ผลการ ตอบข้อสอบของผู้สอบจะถูกบันทึกด้วยโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะ ทั้งคำตอบที่ตอบถูกและคำตอบที่ตอบผิด เพื่อประโยชน์ 2 ประการ ประการแรกใช้ในการประมาณ ค่าความสามารถของผู้สอบและคัดเลือกข้อสอบข้อต่อไปเพื่อให้ได้ข้อสอบที่สอดคล้องกับ ความสามารถของผู้สอบ ประการที่สองใช้เป็นข้อมูลเพื่อการวินิจฉัยเมื่อการทดสอบเสร็จสิ้นลง

จุดเด่นของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นมี 3 ประการ ประการแรก เป็นการพัฒนา แบบทดสอบซึ่งเป็นเครื่องมือการประเมินเชิงวินิจฉัยที่ใช้ลำดับขั้นของคุณลักษณะตามแนวคิดทฤษฎี พุทธิปัญญาทำให้สามารถกำหนดลักษณะของข้อสอบและจำนวนข้อสอบได้อย่างชัดเจน รวมทั้งให้ สารสนเทศเชิงวินิจฉัยที่เป็นสถานะความรู้ซึ่งอธิบายได้ด้วยความรู้ ทักษะและกระบวนการทางปัญญา ตามแนวคิดทฤษฎีพุทธิปัญญา อันเป็นประโยชน์ต่อการพิจารณาแนวทางการพัฒนาการจัดการเรียนรู้ และการพัฒนานักเรียนมากกว่าการพัฒนาแบบทดสอบโดยใช้กรอบเนื้อหาเหมือนในอดีตที่ผ่านมา ประการที่สอง เป็นการประยุกต์ใช้ทฤษฎีการทดสอบแนวใหม่ (modern test theory) ในการประเมิน เชิงวินิจฉัย ซึ่งช่วยเพิ่มประสิทธิภาพของการประเมินทั้งในส่วนของการบริหารจัดการการทดสอบและ ผลการทดสอบ ประการสุดท้าย เป็นการประเมินเชิงวินิจฉัยที่ให้สารสนเทศทั้งจากแบบแผนการตอบ ข้อสอบที่ตอบถูกและแบบแผนการตอบข้อสอบที่ตอบผิด ซึ่งแสดงให้เห็นถึงจุดเด่นและจุดที่ควรพัฒนา ของผู้สอบรายบุคคลทันทีที่การทดสอบเสร็จสิ้น ทำให้ผู้ที่เกี่ยวข้องได้สารสนเทศที่เป็นประโยชน์ต่อการ แก้ไขปรับปรุงการเรียนรู้ของนักเรียนที่ชัดเจน รวดเร็ว นอกจากนี้ยังสามารถประมวลผลการทดสอบซึ่ง แสดงให้เห็นถึงจุดเด่น จุดด้อยของผู้สอบรายกลุ่ม อันจะเป็นประโยชน์ต่อการวางนโยบายและการ วางแผนเพื่อการแก้ไขปรับปรุงการจัดการเรียนรู้และการพัฒนานักเรียนในภาพรวมด้วย

สำหรับคุณลักษณะที่ใช้ในการวินิจฉัย ผู้วิจัยเลือกใช้คุณลักษณะทางคณิตศาสตร์ เรื่อง การบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) เนื่องจาก คณิตศาสตร์เป็นวิชาที่มีลักษณะธรรมชาติเป็นนามธรรมและมีโครงสร้างที่ชัดเจน นักเรียนจะเกิด



การเรียนรู้และมีความเข้าใจคณิตศาสตร์ได้ดีจะต้องใช้ความคิดอย่างสมเหตุสมผลและศึกษาตามลำดับขั้นอย่างมีกระบวนการ กล่าวคือ นักเรียนจะต้องมีทักษะขั้นพื้นฐานที่เหมาะสมจึงจะสามารถถ่ายโยงไปสู่ทักษะหรือเนื้อหาที่สูงกว่าได้อย่างต่อเนื่อง ประกอบกับการบวกลบเศษส่วนเป็นพื้นฐานสำคัญในการเรียนเรื่องทศนิยม และร้อยละซึ่งมีความจำเป็นต่อการเรียนคณิตศาสตร์ในระดับสูง รวมทั้งเป็นเรื่องที่เข้าใจยาก โดยเฉพาะสำหรับนักเรียนระดับการศึกษาขั้นพื้นฐาน ซึ่งมีโมโนทัศน์เกี่ยวกับนามธรรมที่ยังไม่ชัดเจน เนื่องจากเศษส่วนมีลักษณะเป็นสัญลักษณ์แทนจำนวนที่มีทั้งตัวเศษและตัวส่วน การเรียนรู้เรื่องการบวกลบเศษส่วนสำหรับระดับการศึกษาขั้นพื้นฐานจึงมักจะไม่สามารถประสบความสำเร็จ ดังนั้นการใช้ประโยชน์จากลำดับขั้นของคุณลักษณะในการพัฒนาวิธีการประเมินเชิงวินิจฉัยเนื้อหาวิชาคณิตศาสตร์จึงน่าจะมีความเหมาะสมและให้สารสนเทศการประเมินเชิงวินิจฉัยที่เป็นประโยชน์สอดคล้องกับธรรมชาติของวิชามากกว่าเนื้อหาวิชาอื่น ๆ โดยเฉพาะในเรื่องการบวกลบเศษส่วน สำหรับนักเรียนระดับการศึกษาขั้นพื้นฐานซึ่งมักพบปัญหา อุปสรรคหรือข้อบกพร่องในการเรียนรู้อยู่เสมอ

การพัฒนาวิธีการประเมินเชิงวินิจฉัยโดยการประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบปรับเหมาะโดยใช้คอมพิวเตอร์ในครั้งนี้จึงเป็นประโยชน์ต่อการนำลำดับขั้นของคุณลักษณะตามแนวคิดทฤษฎีทางพุทธิปัญญาและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ไปประยุกต์ใช้ให้เกิดประโยชน์สูงสุดต่อการประเมินเชิงวินิจฉัยทางการศึกษา รวมทั้งสารสนเทศที่ได้จากการประเมินเชิงวินิจฉัยในครั้งนี้ยังเป็นประโยชน์ต่อการปรับปรุงพัฒนาการจัดการเรียนรู้และการพัฒนาคุณลักษณะทางคณิตศาสตร์ของนักเรียนให้เป็นที่ไปตามมาตรฐานที่กำหนดต่อไป

### คำถามการวิจัย

1. การประเมินเชิงวินิจฉัยโดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มีวิธีการอย่างไร
2. การประเมินเชิงวินิจฉัยโดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มีคุณภาพเพียงไร
3. เกณฑ์ที่ใช้ในการตัดสินความรอบรู้ในแต่ละลำดับขั้นที่แตกต่างกันส่งผลให้คุณภาพของการประเมินเชิงวินิจฉัยโดยใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แตกต่างกันหรือไม่
4. ผู้ที่เกี่ยวข้องมีความคิดเห็นต่อวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นอย่างไร

## วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาวิธีการประเมินเชิงวิจักษ์ โดยการประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะ และการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ในการประเมินสถานะความรู้และแบบการคิดที่ผิด เรื่อง การบวกลบเศษส่วน ของนักเรียนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) โดยมีวัตถุประสงค์เฉพาะ ดังนี้

เพื่อกำหนดลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน

เพื่อพัฒนาชุดของข้อสอบและโปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ

เพื่อวิจักษ์สถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วนของนักเรียนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6)

2. เพื่อศึกษาคุณภาพของวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น โดยมีวัตถุประสงค์เฉพาะ ดังนี้

เพื่อศึกษาความเที่ยงของผลการประเมินเชิงวิจักษ์จากวิธีการที่

พัฒนาขึ้น

เพื่อศึกษาความตรงของผลการประเมินเชิงวิจักษ์จากวิธีการที่

พัฒนาขึ้น

เพื่อศึกษาจำนวนเวลาที่ใช้ในการทดสอบ

เพื่อศึกษาจำนวนข้อสอบที่ใช้ในการทดสอบ

3. เพื่อเปรียบเทียบคุณภาพของวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น เมื่อเกณฑ์ที่ใช้ในการตัดสินความรอบรู้ในแต่ละลำดับชั้นแตกต่างกัน

4. เพื่อศึกษาความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น ในด้านความยากง่ายในการใช้ ความเป็นไปได้ในการนำไปใช้ ประโยชน์ที่ได้รับจากการวิจักษ์และความพึงพอใจที่มีต่อการใช้

## ขอบเขตของการวิจัย

1. การพัฒนาวิธีการประเมินเชิงวิจักษ์โดยประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ในครั้งนี้ จำกัดขอบเขตการวิจักษ์เฉพาะคุณลักษณะทางคณิตศาสตร์ เรื่องการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6)

2. ตัวแปรที่ใช้ในการศึกษาครั้งนี้ ประกอบด้วย

2.1 ตัวแปรอิสระ ได้แก่ เกณฑ์ในการตัดสินความรอบรู้ในแต่ละลำดับชั้น จำแนกเป็น

2.1.1 ทำข้อสอบในแต่ละลำดับชั้นอย่างน้อย 3 ข้อและตอบข้อสอบได้ถูกต้อง หรือมีแบบการคิดที่ผิดแบบเดียวกัน จำนวน 2 ใน 3 (ร้อยละ 67)

2.1.2 ทำข้อสอบในแต่ละลำดับชั้นอย่างน้อย 4 ข้อและตอบข้อสอบได้ถูกต้อง หรือมีแบบการคิดที่ผิดแบบเดียวกัน จำนวน 3 ใน 4 (ร้อยละ 75)

2.1.3 ทำข้อสอบในแต่ละลำดับชั้นอย่างน้อย 5 ข้อและตอบข้อสอบได้ถูกต้อง หรือมีแบบการคิดที่ผิดแบบเดียวกัน จำนวน 4 ใน 5 (ร้อยละ 80)

2.1.4 ทำข้อสอบในแต่ละลำดับชั้นอย่างน้อย 3 ข้อและตอบข้อสอบได้ถูกต้อง หรือมีแบบการคิดที่ผิดแบบเดียวกัน จำนวน 2 ใน 3 (ร้อยละ 67) และมีความคลาดเคลื่อนมาตรฐาน ในการประมาณค่าต่ำกว่า 0.3

2.1.5 ทำข้อสอบในแต่ละลำดับชั้นอย่างน้อย 4 ข้อและตอบข้อสอบได้ถูกต้อง หรือมีแบบการคิดที่ผิดแบบเดียวกัน จำนวน 3 ใน 4 (ร้อยละ 75) และมีความคลาดเคลื่อนมาตรฐาน ในการประมาณค่าต่ำกว่า 0.3

2.1.6 ทำข้อสอบในแต่ละลำดับชั้นอย่างน้อย 5 ข้อและตอบข้อสอบได้ถูกต้อง หรือมีแบบการคิดที่ผิดแบบเดียวกัน จำนวน 4 ใน 5 (ร้อยละ 80) และมีความคลาดเคลื่อนมาตรฐาน ในการประมาณค่าต่ำกว่า 0.3

## 2.2 ตัวแปรตาม ได้แก่

2.2.1 สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน

2.2.2 คุณภาพของวิธีการประเมินเชิงวิจจัยที่พัฒนาขึ้น จำแนกเป็น

- 1) ความเที่ยงของผลการประเมินเชิงวิจจัยจากวิธีการที่พัฒนาขึ้น
- 2) ความตรงของผลการประเมินเชิงวิจจัยจากวิธีการที่พัฒนาขึ้น
- 3) จำนวนเวลาที่ใช้ในการทดสอบ
- 4) จำนวนข้อสอบที่ใช้ในการทดสอบ

2.2.3 ความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจจัยที่พัฒนาขึ้น

## ข้อจำกัดของการวิจัย

การพัฒนาวิธีการประเมินเชิงวิจจัยโดยประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ในครั้งนี้มีข้อจำกัดที่ไม่สามารถวิจจัยคำตอบผิดซึ่งเกิดจากความไม่รอบคอบของผู้สอบ เช่น การดูเครื่องหมายผิดหรือการบวกลบเลขจำนวนนับผิดได้ เนื่องจากไม่สามารถหาคำตอบผิดซึ่งเกิดจากความไม่รอบคอบเหล่านั้นได้อย่างครบถ้วนและครอบคลุม

## คำจำกัดความที่ใช้ในการวิจัย

**วิธีการประเมินเชิงวินิจัยที่พัฒนาขึ้น** หมายถึง ขั้นตอนและการดำเนินการที่ใช้ในการวัดและตัดสินสถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 - 6) ซึ่งพัฒนาขึ้นโดยประยุกต์ใช้แนวคิดโมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

**โมเดลลำดับขั้นของคุณลักษณะ** หมายถึง แบบจำลองด้านทฤษฎีการตอบสนองข้อสอบทางพุทธิปัญญา (Cognitive item response theory model) ซึ่งเป็นกรนำทฤษฎีทางพุทธิปัญญา (cognitive theory) มาใช้ประโยชน์ในการวัดทางจิต (psychometric practice) ตั้งแต่การพัฒนาแบบทดสอบจนถึงการแปลผลการทดสอบ

**คุณลักษณะ** หมายถึง ความรู้ ทักษะหรือกระบวนการที่ใช้เพื่อให้สามารถแก้ปัญหาเรื่องการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) ได้อย่างถูกต้อง

**ลำดับขั้นของคุณลักษณะ** หมายถึง ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) ซึ่งได้จัดเรียงตามลักษณะความสัมพันธ์เชิงลำดับขั้นจากคุณลักษณะขั้นพื้นฐานไปยังคุณลักษณะเฉพาะขั้นสูงกว่า

**สถานะความรู้** หมายถึง ระดับของความรู้ ทักษะหรือกระบวนการทางปัญญาที่ผู้สอบใช้ในการแก้ปัญหาเรื่องการบวกลบเศษส่วนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) ได้อย่างถูกต้อง ซึ่งพิจารณาได้จากการใช้แบบการคิดในการแก้ปัญหานั้น ๆ

**แบบการคิดที่ถูก** หมายถึง หลักการในการแก้ปัญหาที่ให้คำตอบที่ถูกต้องในทุกเงื่อนไขของปัญหาในเรื่องการบวกลบเศษส่วนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6)

**แบบการคิดที่ผิด** หมายถึง หลักการในการแก้ปัญหาที่ไม่ได้ให้คำตอบที่ถูกต้องในทุกเงื่อนไขของปัญหาในเรื่องการบวกลบเศษส่วนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6)

**คุณภาพในการประเมินเชิงวินิจัย** หมายถึง การให้ผลการประเมินเชิงวินิจัยที่เชื่อถือได้ คุ่มค่าและสอดคล้องกับวัตถุประสงค์ของการประเมิน โดยพิจารณาจาก ความเที่ยงและความตรงของผลการประเมินเชิงวินิจัยจากวิธีการที่พัฒนาขึ้น จำนวนเวลาและจำนวนข้อสอบที่ใช้ในการทดสอบ

**ความคิดเห็นที่มีต่อวิธีการประเมินเชิงวินิจัยที่พัฒนาขึ้น** หมายถึง ทศนะที่มีต่อการประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์เพื่อการประเมินเชิงวินิจัย ในด้านความยากง่ายในการใช้ ความเป็นไปได้ในการนำไปใช้ ประโยชน์ที่ได้รับจากการวินิจฉัยและความพึงพอใจที่มีต่อการใช้ พิจารณาได้จากผลการประเมินโดยใช้แบบสอบถาม

**ความเที่ยงของผลการประเมินเชิงวินิจัยจากวิธีการที่พัฒนาขึ้น** หมายถึง คุณสมบัติของวิธีการประเมินเชิงวินิจัยที่พัฒนาขึ้นในด้านความคงที่ของผลการประเมินเชิง

วินิจฉัย พิจารณาได้จากความสอดคล้องระหว่างผลการประเมินเชิงวินิจฉัยจากการทดสอบซ้ำ (Test-retest reliability)

**ความตรงของผลการประเมินเชิงวินิจฉัยจากวิธีการที่พัฒนาขึ้น** หมายถึง คุณสมบัติของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น ในด้านความสามารถในการประเมินเชิงวินิจฉัยได้ตรงตามสภาพของผู้สอบ (Concurrent validity) พิจารณาได้จากความสอดคล้องระหว่างผลการประเมินเชิงวินิจฉัยจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (computerized adaptive testing) กับผลการประเมินเชิงวินิจฉัยจากการทดสอบโดยใช้เทคนิคการคิดออกเสียง (Think aloud technique)

**เวลาที่ใช้ในการทดสอบ** หมายถึง จำนวนเวลารวมที่ใช้ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ตั้งแต่เริ่มการทดสอบจนสิ้นสุดการทดสอบ

**จำนวนข้อสอบที่ใช้ในการทดสอบ** หมายถึง จำนวนข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ทั้งข้อที่ตอบถูกและข้อที่ตอบผิด

### ประโยชน์ที่คาดว่าจะได้รับ

1. ได้วิธีการประเมินเชิงวินิจฉัยและโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะที่สามารถนำไปใช้ในการประเมินเชิงวินิจฉัยได้อย่างมีประสิทธิภาพ
2. ได้ทราบลำดับขั้นของความรู้ ทักษะและกระบวนการที่ใช้ในการบวกลบเศษส่วนตามแนวคิดทฤษฎีทางพุทธิปัญญา ซึ่งเป็นประโยชน์ต่อการพัฒนาการจัดการเรียนรู้และการพัฒนาให้นักเรียนให้มีคุณภาพและเกิดประสิทธิภาพสูงสุด
3. ได้ทราบแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนในแต่ละลำดับขั้นของคุณลักษณะ ซึ่งเป็นกระบวนการคิดแก้ปัญหาที่แสดงให้เห็นถึงการมีโมโนทัศน์ที่คลาดเคลื่อนของนักเรียน อันจะเป็นประโยชน์ต่อผู้ที่เกี่ยวข้องในการนำแบบการคิดดังกล่าวไปใช้พัฒนาเครื่องมือเพื่อการประเมินเชิงวินิจฉัยและใช้ในการวางแผนพัฒนานักเรียนต่อไป
4. ได้ข้อมูลเชิงวินิจฉัยเกี่ยวกับสถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วนของนักเรียน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) อันจะเป็นประโยชน์ต่อการพิจารณาแนวทางปรับปรุงพัฒนาความรู้ ทักษะและกระบวนการทางปัญญาในการบวกลบเศษส่วนของนักเรียนให้เป็นไปตามมาตรฐานที่กำหนด
5. เป็นแนวทางเบื้องต้นในการนำโมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะไปใช้ให้เกิดประโยชน์ต่อการพัฒนาวิธีวิทยาการด้านการประเมินเชิงวินิจฉัยในประเด็นและบริบทอื่น ๆ

## บทที่ 2

### เอกสารและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการวิจัยในครั้งนี้ จำแนกเป็น 4 ตอน ตอนที่ 1 มโนทัศน์เบื้องต้นเกี่ยวกับการวินิจฉัยทางการศึกษา ประกอบด้วย ความหมายของการวินิจฉัยทางการศึกษา ประเภทของการวินิจฉัยทางการศึกษา รูปแบบและวิธีการที่ใช้ในการวินิจฉัยทางการศึกษาและงานวิจัยเกี่ยวกับการวินิจฉัยทางการศึกษาในประเทศไทย ตอนที่ 2 การประเมินเชิงวินิจฉัยโดยใช้โมเดลลำดับชั้นของคุณลักษณะ ประกอบด้วย การพัฒนา Rule space model ซึ่งเป็นแนวคิดที่เกี่ยวข้องกับโมเดลลำดับชั้นของคุณลักษณะ มโนทัศน์เบื้องต้นเกี่ยวกับโมเดลลำดับชั้นของคุณลักษณะ และตัวอย่างการนำโมเดลลำดับชั้นของคุณลักษณะไปใช้ ตอนที่ 3 การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ประกอบด้วย มโนทัศน์เบื้องต้นเกี่ยวกับการทดสอบแบบปรับเหมาะ ประเภทของการทดสอบแบบปรับเหมาะ ยุทธวิธีที่ใช้ในการทดสอบแบบปรับเหมาะ การให้คะแนนการทดสอบแบบปรับเหมาะ การประยุกต์ใช้คอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ องค์ประกอบของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ประโยชน์และข้อจำกัดของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์และงานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ในประเทศไทย และตอนที่ 4 งานวิจัยที่เกี่ยวข้องประกอบด้วยงานวิจัยเกี่ยวกับลำดับชั้นของคุณลักษณะและงานวิจัยเกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์เพื่อการประเมินเชิงวินิจฉัย มีรายละเอียดในแต่ละตอน ดังนี้

#### ตอนที่ 1 มโนทัศน์เบื้องต้นเกี่ยวกับการวินิจฉัยทางการศึกษา

##### ความหมายของการวินิจฉัยทางการศึกษา

“การวินิจฉัย” มีความหมายทั่วไปหมายถึง การไต่ตรอง การใคร่ครวญหรือการตัดสินใจขาด (ราชบัณฑิตยสถาน, 2546) ส่วนใหญ่ใช้ในทางการแพทย์และทางพฤติกรรมศาสตร์เพื่อศึกษาอาการหรือพฤติกรรมเบี่ยงเบนที่เป็นปัญหาและค้นหาสาเหตุเพื่อรักษาอาการหรือแก้ไขปรับปรุงพฤติกรรมเบี่ยงเบนนั้น ๆ ส่วนทางการศึกษามีการใช้การวินิจฉัยอย่างแพร่หลายมากขึ้น ทั้งในบริบทของการศึกษาสำหรับบุคคลที่มีความต้องการพิเศษ เช่น ความบกพร่องทางการเรียนรู้ (learning disability) และบริบทของการศึกษาสำหรับบุคคลทั่วไป มีนักการศึกษาและนักจิตวิทยาให้ความหมายการวินิจฉัยไว้หลายท่าน แต่ในที่นี้จะนำเสนอเฉพาะความหมายที่ให้โดยนักการศึกษาและนักจิตวิทยาในต่างประเทศ เนื่องจากในบริบทของการศึกษาในประเทศไทยพบเอกสารทางการวัดและประเมินผล

ที่ให้ความหมายการวินิจฉัยไว้จำนวนน้อยมาก ส่วนใหญ่จะให้ความหมายในลักษณะเครื่องมือที่ใช้เพื่อการวินิจฉัยมากกว่า เช่น ความหมายของแบบทดสอบเชิงวินิจฉัย ความหมายของการวินิจฉัยที่ให้โดยนักการศึกษาและนักจิตวิทยาในต่างประเทศ” สรุปได้ดังนี้

Baron และ Bernard (1958) ให้ความหมายไว้ว่า หมายถึง การแปลความหมายข้อมูลเพื่อกำหนดสาเหตุเฉพาะที่เป็นอุปสรรคต่อการเรียนรู้ของนักเรียน

Good (1973) ให้ความหมายไว้ว่า หมายถึง การค้นหาอุปสรรคหรือข้อบกพร่องในการเรียนรู้

Brown (1983) ให้ความหมายไว้ว่า หมายถึง การเปรียบเทียบการแสดงผลของบุคคลเพื่อกำหนดจุดเด่น จุดด้อย

Delandshere (1990) ให้ความหมายไว้ว่า เป็นการทดสอบหรือการเปรียบเทียบลักษณะ (Profile) ของแต่ละบุคคลกับบรรทัดฐาน (norms) หรือเกณฑ์ (criteria) โดยเน้นการวินิจฉัยจุดเด่นและจุดด้อยของบุคคลที่เกี่ยวข้องกับกระบวนการของโรงเรียน เช่น หลักสูตร โปรแกรม การบริหารจัดการ

Hopkins และ Antes (1990) ให้ความหมายไว้ว่า หมายถึง การจำแนกสภาวะที่ไม่ดีหรือไม่เป็นไปตามที่ต้องการในลำดับการเรียนรู้ เพื่อปรับปรุงการเรียนการสอนให้ดีขึ้น

Cohen, Swerdlik และ Smith (1992) ให้ความหมายไว้ว่า หมายถึง การกระทำหรือกระบวนการในการจำแนกหรือกำหนดความผิดปกติ โดยการตรวจพิจารณาหรือการทดสอบ

Gregory (1992) ให้ความหมายไว้ว่า หมายถึง การกำหนดลักษณะที่เป็นอยู่หรือแหล่งของพฤติกรรมที่ไม่ปกติเพื่อจำแนกแบบแผนของพฤติกรรมตามระบบการวินิจฉัยที่ได้รับการยอมรับ

Gipps (1995) ให้ความหมายไว้ว่า หมายถึง การระบุปัญหาเฉพาะที่มีต่อความคิดหรือทักษะของนักเรียนแต่ละบุคคล

จากความหมายการวินิจฉัยข้างต้นสามารถวิเคราะห์จุดเน้นการให้ความหมายจำแนกตามวิธีการที่ใช้ สิ่งทีวินิจฉัย เป้าหมายในการวินิจฉัยและประโยชน์ในการวินิจฉัยได้ดังนี้

1. วิธีการที่ใช้ ได้แก่ การระบุ การกำหนด การจำแนก การเปรียบเทียบ การค้นหา การทดสอบ การตรวจพิจารณา และการแปลความหมายข้อมูล ส่วนสิ่งที่ใช้ประกอบการดำเนินการอาจเป็นแบบแผนของพฤติกรรม บรรทัดฐาน (norms) เกณฑ์ (criteria) หรือ ระบบการวินิจฉัยที่ได้รับการยอมรับ

2. สิ่งทีวินิจฉัย ได้แก่ ปัญหา อุปสรรค ข้อบกพร่อง จุดเด่น จุดด้อย สภาวะที่ไม่ดีหรือไม่เป็นไปตามที่ต้องการ รวมทั้งสาเหตุที่ทำให้เกิดสิ่งต่าง ๆ เหล่านั้น

3. เป้าหมายในการวินิจฉัย ได้แก่ นักเรียนและกระบวนการของโรงเรียน ซึ่งเน้นไปที่แต่ละบุคคลมากกว่ากลุ่มบุคคล

4. ประโยชน์ในการวินิจฉัย ได้แก่ การปรับปรุงการเรียนการสอน

เมื่อพิจารณาผลการวิเคราะห์จุดเน้นการให้ความหมายดังกล่าว สรุปได้ว่า การวินิจฉัย หมายถึง การใช้วิธีการหรือกระบวนการเพื่อจำแนก ระบุ กำหนด เปรียบเทียบ ค้นหา ทดสอบหรือ ตรวจสอบพิจารณาซึ่งแสดงให้เห็นถึงจุดเด่น จุดด้อย ปัญหา อุปสรรคและสาเหตุที่เกี่ยวข้องกับ กระบวนการเรียนรู้ของบุคคล รวมทั้งกระบวนการดำเนินงานของโรงเรียน โดยมีเป้าหมายเพื่อการ ปรับปรุงแก้ไขและการพัฒนา

## ประเภทของการวินิจฉัยทางการศึกษา

Underhill (1972 อ้างถึงใน ปิยาพร ชาวสะอาด, 2541) และดวงเดือน อ่อนน้อม (2533) ได้ จำแนกประเภทของการวินิจฉัยตามระดับความละเอียดในการวินิจฉัยเป็น 3 ประเภท สรุปได้ดังนี้

### 1. การวินิจฉัยแบบทั่วไป (general diagnosis)

การวินิจฉัยแบบทั่วไปเป็นการวินิจฉัยระดับทั่วไป (General level) ขึ้นการสำรวจเพื่อ ทราบระดับความสามารถทั่ว ๆ ไป ของนักเรียนทั้งรายบุคคลและรายกลุ่ม เครื่องมือที่ใช้ส่วนใหญ่ เป็นแบบทดสอบทั้งแบบทดสอบวัดผลสัมฤทธิ์ที่ครูสร้างขึ้นและแบบทดสอบมาตรฐาน อาจใช้การ วินิจฉัยโดยการเปรียบเทียบกับเกณฑ์ปกติ (norm) ซึ่งจะทำให้ทราบข้อบกพร่องในด้านต่าง ๆ ของ นักเรียนอย่างกว้าง ๆ

### 2. การวินิจฉัยแบบวิเคราะห์ (analytical diagnosis)

การวินิจฉัยแบบวิเคราะห์เป็นการวินิจฉัยระดับเฉพาะ (Specific level) โดยการ เก็บรวบรวมข้อมูลอย่างละเอียดเกี่ยวกับความสามารถของนักเรียนเพื่อทราบข้อบกพร่องของ นักเรียนในเรื่องใดเรื่องหนึ่ง เครื่องมือที่ใช้ส่วนใหญ่เป็นแบบทดสอบที่มีเนื้อหาเฉพาะเจาะจง ในเรื่องใดเรื่องหนึ่ง การวินิจฉัยแบบวิเคราะห์มักดำเนินการภายหลังจากการวินิจฉัยแบบทั่วไป เพื่อให้ได้ข้อมูลที่มีรายละเอียดที่ลึกซึ้งมากขึ้น

### 3. การวินิจฉัยแบบคลินิก (Clinical Diagnosis)

เป็นการวินิจฉัยระดับละเอียดลึกซึ้ง (Intensive level) เพื่อให้เห็นถึงสาเหตุของปัญหา หรือข้อบกพร่องที่พบในตัวนักเรียนซึ่งมีความซับซ้อน การใช้ข้อมูลจากแหล่งใดแหล่งหนึ่งเพียง แหล่งเดียวจึงไม่เพียงพอ จำเป็นต้องใช้ข้อมูลจากหลายแหล่งประกอบกัน เช่น การใช้ข้อมูลจาก แบบทดสอบ ประกอบการสัมภาษณ์ การสังเกต การศึกษาสภาพครอบครัว เป็นต้น



## รูปแบบและวิธีการที่ใช้ในการวินิจฉัยทางการศึกษา

ศิริเดช สุชีวะ (2538) ได้จำแนกรูปแบบการวินิจฉัยทางการศึกษาไว้ 2 รูปแบบ ได้แก่ รูปแบบที่ไม่เป็นทางการและรูปแบบที่เป็นทางการ มีลักษณะและวิธีการวินิจฉัยที่ใช้ในแต่ละรูปแบบ ดังนี้

### 1. การวินิจฉัยที่ไม่เป็นทางการ (informal diagnosis)

การวินิจฉัยที่ไม่เป็นทางการเป็นการค้นหาข้อบกพร่องของนักเรียนด้วยวิธีการง่าย ๆ ไม่มีแบบแผนที่แน่นอน วิธีการที่ใช้ ได้แก่ การสังเกต การสอบถาม การสัมภาษณ์ การตรวจผลงาน

### 2. การวินิจฉัยที่เป็นทางการ (formal diagnosis)

การวินิจฉัยที่เป็นทางการเป็นการค้นหาข้อบกพร่องของนักเรียนโดยใช้วิธีการที่สร้างขึ้นมาอย่างเป็นระบบและมีแบบแผน วิธีการที่ใช้ ได้แก่ การใช้แบบทดสอบเชิงวินิจฉัย หรือการวินิจฉัยจากแบบแผนการตอบข้อสอบ โดยใช้ดัชนีบ่งชี้ความผิดปกติของแบบแผนการตอบข้อสอบ การใช้วิธีการ Rule space หรือการใช้การย้อนรอยกระบวนการคิด เป็นต้น

## งานวิจัยเกี่ยวกับการวินิจฉัยทางการศึกษาในประเทศไทย

การศึกษางานวิจัยที่เกี่ยวกับการวินิจฉัยทางการศึกษาในประเทศไทยมีวัตถุประสงค์เพื่อได้ข้อมูลพื้นฐานเกี่ยวกับการวินิจฉัยทางการศึกษาที่ผ่านมาในบริบทสังคมไทย ดำเนินการโดยรวบรวมเอกสารหรือบทความที่เกี่ยวข้องกับการวินิจฉัยทางการศึกษาจากฐานข้อมูลวิทยานิพนธ์และงานวิจัยของมหาวิทยาลัยต่าง ๆ และฐานข้อมูลงานวิจัยของสำนักงานกองทุนสนับสนุนการวิจัย (สกว.) รวมทั้งฐานข้อมูลงานวิจัยของสำนักงานเลขาธิการสภาการศึกษา (สกศ.) ได้งานวิจัยที่ใช้ในการศึกษา จำนวน 74 เรื่อง ส่วนใหญ่เป็นงานวิจัยในลักษณะของวิทยานิพนธ์ ผลจากการวิเคราะห์เนื้อหา พบว่า ส่วนใหญ่เป็นการพัฒนาหรือสร้างแบบทดสอบเชิงวินิจฉัย (ร้อยละ 86.67) รองลงมา ได้แก่ การพัฒนาหรือใช้ดัชนีเพื่อบ่งชี้ความผิดปกติของแบบแผนการตอบข้อสอบ (ร้อยละ 6.67) การใช้แนวคิดการประเมินกฎในการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อน (ร้อยละ 2.67) การวินิจฉัยที่คำนึงถึงสภาพจริงในการแก้โจทย์ปัญหา (ร้อยละ 1.33) การวินิจฉัยและแก้ไขข้อบกพร่องโดยนักเรียน (ร้อยละ 1.33) และการวินิจฉัยโดยใช้เทคนิคการคิดออกเสียง (ร้อยละ 1.33) เนื้อหาที่ใช้ในการศึกษามากที่สุด ได้แก่ วิชาคณิตศาสตร์ (ร้อยละ 73.61) รองลงมา ได้แก่ วิชาวิทยาศาสตร์ซึ่งมีจำนวนเท่ากับวิชาภาษาไทย (ร้อยละ 11.11) และวิชาภาษาอังกฤษ (ร้อยละ 4.17) ส่วนระดับที่ใช้ในการศึกษา พบว่า ระดับช่วงชั้นที่ใช้ในการศึกษามากที่สุด ได้แก่ ระดับช่วงชั้นที่ 2 ชั้นประถมศึกษาปีที่ 4-6 (ร้อยละ 41.43) รองลงมา ได้แก่ ระดับช่วงชั้นที่ 3 ชั้นมัธยมศึกษาปีที่ 1-3 (ร้อยละ 34.29) ระดับช่วงชั้นที่ 4 ชั้นมัธยมศึกษาปีที่ 4-6 (ร้อยละ 15.71) ระดับช่วงชั้นที่ 1 ชั้นประถมศึกษาปีที่ 1-3 (ร้อยละ 7.14) และระดับอุดมศึกษา (ร้อยละ 1.43) มีรายละเอียด ดังนี้

## 1. แบบทดสอบเชิงวินิจฉัย

งานวิจัยที่เกี่ยวกับแบบทดสอบเชิงวินิจฉัย ประกอบด้วย

### 1.1 การสร้างหรือพัฒนาแบบทดสอบเชิงวินิจฉัย

งานวิจัยที่เกี่ยวกับการสร้างหรือการพัฒนาแบบทดสอบเชิงวินิจฉัยส่วนใหญ่ใช้ชื่อเรื่องว่า “การสร้างหรือการพัฒนาแบบทดสอบ” มีบางส่วนที่ใช้ชื่อเรื่องว่า “การวินิจฉัย” แต่รายละเอียดเป็นการสร้างหรือการพัฒนาแบบทดสอบ แบบทดสอบส่วนใหญ่เป็นเนื้อหาในวิชาคณิตศาสตร์ ทั้งในภาพรวมแต่ละระดับชั้นและการวินิจฉัยรายวิชาหรือเนื้อหาย่อยในแต่ละระดับชั้น แบบทดสอบที่พัฒนาขึ้นมีการดำเนินการ 2 ขั้นตอน ขั้นตอนที่ 1 เป็นการสำรวจข้อบกพร่อง วิธีการที่ใช้ได้แก่ การใช้แบบทดสอบและการใช้การสัมภาษณ์ แบบทดสอบที่ใช้เป็นแบบอัตนัยและแบบเติมคำ ส่วนใหญ่ใช้แบบทดสอบเพียง 1 ฉบับ แต่ละฉบับมีจำนวนข้อแตกต่างกันขึ้นอยู่กับเนื้อหาที่ใช้ในการวินิจฉัย ขั้นตอนที่ 2 เป็นการวินิจฉัยข้อบกพร่องวิธีการที่ใช้ส่วนใหญ่ ใช้แบบทดสอบแบบเลือกตอบที่มีตัวเลือกจำนวน 4 หรือ 5 ตัวเลือก แบบทดสอบจะแยกเป็นฉบับย่อย ๆ ตามเนื้อหาย่อย มีจำนวน ตั้งแต่ 1 ถึง 53 ฉบับ ขึ้นอยู่กับลักษณะของเนื้อหาที่ใช้ในการวินิจฉัย สำหรับการวิเคราะห์ข้อสอบ พบว่า ส่วนใหญ่ใช้ทฤษฎีการทดสอบแบบดั้งเดิม (classical test theory) มีการใช้ทฤษฎีการตอบสนองของข้อสอบ (item response theory) จำนวนน้อยมาก ซึ่งในจำนวนนั้นเป็นแบบตรวจให้คะแนน 2 ค่า (dichotomous item response theory) ทั้งหมด ผลการวิจัยส่วนใหญ่พบว่าแบบทดสอบที่พัฒนาขึ้นสามารถใช้ในการวินิจฉัยได้จริง โดยมีค่าความยาก ค่าอำนาจจำแนก ความเที่ยงและความตรงของข้อสอบอยู่ในเกณฑ์ที่สามารถยอมรับได้ (รายชื่องานวิจัยและแหล่งที่มา แสดงในภาคผนวก ก ตารางที่ 1)

### 1.2 การสร้างหรือพัฒนาแบบทดสอบและการแก้ไขข้อบกพร่อง

งานวิจัยที่มีรายละเอียดเกี่ยวกับการสร้างหรือพัฒนาแบบทดสอบและการแก้ไขข้อบกพร่องมีจำนวน 2 เรื่อง เป็นเนื้อหาในวิชาคณิตศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 2 และภาษาไทย ระดับชั้นมัธยมศึกษาปีที่ 2 การวิจัยทั้ง 2 เรื่อง มีการดำเนินการจำแนกเป็น 2 ขั้นตอน ขั้นตอนที่ 1 เป็นขั้นวินิจฉัยข้อบกพร่องทางการเรียน โดยใช้แบบทดสอบจำแนกตามจุดประสงค์หรือเนื้อหาย่อย ขั้นตอนที่ 2 เป็นขั้นแก้ไขข้อบกพร่องทางการเรียน ซึ่งวิชาภาษาไทยใช้แบบฝึกหัด ส่วนวิชาคณิตศาสตร์ใช้บทเรียนสำเร็จรูป ผลการวิจัย ในวิชาคณิตศาสตร์พบว่านักเรียนที่ได้รับการวินิจฉัยและแก้ไขข้อบกพร่องมีผลสัมฤทธิ์ทางการเรียนสูงขึ้น แต่ยังมีบางจุดประสงค์ที่มีคะแนนไม่ผ่านเกณฑ์ สำหรับผลการวิจัยในวิชาภาษาไทย พบว่า นักเรียนส่วนใหญ่มีผลการทดสอบผ่านเกณฑ์ที่กำหนดหลังจากได้รับการวินิจฉัยและการแก้ไขข้อบกพร่อง (รายชื่องานวิจัยและแหล่งที่มา แสดงในภาคผนวก ก ตารางที่ 2)

### 1.3 การเขียนตัวลง

งานวิจัยที่เกี่ยวกับการเขียนตัวลงในแบบทดสอบเชิงวินิจฉัยมีจำนวน 1 เรื่อง มีวัตถุประสงค์เพื่อเปรียบเทียบค่าความยาก ค่าอำนาจจำแนกและความเชื่อมั่นของแบบทดสอบเชิงวินิจฉัยวิชาคณิตศาสตร์ซึ่งเป็นแบบทดสอบแบบเลือกตอบชนิด 5 ตัวเลือก จำนวน 4 ฉบับ ฉบับละ 42 ข้อ โดยสร้างตัวลงด้วยวิธีที่ต่างกัน 4 วิธี ได้แก่ การใช้คำตอบผิดของนักเรียนที่ซ้ำกันมากที่สุด การวิเคราะห์หาคุณภาพจากแบบทดสอบเพื่อการสำรวจ การสุ่มคำตอบผิดของนักเรียนและการให้ครูผู้สอนกำหนด ผลการวิจัย พบว่า แบบทดสอบที่มีค่าความยากมากที่สุด ได้แก่ แบบทดสอบที่สร้างตัวลงโดยใช้คำตอบผิดของนักเรียนที่ซ้ำกันมากที่สุด รองลงมา ได้แก่ แบบทดสอบที่สร้างตัวลงจากการวิเคราะห์หาคุณภาพจากแบบทดสอบเพื่อการสำรวจ ตัวลงซึ่งกำหนดโดยครูและตัวลงโดยวิธีสุ่มคำตอบผิดของนักเรียน แบบทดสอบที่มีค่าอำนาจจำแนกและความเชื่อมั่นสูงสุด ได้แก่ จากการวิเคราะห์หาคุณภาพจากแบบทดสอบเพื่อการสำรวจ รองลงมา ได้แก่ แบบทดสอบที่สร้างตัวลงซึ่งกำหนดโดยครู แบบทดสอบที่สร้างตัวลงโดยใช้คำตอบผิดของนักเรียนที่ซ้ำกันมากที่สุดและแบบทดสอบที่สร้างตัวลงโดยวิธีสุ่มคำตอบผิดของนักเรียน นอกจากนี้ยัง พบว่า การสร้างตัวลงโดยวิธีที่แตกต่างกัน ส่งผลให้ค่าความยากเฉลี่ย ค่าอำนาจจำแนกเฉลี่ยและค่าความเชื่อมั่นของแบบทดสอบไม่แตกต่างกัน (รายชื่องานวิจัยและแหล่งที่มา แสดงในภาคผนวก ก ตารางที่ 3)

## 2. ดัชนีบ่งชี้ความผิดปกติของแบบแผนการตอบข้อสอบ

งานวิจัยที่เกี่ยวกับดัชนีบ่งชี้ความผิดปกติของแบบแผนการตอบข้อสอบ ประกอบด้วย

### 2.1 การพัฒนาดัชนีบ่งชี้ความผิดปกติของแบบแผนการตอบข้อสอบ

งานวิจัยที่เกี่ยวกับการพัฒนาดัชนีบ่งชี้ความผิดปกติของแบบแผนการตอบข้อสอบมีจำนวน 2 เรื่อง งานวิจัยเรื่องแรกเป็นการพัฒนาดัชนีเพื่อแก้ไขข้อบกพร่องของดัชนีของกัตแมนและซาได้ในกรณีการใช้ค่าความยากซึ่งขึ้นอยู่กับความสามารถของกลุ่มในการจัดเรียงข้อสอบ โดยได้พัฒนาดัชนีชี้ความบกพร่องของผู้ตอบแบบทดสอบโดยการวิเคราะห์แบบแผนการตอบข้อสอบซึ่งจัดเรียงข้อสอบตามค่าความยากและเรียงผู้สอบตามระดับความสามารถที่ไม่เกี่ยวกับคะแนนรวมของกลุ่มแต่ขึ้นอยู่กับคุณสมบัติของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ ดัชนีที่พัฒนาขึ้นมีจำนวน 2 ดัชนี ได้แก่ ดัชนีดีเอฟซี 1 และดัชนีดีเอฟซี 2 ซึ่งมีสูตรที่ใช้ในการคำนวณที่แตกต่างกัน ดัชนีดีเอฟซี 1 คำนวณจากค่าความแปรปรวนร่วมระหว่างคะแนนรายข้อกับค่าโอกาสการตอบข้อสอบแต่ละข้อได้ถูกต้องหารด้วยค่าความแปรปรวนร่วมระหว่างคะแนนรายข้อที่มีลักษณะตามกัตแมนสมบูรณ์กับค่าโอกาสการตอบข้อสอบแต่ละข้อได้ถูกต้อง แล้วนำค่าสัดส่วน

ที่คำนวณได้มาลบออกจาก 1 ส่วนดัชนีดีเอฟซี 2 คำนวณจากค่าความแปรปรวนร่วมระหว่างคะแนนรายข้อกับค่าเฉลี่ยของโอกาสการตอบข้อสอบข้อนั้นได้ถูกต้องของทุกคนในกลุ่มหารด้วยค่าความแปรปรวนร่วมระหว่างคะแนนรายข้อที่มีลักษณะตามกัตแมนสมบรูณ์กับค่าเฉลี่ยของโอกาสการตอบข้อสอบข้อนั้นได้ถูกต้องของทุกคนในกลุ่ม แล้วนำค่าสัดส่วนที่คำนวณได้ลบออกจาก 1 ผลจากการศึกษาความตรงของดัชนีที่พัฒนาขึ้น พบว่า ดัชนีที่พัฒนาขึ้นทั้ง 2 มีความตรงในการจำแนกผู้ตอบแบบทดสอบที่มีความบกพร่องและมีความสัมพันธ์กับดัชนีของชาโต้สูง ดัชนีดีเอฟซี 1 มีประสิทธิภาพในการวินิจฉัยความบกพร่องของผู้ตอบแบบทดสอบในแบบแผนการตอบสูงกว่าดัชนีดีเอฟซี 2 และดัชนีของชาโต้ เมื่อผู้สอบมีความสามารถปานกลาง แต่สำหรับผู้สอบที่มีความสามารถสูงและต่ำ ดัชนีดีเอฟซี 2 มีประสิทธิภาพในการวินิจฉัยสูงกว่าดัชนีดีเอฟซี 1 และดัชนีของชาโต้

งานวิจัยเรื่องที่สองเป็นการพัฒนาดัชนีเอฟซี (actual pattern) เพื่อบ่งชี้ว่าแบบแผนการตอบข้อสอบของผู้สอบมีระดับความสอดคล้องกับแบบแผนการตอบข้อสอบที่มีลักษณะเป็นกัตแมนสมบรูณ์มากน้อยเพียงใด คำนวณจากค่าสัดส่วนระหว่างผลรวมของค่าประจำข้อของคะแนนรายข้อซึ่งมีความคลาดเคลื่อนที่เกิดจากการตอบข้อสอบที่ง่ายกว่าหรือมีค่าความยากเท่ากับความสามารถของเขาได้ไม่ถูกต้องและผลรวมของค่าประจำข้อของคะแนนรายข้อซึ่งมีความคลาดเคลื่อนที่เกิดจากการตอบข้อสอบที่ยากกว่าความสามารถของเขาได้ถูกต้องกับผลรวมของค่าประจำข้อของคะแนนรายข้อทั้งหมด ซึ่งมีค่าเท่ากับค่าประจำข้อของคะแนนรายข้อที่มีความคลาดเคลื่อนสูงสุดของผู้สอบ ค่าดัชนีมีค่าอยู่ระหว่าง 0 ถึง 1 ถ้าดัชนีมีค่าเข้าใกล้ 1 แสดงว่าแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะใกล้เคียงกับกัตแมนสมบรูณ์ ถ้าดัชนีมีค่าเท่ากับ 1 แสดงว่าแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะเป็นกัตแมนสมบรูณ์ ถ้าดัชนีมีค่าเข้าใกล้ 0 แสดงว่าแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะแตกต่างไปจากกัตแมนสมบรูณ์ และถ้ามีค่าดัชนีเท่ากับ 0 แสดงว่าแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะตรงข้ามกับกัตแมนสมบรูณ์ ผลจากการตรวจสอบความตรงของดัชนีที่พัฒนาขึ้น พบว่า ดัชนีที่พัฒนาขึ้นมีความตรงในการบ่งชี้ความสอดคล้องของแบบแผนการตอบข้อสอบที่ผู้สอบทำได้กับที่มีลักษณะเป็นกัตแมนสมบรูณ์ เมื่อใช้ความสัมพันธ์ระหว่างค่าของดัชนีเอฟซีกับค่าสัดส่วนของจำนวนข้อสอบที่ผู้สอบตอบได้ถูกต้องด้วยความมั่นใจต่อจำนวนข้อสอบทั้งหมดที่ตอบได้ถูกต้องและค่าของดัชนีเอฟซีกับค่าสัดส่วนของจำนวนข้อสอบที่ผู้สอบตอบได้ไม่ถูกต้องเพราะมีความรู้ไม่เพียงพอต่อจำนวนข้อสอบทั้งหมดที่ผู้สอบตอบได้ไม่ถูกต้องเป็นเกณฑ์ในการพิจารณาความตรง รวมทั้งมีความตรงในการจำแนกแบบแผนการตอบข้อสอบที่มีความสอดคล้องกับลักษณะกัตแมนสูงและต่ำ เมื่อใช้คะแนนรวม

ความมั่นใจในการตอบข้อสอบและค่าความสามารถของผู้สอบตามแนวโลจิสติกโมเดลเป็นเกณฑ์ในการจำแนกแบบแผนการตอบข้อสอบทั้ง 2 ลักษณะ (รายชื่องานวิจัยและแหล่งที่มา แสดงในภาคผนวก ก ตารางที่ 4)

## 2.2 การใช้ดัชนีบ่งชี้ความผิดปกติของแบบแผนการตอบข้อสอบ

งานวิจัยที่เกี่ยวกับการใช้ดัชนีบ่งชี้ความผิดปกติของแบบแผนการตอบข้อสอบ มีจำนวน 2 เรื่อง เป็นการชี้แจงแผนภูมิเอส-พีและดัชนีบ่งชี้ของชาโต้ในการวินิจฉัยข้อบกพร่องทางการเรียนคณิตศาสตร์ เรื่องทศนิยมและเรื่องอัตราส่วนและร้อยละ ผลการวิจัยในการวินิจฉัยข้อบกพร่อง เรื่องทศนิยม พบว่า นักเรียนในโรงเรียนขนาดกลางมีข้อบกพร่องในแบบแผนการตอบมากที่สุด เนื้อหาที่มีข้อบกพร่องของแบบแผนการตอบมากที่สุด คือ พฤติกรรมด้านความเข้าใจและพฤติกรรมด้านการนำไปใช้ในเรื่องการลบทศนิยมที่ไม่เกิน 2 ตำแหน่ง นอกจากนี้ยังพบว่า โรงเรียนขนาดใหญ่และขนาดกลางนักเรียนมีข้อบกพร่องทางการเรียนคณิตศาสตร์แตกต่างกัน ส่วนในเรื่องอัตราส่วนและร้อยละ พบว่า นักเรียนที่เรียนแผนการเรียนช่างมีข้อบกพร่องในแบบแผนการตอบมากที่สุด รองลงมา ได้แก่ แผนการเรียนศิลปศึกษา แผนการเรียนเกษตร แผนการเรียนธุรกิจ แผนการเรียนวิทยาศาสตร์ - คณิตศาสตร์ และแผนการเรียนคหกรรม ด้านข้อบกพร่องทางการเรียน พบว่า นักเรียนที่เรียนแผนการเรียนวิทยาศาสตร์ - คณิตศาสตร์ไม่มีข้อบกพร่องทางการเรียนในเนื้อหาใด ส่วนแผนการเรียนอื่น ๆ มีข้อบกพร่องทางการเรียน เรื่องอัตราส่วนที่เท่ากันและเรื่องร้อยละ (รายชื่องานวิจัยและแหล่งที่มา แสดงในภาคผนวก ก ตารางที่ 5)

## 2.3 การพัฒนาโปรแกรมคอมพิวเตอร์เพื่อวิเคราะห์แบบแผนการตอบข้อสอบ

งานวิจัยที่เกี่ยวกับการพัฒนาโปรแกรมเพื่อวิเคราะห์แบบแผนการตอบข้อสอบ มีจำนวน 1 เรื่อง มีวัตถุประสงค์เพื่อพัฒนาโปรแกรมไมโครคอมพิวเตอร์เพื่อวิเคราะห์แบบแผนการตอบข้อสอบโดยใช้สูตรดัชนีชี้เตือนที่ปรับปรุงโดยฮาร์นิสและลินน์ มีขั้นตอนในการวิจัยและพัฒนาโปรแกรม 7 ขั้นตอน ได้แก่ กำหนดปัญหา วิเคราะห์ปัญหาและสิ่งที่ต้องการ ออกแบบโปรแกรมคอมพิวเตอร์ เลือกภาษาคอมพิวเตอร์และเขียนโปรแกรม ทดสอบและปรับปรุงแก้ไขโปรแกรม ประเมินประสิทธิภาพของโปรแกรมและจัดทำคู่มือการใช้โปรแกรม ผลจากการพัฒนาโปรแกรมทำให้ได้โปรแกรมคอมพิวเตอร์สำเร็จรูปและคู่มือการใช้ ซึ่งเป็นโปรแกรมที่สามารถวิเคราะห์คำตอบจากการสอบที่ใช้แบบทดสอบชนิดเลือกตอบไม่เกิน 5 ตัวเลือก โดยสามารถแสดงผลทั้งบนจอภาพและทางเครื่องพิมพ์ ได้แก่ แผนภูมิเอส-พี คะแนน ร้อยละของการตอบถูก ดัชนีชี้เตือนที่ปรับปรุงของนักเรียนและของข้อสอบ สัญลักษณ์ชี้เตือนที่ปรับปรุงของนักเรียนและของข้อสอบ ค่าความยาก

คำอำนาจจำแนก แผนภูมิเอส-พี ที่แบ่งข้อสอบตามกลุ่มเนื้อหาย่อย ตารางจำแนกและแจกแจงความถี่ของนักเรียนและของข้อสอบและผลสรุปค่าสถิติจากการวิเคราะห์ ได้แก่ สัมประสิทธิ์ความแตกต่าง ค่าความเบ้และความโด่งของคะแนน ความเที่ยงของแบบทดสอบและความคลาดเคลื่อนมาตรฐานของการวัด ผลการประเมินโปรแกรม พบว่า โปรแกรมสามารถทำงานตามคำสั่งได้อย่างถูกต้องและมีประสิทธิภาพ (รายชื่องานวิจัยและแหล่งที่มา แสดงในภาคผนวก ก ตารางที่ 6)

### 3. แนวคิดการประเมินกฎในการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อน

งานวิจัยที่เกี่ยวข้องกับแนวคิดการประเมินกฎในการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อน มีจำนวน 2 เรื่อง เป็นเนื้อหาในวิชาคณิตศาสตร์ เรื่องการบวกเลขจำนวนเต็มลบและเรื่องการทำงานวิจัยเรื่องแรกมีกระบวนการวิจัย 3 ขั้นตอน ได้แก่ สํารวจแบบการคิดทั้งหมดที่เป็นไปได้ สร้างข้อสอบจากรูปแบบข้อสอบทั้งหมดที่เป็นไปได้ในแบบแผนการตอบทั้งหมดซึ่งสามารถย้อนรอยกระบวนการคิดของผู้สอบแต่ละคนและวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนใน 2 ขั้นตอน ได้แก่ วินิจฉัยเชิงสำรวจและวินิจฉัยเพื่อยืนยันแบบการคิด ผลการวิจัย พบว่า วิธีที่พัฒนาขึ้นมีความคงที่ในการวินิจฉัยและมีความตรงเชิงเกณฑ์สัมพันธ์เมื่อใช้การวินิจฉัยของครูเป็นเกณฑ์ นอกจากนี้ยังพบว่า ผลการวินิจฉัยจากวิธีที่พัฒนาขึ้นมีความสอดคล้องกับผลการวินิจฉัยจากวิธีการของทาซุโอะกะและลดข้อจำกัดของวิธีการของทาซุโอะกะเกี่ยวกับการจำแนกมโนทัศน์ที่คลาดเคลื่อนซึ่งมีแบบแผนคะแนนการตอบเหมือนกันหรือมีแบบแผนคะแนนการตอบเป็นศูนย์หมดทุกข้อ สำหรับโปรแกรมคอมพิวเตอร์ที่พัฒนาขึ้นพบว่าใช้งานได้ตามวัตถุประสงค์ที่กำหนด

งานวิจัยเรื่องที่ 2 ใช้เครื่องมือในการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อน จำนวน 4 ฉบับ ได้แก่ แบบวัดเพื่อสำรวจข้อบกพร่อง แบบวัดมโนทัศน์ที่คลาดเคลื่อน แบบวัดเพื่อย้อนรอยการคิด และแบบวัดพื้นฐานความเข้าใจ การวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนมีขั้นตอนการดำเนินการ ดังนี้ สำรวจข้อบกพร่องเพื่อสร้างเป็นแบบการคิดที่เป็นไปได้ทั้งหมด เปรียบเทียบชุดคำตอบของนักเรียนกับชุดคำตอบของแบบการคิดในการทดสอบด้วยแบบวัดมโนทัศน์ที่คลาดเคลื่อนและแบบวัดเพื่อย้อนรอยการคิด กรณีที่ไม่สามารถวินิจฉัยได้ สำหรับนักเรียนในกลุ่มเรียนอ่อนและกลุ่มที่ตอบแบบเดาสุ่มใช้แบบวัดพื้นฐานความเข้าใจ ผลการวิจัย พบว่า มีแบบการคิดทั้งหมด 15 แบบการคิด นักเรียนมีมโนทัศน์ที่คลาดเคลื่อนจากการใช้แบบการคิดที่ 1 คือ ไม่เข้าใจการกระจายเศษมีจำนวนมากที่สุด การวินิจฉัยนักเรียนในกลุ่มเดาสุ่ม พบว่า นักเรียนส่วนใหญ่มีความบกพร่องในการหารเหลือเศษ กรณีตัวตั้ง 1 หลักที่มีค่าน้อยกว่าตัวหาร การหารเหลือเศษกรณีตัวตั้ง 2 หลัก ผลหาร 2 หลัก มีการกระจายเศษ ผลการประเมินวิธีการวินิจฉัยที่พัฒนาขึ้น พบว่า เป็นวิธีการวินิจฉัยที่สามารถ

นำไปใช้ได้และมีความตรงเชิงเกณฑ์สัมพัทธ์เมื่อใช้ผลการวินิจฉัยของครูเป็นเกณฑ์ (รายชื่อ งานวิจัยและแหล่งที่มา แสดงในภาคผนวก ก ตารางที่ 7)

#### 4. การคำนึงถึงสภาพจริงในการแก้โจทย์ปัญหา

งานวิจัยที่เกี่ยวกับการใช้การคำนึงถึงสภาพจริงในการแก้โจทย์ปัญหามีจำนวน 1 เรื่อง เป็นเนื้อหาในวิชาคณิตศาสตร์ ขั้นตอนในการวินิจฉัยจำแนกเป็น 3 มิติ ได้แก่ มิติความเข้าใจการแก้โจทย์ปัญหา มิติการคิดคำนวณและมิติการให้คำตอบที่ใช้ได้กับสภาพที่เป็นจริง เครื่องมือที่ใช้ ได้แก่ แบบทดสอบแบบเลือกตอบที่วัดความสามารถของนักเรียนทั้ง 3 มิติ จำนวน 18 ข้อ ผลการวิจัย พบว่า วิธีการที่พัฒนาขึ้นมีความเที่ยงแบบแบบทดสอบคู่ขนาน โดยให้ผลการวินิจฉัย ในการจำแนกความสามารถของนักเรียนได้สอดคล้องกันร้อยละ 81.86 มีความตรงเชิงเกณฑ์สัมพัทธ์โดยให้ผลการวินิจฉัยตรงกับวิธีการคิดออกเสียงร้อยละ 86.87 (รายชื่อ งานวิจัยและแหล่งที่มา แสดงในภาคผนวก ก ตารางที่ 8)

#### 5. การวินิจฉัยและแก้ไขข้อบกพร่องโดยนักเรียน

งานวิจัยที่เกี่ยวกับการวินิจฉัยและแก้ไขข้อบกพร่องโดยนักเรียน มีจำนวน 1 เรื่อง เป็นการสอนวิธีการวินิจฉัยและแก้ไขงานเขียนเรียงความที่ไม่ชัดเจน การดำเนินการสอนวิธีการวินิจฉัยและแก้ไขข้อบกพร่องใช้เวลา 2 คาบต่อสัปดาห์ เป็นเวลา 8 สัปดาห์ เครื่องมือที่ใช้ ได้แก่ แบบทดสอบวัดความสามารถในการเขียนเรียงความภาษาไทยซึ่งเป็นแบบปรนัยและแบบอัตนัย ผลการวิจัย พบว่า นักเรียนที่ได้รับการสอนวิธีการวินิจฉัยและแก้ไขงานเขียนเรียงความที่ไม่ชัดเจนมีความสามารถทางการเขียนเรียงความภาษาไทยผ่านเกณฑ์ร้อยละ 70 ของคะแนนสอบทั้งฉบับ โดยมีความสามารถสูงกว่าก่อนได้รับการสอนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (รายชื่อ งานวิจัยและแหล่งที่มา แสดงในภาคผนวก ก ตารางที่ 9)

#### 6. เทคนิคการคิดออกเสียง

งานวิจัยที่เกี่ยวกับการใช้เทคนิคการคิดออกเสียงในการวินิจฉัยมีจำนวน 1 เรื่อง เป็นการใช้เทคนิคการคิดออกเสียงเป็นเครื่องมือในการวินิจฉัยความสามารถในการแก้ปัญหาโจทย์คณิตศาสตร์เพื่อจัดซ่อมเสริมและเพื่อศึกษากระบวนการคิดและวิธีแก้ปัญหาโจทย์คณิตศาสตร์ วิเคราะห์วิธีคิดของนักเรียนที่ได้คะแนนผ่านเกณฑ์ที่กำหนด วินิจฉัยข้อบกพร่องในวิธีคิดและศึกษาผลของการสอนซ่อมเสริมตามวิธีคิดที่บกพร่องที่ค้นพบ ผลการวิจัย พบว่า วิธีคิดของนักเรียนในการนำไปใช้แก้ปัญหาโจทย์คณิตศาสตร์ มี 6 วิธีการ ได้แก่ วิธีการกำหนดตัวแปรไม่ทราบค่า วิธีการสร้างภาพ วิธีการสร้างตาราง วิธีการให้เหตุผล วิธีการทำย้อนกลับและวิธีการคาดคะเนและ

ตรวจสอบ วิธีที่นักเรียนกลุ่มที่ได้คะแนนผ่านเกณฑ์นำมาใช้ในการแก้ปัญหาโจทย์คณิตศาสตร์ในแต่ละข้อมากที่สุด ได้แก่ ข้อที่ 1 วิธีการคาดคะเนและตรวจสอบ ข้อที่ 2 วิธีการให้เหตุผล ข้อที่ 3 วิธีการสร้างภาพ ข้อที่ 4 วิธีการสร้างตารางและข้อที่ 5 วิธีการสร้างภาพ/วิธีการทำย้อนกลับ ด้านข้อบกพร่องในวิธีคิด พบว่า ก่อนการสอนซ่อมเสริมนักเรียนมีข้อบกพร่องในวิธีคิด โดยมีสาเหตุจากการไม่รู้จักรูปวิธีคิดมากที่สุด หลังการสอนซ่อมเสริม พบว่า นักเรียนมีข้อบกพร่องในการคิดโดยมีสาเหตุจากการใช้วิธีคิดไม่ถูกวิธีมากที่สุด หลังการสอนซ่อมเสริมพบว่านักเรียนยังคงใช้วิธีคิดวิธีเดิมมากที่สุด ด้านสาเหตุที่ทำให้คะแนนเพิ่มขึ้น พบว่า การเปลี่ยนวิธีคิดทำให้นักเรียนมีคะแนนเพิ่มขึ้นมากที่สุด ผลของการสอนซ่อมเสริมทำให้นักเรียนมีคะแนนเพิ่มขึ้น ร้อยละ 40.0 มีคะแนนลดลง ร้อยละ 12.0 และมีคะแนนเท่าเดิม ร้อยละ 48.0 (รายละเอียดงานวิจัยและแหล่งที่มา ก แสดงในภาคผนวก ตารางที่ 10)

จากการศึกษางานวิจัยที่เกี่ยวกับการวิจัยทางการศึกษาในประเทศไทยดังกล่าวข้างต้น จะเห็นได้ว่า การวิจัยทางการศึกษายังจำกัดอยู่เฉพาะการศึกษาในลักษณะวิทยานิพนธ์ซึ่งเน้นที่การสร้างหรือพัฒนาแบบทดสอบเชิงวินิจฉัยมากกว่าการใช้รูปแบบและวิธีการอื่น ๆ มีการใช้ข้อมูลเชิงวินิจฉัยในการแก้ไขปรับปรุงนักเรียน การนำแนวคิดทฤษฎีทางพุทธิปัญญามาใช้ในการวินิจฉัยและการใช้ทฤษฎีการทดสอบแนวใหม่ในการวิเคราะห์ข้อสอบจำนวนน้อยมาก ตลอดจนยังไม่มีการนำคอมพิวเตอร์มาใช้ในการทดสอบ

## ตอนที่ 2 การประเมินเชิงวินิจฉัยโดยใช้โมเดลลำดับขั้นของคุณลักษณะ

**แนวคิดที่เกี่ยวข้องกับโมเดลลำดับขั้นของคุณลักษณะ : การพัฒนา Rule space model**

Rule space model เป็นโมเดลที่พัฒนาขึ้นโดย Kikumi K. Tatsuoka ซึ่งเป็นศาสตราจารย์ที่มีชื่อเสียงด้านการวัดผลทางการศึกษา (Educational Measurement) อยู่ที่ Teachers College, Columbia University มีความเชี่ยวชาญด้านการวิจัยเกี่ยวกับการวัดผล การประเมินผลและสถิติ โดยเฉพาะด้านทฤษฎีการทดสอบทางการศึกษาและการวินิจฉัยทางพุทธิปัญญา (Educational testing theory and cognitive diagnosis) การพัฒนา Rule space model มีรายละเอียดดังนี้

### 1. การพัฒนาวิธีวินิจฉัยความผิดปกติของแบบแผนการตอบข้อสอบก่อนการพัฒนา Rule space model

Tatsuoka ได้ดำเนินการศึกษาวิธีการวินิจฉัยความผิดปกติของแบบแผนการตอบข้อสอบซึ่งเป็นจุดเริ่มต้นแนวคิดในการพัฒนา Rule space model ตั้งแต่ปี ค.ศ.1980 ในขณะที่เป็นผู้ช่วยศาสตราจารย์ด้านจิตวิทยาการศึกษา (Educational Psychology) อยู่ที่ University of



Illinois โดยในระยะแรก ระหว่างปี ค.ศ. 1980 ถึง ปี ค.ศ. 1983 ได้ร่วมกับ Maurice M. Tatsuoka ผู้ช่วยศาสตราจารย์ด้านจิตวิทยาและจิตวิทยาการศึกษา (Psychology and educational psychology) อยู่ที่ University of Illinois พัฒนาดัชนีเพื่อบ่งชี้ระดับความผิดปกติของแบบแผนการตอบข้อสอบ จำนวน 2 ดัชนี ดัชนีแรก ได้แก่ ดัชนีการคล้อยตามกลุ่ม (Norm Conformity Index : NCI) พัฒนารุ่นเพื่อบ่งชี้ถึงระดับความคงเส้นคงวาระหว่างแบบแผนการตอบข้อสอบรายบุคคล (response pattern of an Individual) กับลำดับความยากง่ายของข้อสอบ (difficulty ordering of the items) โดยการเปรียบเทียบผลรวมของคะแนนรายข้อที่สอดคล้องกับลักษณะกัตแมนสมบูรณ์และผลรวมของคะแนนรายข้อที่มีลักษณะไม่สอดคล้องกับลักษณะกัตแมนสมบูรณ์ ดัชนีที่สอง ได้แก่ ดัชนีความสอดคล้องรายบุคคล (Individual Consistency Index : ICI) พัฒนารุ่นโดยใช้แนวคิดเดียวกับดัชนีการคล้อยตามกลุ่ม แต่ปรับเปลี่ยนวิธีการเรียงข้อสอบตามลำดับความยากง่ายของข้อสอบมาเป็นการเรียงข้อสอบตามลำดับจากสถานะความรู้รายบุคคล (individual student state of knowledge) ซึ่งได้จากการวัดการเปลี่ยนแปลงโดยการวัดซ้ำ (repeated measures) เพื่อแก้ไขข้อบกพร่องจากการเรียงข้อสอบตามลำดับความยากง่ายของข้อสอบซึ่งขึ้นอยู่กับความสามารถของกลุ่มผู้สอบ (Harnisch and Linn, 1981; Tatsuoka and Tatsuoka, 1982; Tatsuoka and Tatsuoka, 1983; พนิชา สังข์เพชร, 2535)

นอกจากนี้ ในระหว่างปี ค.ศ. 1980 ถึง ค.ศ. 1982 Tatsuoka, Baillie และ Yamamoto ยังได้พัฒนาโปรแกรม "SIGNBUG" และ โปรแกรม "SIGNBUG 2" เพื่อใช้วินิจฉัยจำนวนการใช้กฎที่ผิดเนื่องจากการมีมโนทัศน์ที่คลาดเคลื่อนในการใช้เครื่องหมายวงเล็บเพื่อแก้ปัญหาเกี่ยวกับเลขจำนวนขึ้น แต่เป็นเพียงการศึกษาเพื่อสำรวจลักษณะของความคลาดเคลื่อนในบริบทของการวัดทางจิต เช่น การเปลี่ยนแปลงประเภทของความคลาดเคลื่อนหรือความสม่ำเสมอของมโนทัศน์ที่คลาดเคลื่อน การนำโปรแกรมไปใช้ยังคงมีข้อจำกัดในเรื่องค่าใช้จ่าย (expensive) รวมทั้งสามารถนำไปใช้ได้จำกัดเฉพาะเนื้อหาทางคณิตศาสตร์เท่านั้น (Tatsuoka and Tatsuoka, 1982; Tatsuoka, 1983; Tatsuoka and Tatsuoka, 1983)

## 2. การพัฒนา Rule space model โดยใช้แนวคิดการประเมินกฎ

แม้ว่าจะมีการพัฒนาวิธีวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนมาอย่างต่อเนื่อง โดยเฉพาะการใช้แนวคิดการประเมินกฎ แต่วิธีการที่ผ่านมายังคงมีข้อจำกัดที่ไม่สามารถวินิจฉัยแหล่งของมโนทัศน์ที่คลาดเคลื่อนซึ่งทำให้สารสนเทศที่ได้จากการวินิจฉัยไม่สามารถใช้ประโยชน์เพื่อการปรับปรุงแก้ไขนักเรียนและการเรียนการสอนของครูได้อย่างแท้จริง ด้วยเหตุนี้ Tatsuoka (1983) จึงได้พัฒนาโมเดลการประเมินเชิงวินิจฉัยจากแนวคิดวิธีการประเมินกฎ (rule-assessment method)

และวิธีดัชนีบุคคล (personal-index approach) ขึ้น โดยประยุกต์ใช้ Rule space model ซึ่งเป็นโมเดลที่พัฒนาเพื่อใช้ในบริบทของปัญหาทางวิศวกรรมและวิทยาศาสตร์กายภาพ ในการเปรียบเทียบความใกล้เคียงของแบบแผนการตอบข้อสอบกับแบบการคิดผิดแบบต่าง ๆ โดยใช้จุดในเวกเตอร์สเปซสองมิติระหว่างดัชนีบ่งชี้ความผิดปกติของแบบแผนการตอบข้อสอบกับระดับความสามารถของผู้สอบ ซึ่งจากการศึกษาในระยะแรกพบว่าการใช้ดัชนีบ่งชี้ความผิดปกติจากแบบแผนการสอบที่อิงทฤษฎีการทดสอบแบบดั้งเดิม (classical test theory) และคะแนนรวมไม่สามารถแยกตำแหน่งของจุดในเวกเตอร์สเปซได้อย่างชัดเจน Tasuoka จึงได้พัฒนาดัชนีที่มีพื้นฐานแนวคิดจากทฤษฎีการตอบสนองของข้อสอบมาใช้ โดยการใช้ค่าความสามารถของผู้สอบ ( $\theta$ ) แทนคะแนนรวมและค่าความผิดปกติของแบบแผนการตอบข้อสอบ ( $\zeta$ ) แทนดัชนีบ่งชี้ความผิดปกติของแบบแผนการตอบข้อสอบที่อิงทฤษฎีการทดสอบแบบดั้งเดิม ซึ่งต่อมามีการพัฒนาดัชนีชี้เตือนมาตรฐานหมายเลข 4 (standardized extended caution index IV : SECI 4) ขึ้น ดัชนีบ่งชี้ความผิดปกติของแบบแผนการตอบข้อสอบกับระดับความสามารถของผู้สอบจะนำไปใช้เพื่อกำหนดจุดของแบบแผนการตอบข้อสอบของนักเรียนแต่ละคน (student point) พร้อมกับจุดของแบบแผนการคิดแต่ละแบบ (rule point) หลังจากนั้นจึงคำนวณค่าความใกล้เคียงระหว่างจุดของแบบแผนการตอบข้อสอบของนักเรียนแต่ละคนกับจุดของแบบแผนการคิดแต่ละแบบ โดยคำนวณระยะทางน้อยที่สุดตามกฎระยะทางกำลังสองน้อยที่สุด (minimum- $D^2$  rule ;  $D^2$  = Mahalanobis generalized square-distance) และใช้การตัดสินใจแบบเบย์ (Bayes decision rule for minimum error) สำหรับจุดของผู้สอบที่มีระยะทางระหว่างแบบการคิดที่ผิดสองแบบเท่ากัน (Tasuoka, 1984; ศิริเดช สุชีวะ, 2538)

สำหรับขั้นตอนในการวินิจฉัยคนที่คลาดเคลื่อนโดยใช้ Rule space model ศิริเดช สุชีวะ (2538) นำเสนอไว้ สรุปได้ดังนี้

- 1) สํารวจแบบการคิดที่ผิดที่เป็นไปได้ทั้งหมดในเนื้อหา
- 2) สร้างข้อสอบให้ครอบคลุมวัตถุประสงค์เชิงพฤติกรรม รวบรวมเป็นแบบทดสอบปลายเปิดชนิดเติมคำ
- 3) คํานวณหาคําตอบโดยใช้แบบการคิดที่ผิดที่ละแบบ
- 4) ตรวจสอบให้คะแนน 0-1 ซึ่งจะได้แบบแผนคะแนนการตอบประจำแบบการคิด
- 5) นำแบบทดสอบไปให้กลุ่มตัวอย่างที่จะวินิจฉัยทำและตรวจสอบให้คะแนนแบบ 0-1 ซึ่งจะได้แบบแผนคะแนนการตอบจากแต่ละแบบแผนการตอบข้อสอบ
- 6) คํานวณค่าดัชนีที่บ่งชี้ความใกล้เคียงระหว่างแบบแผนการตอบข้อสอบกับแบบแผนการคิดที่ผิดแบบต่าง ๆ
- 7) ระบุแบบแผนการคิดที่ผิดของผู้สอบแต่ละคน

### 3. การพัฒนา Rule space model โดยใช้แนวคิดการประเมินเชิงวินิจฉัยด้านพุทธิปัญญา

เมื่อพิจารณาการนำแนวคิดในการพัฒนา Rule space model ของ Tatsuoka มาใช้ในการวินิจฉัย พบว่า Tatsuoka มีการประยุกต์แนวคิดจิตวิทยาด้านพุทธิปัญญามาใช้ในการวินิจฉัย ร่วมกับทฤษฎีการตอบสนองข้อสอบตั้งแต่ปี ค.ศ. 1990 จะเห็นได้จากการเสนอแนวคิดเรื่อง Toward an Integration of item response theory and cognitive analysis ในเอกสาร Diagnostic monitoring of skill and knowledge acquisition (Tatsuoka, 2004) ประกอบกับ Tatsuoka (2002) ได้แสดงข้อสังเกตไว้ในหน้าข่าวของ website ของ Columbia University เรื่อง “The rule-space model for educational measurement” ว่าการใช้วิธีการเชิงกำหนด (deterministic approach) เพื่อค้นหาในทัศนที่คลาดเคลื่อนของนักเรียนจากแบบการคิดที่ผิด (students error or bugs) ไม่สามารถใช้ได้ดี เนื่องจากบางครั้งนักเรียนเปลี่ยนแบบการคิดในการเลือกคำตอบโดยไม่มีเหตุผลและไม่สามารถบอกได้ว่าทำไมจึงเลือกคำตอบนั้น ทำให้ไม่สามารถวินิจฉัยสาเหตุของการเลือกคำตอบจากแบบการคิดที่ผิดของนักเรียนได้

Tatsuoka ได้ปรับเปลี่ยนแนวคิดในการวินิจฉัยในทัศนที่คลาดเคลื่อนจากแบบการคิดที่ผิดมาเป็นการวินิจฉัยที่ให้สารสนเทศเกี่ยวกับจุดเด่นจุดด้อยซึ่งเป็นสถานะความรอบรู้จากการพิจารณาว่าการทำข้อสอบได้ถูกต้องต้องใช้ความรู้และทักษะกระบวนการทางปัญญาใดบ้าง Rule space model ที่พัฒนาขึ้นใหม่จึงมีคุณลักษณะเป็นโมเดลทางคณิตศาสตร์เพื่อจำแนกแบบแผนการตอบข้อสอบของผู้สอบไปสู่แบบแผนความรอบรู้ (attribute-mastery patterns) จากการใช้ทักษะทางด้านพุทธิปัญญาที่แตกต่างกัน โดยมีข้อตกลงเบื้องต้นว่าโมเดลที่ได้จากข้อสอบอาจจะบรรยายได้ด้วยทักษะทางพุทธิปัญญาเฉพาะซึ่งสามารถจำแนกวิธีการ ทักษะหรือกระบวนการของผู้สอบได้และเรียกทักษะทางพุทธิปัญญาเฉพาะนั้นว่า คุณลักษณะ (attribute) คุณลักษณะจึงเป็นความรู้หรือทักษะการคิดที่มีอยู่และใช้เพื่อการทำข้อสอบให้ถูกต้อง เปรียบเสมือนสถานะความรู้ของผู้สอบ (knowledge states or latent knowledge states) คุณลักษณะหรือสถานะความรู้ไม่สามารถสังเกตได้โดยตรงแต่แสดงให้เห็นได้จากการตอบข้อสอบในแบบทดสอบ การใช้ rule space model ในการประเมินเพื่อกำหนดความใกล้เคียงของสถานะความรู้ของผู้สอบกับแบบแผนของสถานะความรู้ จึงต้องใช้การพิจารณาจากแบบแผนการตอบข้อสอบของผู้สอบ โดยกำหนดการให้คะแนนแบบ 0-1 ให้คะแนน 1 สำหรับข้อสอบที่ตอบถูกซึ่งหมายถึงมีความรอบรู้ในคุณลักษณะหรือสถานะความรู้นั้น และให้คะแนน 0 สำหรับข้อสอบที่ตอบผิดซึ่งหมายถึงความไม่รอบรู้ในคุณลักษณะหรือสถานะความรู้ นั้น นอกจากนี้ยังได้นำหลักการของ Boolean Algebra มาใช้ในการกำหนดเมทริกซ์ของแบบแผนการตอบข้อสอบเพื่อใช้ในการคำนวณด้วย (Gierl, Leighton and Hunka, 2000; Tatsuoka, 2004)

Rule space model สร้างขึ้นโดยการรวบรวมและจัดอันดับคุณลักษณะที่ต้องใช้ในการตอบข้อสอบแล้วใช้วิธีการทางสถิติในการจำแนกแบบแผนการตอบข้อสอบไปสู่แบบแผนความรู้ซึ่งเป็นแบบแผนความสำเร็จทางพุทธิปัญญาแบบแผนเดียว (unique cognitive blueprint) การใช้ Rule space model ในการประเมินเชิงวินิจฉัยจึงมีวัตถุประสงค์สำคัญ 2 ประการ คือ เพื่อระบุลักษณะสำคัญของตัวแปร โดยการกำหนดคุณลักษณะและเพื่อจำแนกแบบแผนทางสถิติ โดยใช้ two-dimensional cartesian coordinate system ที่เรียกว่า rule space ซึ่งแบ่งเป็นแกนระดับความสามารถของผู้สอบ (Theta :  $\theta$  or ability) และแบบแผนการตอบที่ผิดปกติ (Zeta:  $\zeta$  or response unusualness) สรุปขั้นตอนการใช้ Rule space model ในการประเมินเชิงวินิจฉัยด้านพุทธิปัญญาได้ดังนี้ (Gierl, Leighton and Hunka, 2000; Lohman, 2002; Yan, Almond and Mislevy, 2003; Tatsuoka, 2004)

- 1) กำหนดคุณลักษณะ โดยพิจารณาจากชุดของทักษะเฉพาะหรือความสามารถที่ต้องใช้ในการตอบข้อสอบให้ถูกต้อง
- 2) กำหนด matrix (Q) โดยให้แถวแทนข้อสอบ (j) และคอลัมน์แทนคุณลักษณะ (k)
- 3) กำหนดให้ข้อที่ j คุณลักษณะที่ k ( $Q_{jk}$ ) เป็น 1 ถ้าข้อที่ j โดยใช้ คุณลักษณะที่ k ได้ถูกต้องและเป็น 0 เมื่อข้อที่ j โดยใช้คุณลักษณะที่ k ได้ไม่ถูกต้องซึ่งจะทำให้ได้เมทริกซ์ที่แสดงถึงแบบแผนของสถานะความรู้จากแบบแผนการตอบข้อสอบตามอุดมคติ (ideal item response patterns)
- 4) คำนวณค่าความน่าจะเป็นของแบบแผนของสถานะความรู้แต่ละแบบ โดยใช้หลัก Boolean algebra ซึ่งสามารถใช้โปรแกรม BUGSHELL ในการคำนวณได้
- 5) ตรวจสอบให้คะแนนแบบแผนการตอบข้อสอบของผู้สอบ แบบ 0-1 และกำหนด matrix (Q) ที่แสดงถึงแบบแผนของสถานะความรู้จากแบบแผนการตอบข้อสอบที่สังเกตได้ (observe item response patterns)
- 6) เปรียบเทียบความใกล้เคียงระหว่างแบบแผนการตอบข้อสอบตามอุดมคติกับแบบแผนการตอบข้อสอบที่สังเกตได้
- 7) จัดกลุ่มผู้สอบตามสถานะความรู้ต่าง ๆ โดยพิจารณาจากสถานะความรู้ที่มากที่สุดซึ่งแสดงให้เห็นจากแบบแผนการตอบข้อสอบของนักเรียน โดยประยุกต์ใช้ Bayesian decision rule

#### 4. ตัวอย่างการนำ Rule space model ไปใช้ในการประเมินเชิงวินิจฉัย

Tatsuoka, Corter และ Tatsuoka (2004) ได้ดำเนินการวิจัยเรื่องรูปแบบการวินิจฉัยเนื้อหาและกระบวนการทางคณิตศาสตร์ในการสอบ The third international math and science student -

revised (TIMSS-R) จากกลุ่มตัวอย่าง จำนวน 20 ประเทศ เพื่อศึกษาคุณลักษณะด้านความรู้ ทักษะ และกระบวนการในการทำแบบทดสอบคณิตศาสตร์ จากการสอบ TIMSS-R สำหรับประเทศที่แตกต่างกันด้านวัฒนธรรมและสิ่งแวดล้อมทางการศึกษา การดำเนินการวิจัยแสดงให้เห็นการประยุกต์ใช้ Rule space model ในการประเมินเชิงวิจนิจฉัย ดังนี้

### **ขอบเขตการศึกษา**

เป็นการวิจัยความรอบรู้ที่แสดงให้เห็นจากคุณลักษณะที่เป็นความรู้และทักษะย่อยทางพุทธิปัญญาที่ใช้ในการทำข้อสอบคณิตศาสตร์จากการสอบ TIMSS-R ของนักเรียนเกรด 8 (eighth-grade) เลือกประเทศที่ใช้ในการเปรียบเทียบโดยใช้ Rule space model จำนวน 20 ประเทศ ตามความแตกต่างกันด้านวัฒนธรรมและผลสัมฤทธิ์ ดังนี้ กลุ่มที่ 1 ประเทศที่มีผลสัมฤทธิ์สูงสุด 6 ประเทศ ได้แก่ สิงคโปร์, เกาหลี, ฮองกง, ญี่ปุ่น, สาธารณรัฐเชค, เบลเยียม-ฟรานเดิส กลุ่มที่ 2 สหรัฐและประเทศที่ใช้ภาษาอังกฤษในกลุ่มคนที่แตกต่างกันเหมือนสหรัฐและมีผลสัมฤทธิ์คณิตศาสตร์สูงกว่าสหรัฐ ได้แก่ แคนาดา และออสเตรเลีย กลุ่มที่ 3 ประเทศที่ใช้ภาษาสเปน ได้แก่ ชิลี กลุ่มที่ 4 ประเทศในทวีปยุโรป ได้แก่ รัสเซีย, อิตาลี, ฟินแลนด์, เนเธอร์แลนด์และอังกฤษ กลุ่มที่ 5 ประเทศในตะวันออกกลางหรืออนึ่งือศาสนาอิสลาม ได้แก่ อิสราเอล, ตุรกี, จอร์แดน, อินโดนีเซียและฟิลิปปินส์

### **จำแนกคุณลักษณะ**

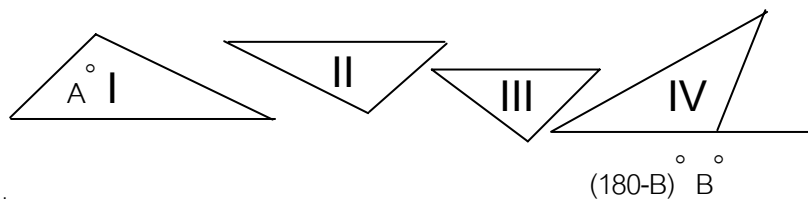
การจำแนกคุณลักษณะที่ต้องใช้ในการตอบข้อสอบให้ถูกต้องในแต่ละข้อวิเคราะห์จากการเขียนต้นร่างของนักเรียน (student protocol) โดยคณะผู้เชี่ยวชาญ ประกอบด้วยศาสตราจารย์หรือนักศึกษาในระดับบัณฑิตศึกษาด้านการวัดซึ่งทั้งหมดเป็นผู้ที่มีประสบการณ์ในการสอนคณิตศาสตร์ในระดับมัธยมศึกษาหรือวิชาสถิติเบื้องต้นในระดับวิทยาลัย หลังจากนั้นจึงนำคุณลักษณะที่ได้ไปสัมภาษณ์ครูคณิตศาสตร์เพื่อตรวจสอบความตรงและประโยชน์ในการศึกษาและนำผลมาปรับปรุงคุณลักษณะ คุณลักษณะที่ได้จำแนกเป็น 3 กลุ่ม ได้แก่ ความรู้ด้านเนื้อหา (content knowledge) กระบวนการทางพุทธิปัญญา (cognitive process) และทักษะหรือประเภทของข้อคำถาม (skill or item-type) รายละเอียด ดังแสดงในตารางที่ 1

ตารางที่ 1 ความรู้ ทักษะและกระบวนการที่ใช้ในการทำแบบทดสอบ TIMSS-R สำหรับนักเรียนเกรด 8

รหัส	คุณลักษณะ
คุณลักษณะด้านเนื้อหา	
C1	มโนทัศน์เบื้องต้นและปฏิบัติการเกี่ยวกับจำนวนและจำนวนเต็ม
C2	มโนทัศน์เบื้องต้นและปฏิบัติการเกี่ยวกับเศษส่วนและทศนิยม
C3	มโนทัศน์เบื้องต้นและปฏิบัติการเกี่ยวกับพีชคณิตเบื้องต้น
C4	มโนทัศน์เบื้องต้นและปฏิบัติการเกี่ยวกับเรขาคณิต 2 มิติ
C5	ข้อมูล, ความน่าจะเป็นและสถิติพื้นฐาน
C6	การวัดหรือการประมาณค่า, ความยาว, เวลา, มุม, อุณหภูมิ ฯลฯ
คุณลักษณะด้านกระบวนการ	
P1	การแปลงหรือการสร้างสูตรและการแก้ปัญหา
P2	การประยุกต์ใช้ความรู้ในการคำนวณทางเลขคณิตและเรขาคณิต
P3	การประยุกต์ใช้ความรู้ในการตัดสินใจทางเลขคณิตและเรขาคณิต
P4	การประยุกต์ใช้กฎในพีชคณิต
P5	เหตุผลเชิงตรรกะ ได้แก่ การรวมเหตุผล การคิดเชิงอนุมาน, ถ้า-แล้ว, ความจำเป็นและความเพียงพอ, การสรุปอ้างอิง
P6	การค้นหาปัญหา, การคิดเชิงวิเคราะห์, การสร้างปัญหาขึ้นมาใหม่, การคิดเชิงอุปนัย
P7	การสร้าง การดูและการอ่านภาพและกราฟ
P8	การประยุกต์ใช้และประเมินความถูกต้องทางคณิตศาสตร์
P9	การจัดการและการดำเนินการเกี่ยวกับข้อมูล
P10	การอ่านเชิงปริมาณและการอ่านเชิงตรรกะ
คุณลักษณะด้านทักษะหรือประเภทของข้อคำถาม	
S1	Unit conversation
S2	การประยุกต์ใช้คุณสมบัติและความสำคัญของจำนวน, ความหมายของจำนวน, เส้นจำนวน
S3	การใช้รูปภาพ, ตาราง, แผนภาพและกราฟ
S4	การคาดคะเนและการประมาณค่า
S5	การประเมิน, การพิสูจน์และการตรวจสอบทางเลือก
S6	แบบแผนและความสัมพันธ์ (ทักษะการคิดเชิงอนุมาน)
S7	การใช้เหตุผลอย่างสมเหตุสมผล
S8	การแก้ปัญหาใหม่หรือปัญหาที่ไม่คุ้นเคย
S9	การเปรียบเทียบข้อมูลในฐานข้อมูล 2 ฐานหรือมากกว่า
S10	คำถามปลายเปิดที่ไม่ให้คำตอบ
S11	ความเข้าใจคำที่ใช้ในการตั้งคำถาม

### การพัฒนา Q Matrix

เมื่อได้ชุดของคุณลักษณะในข้อ 2) คำถามจำนวน 163 ข้อ ที่ใช้ในแบบทดสอบย่อย วิชาคณิตศาสตร์ 8 รูปแบบจะถูกนำมากำหนดคุณลักษณะที่ต้องใช้ในการตอบคำถามแต่ละข้อ โดยผู้เชี่ยวชาญจำนวน 3 คน ประกอบด้วยนักวิจัยที่มีวุฒิการศึกษาระดับปริญญาเอกซึ่งมีประสบการณ์ในการสอนวิชาสถิติและการวัดผลในระดับวิทยาลัย จำนวน 2 คน และนักศึกษา ระดับบัณฑิตศึกษาที่มีพื้นฐานด้านการวัด จำนวน 1 คน ในขั้นแรกผู้เชี่ยวชาญแต่ละคนจะกำหนด คุณลักษณะที่ต้องใช้ในการตอบข้อสอบโดยอิสระหลังจากนั้นจึงมาพบและอภิปรายในข้อที่มีความคิดเห็นไม่สอดคล้องกัน โดยใช้การทบทวนต้นร่างของนักเรียน (student protocols) ประกอบการ พิจารณา จนกระทั่งได้ฉันทามติจึงสร้าง Q Matrix และวิเคราะห์ค่าความยากของข้อสอบ โดยใช้การ วิเคราะห์ถดถอยเชิงเส้นซึ่งมีคุณลักษณะเป็นตัวทำนาย รวมทั้งประมาณค่าความน่าจะเป็นในการ ระบุคุณลักษณะของนักเรียนแต่ละคนและใช้สถิติบรรยาย โดยใช้เมตริกซ์ความสัมพันธ์ของระดับ ความน่าจะเป็นในการระบุเพื่อคำนวณและตัดคุณลักษณะที่มีค่าสถิติไม่เหมาะสมออกเพื่อใช้ในการ คำนวณ Rule space model ในขั้นสุดท้าย เช่น ในแผนภาพที่ 1 แสดงให้เห็นว่าคุณลักษณะที่ใช้ใน การตอบข้อสอบ ประกอบด้วย C4, S3, S5, P3, P5, P7 และ P9 มีค่าสัดส่วนการตอบถูกเท่ากับ .1025 แสดงว่าเป็นปัญหาที่ยากมาก เนื่องจากประกอบด้วย P3, P5, P7 และ P9



สามเหลี่ยม 2 รูปใดเหมือนกัน

- A. I และ IV
- B. I และ II
- C. II และ III
- D. II และ IV
- E. III และ IV

- เป็นปัญหาเกี่ยวกับเรขาคณิต.....C4
- ใช้รูปภาพ.....S4
- ต้องประเมินทางเลือกเพื่อให้ได้คำตอบ.....S5
- ประยุกต์ใช้การตัดสินใจที่เหมาะสม(เหมือนหรือไม่) .....P3
- I และ II มีด้านเหมือนกัน 3 ด้าน และมีมุมเท่ากันจึงเป็น สามเหลี่ยมที่เหมือนกัน.....P5
- การจับคู่อื่น ๆ ไม่เหมาะสม.....P9
- ความเข้าใจความสัมพันธ์ระหว่างรูปภาพ เช่นด้านเหมือนกัน...P7

แผนภาพที่ 1 ตัวอย่างการจำแนกคุณลักษณะที่ใช้ในการตอบคำถาม

#### 4.4 การวินิจฉัยและการจำแนกนักเรียนโดยใช้ rule space model

การวิเคราะห์ rule space ดำเนินการแยกวิเคราะห์ในแต่ละประเทศโดยใช้เฉพาะข้อมูลจากแบบทดสอบย่อยที่ 1, 3, 5 และ 7 เนื่องจากข้อมูลจากแบบทดสอบย่อยอื่น ๆ มีการกระจายคุณลักษณะไม่สม่ำเสมอ ทั้งนี้ได้ตัดคุณลักษณะ C6, S1, S9 และ P8 ออก เนื่องจากพบว่ามีค่าเฉลี่ยน้อยมากจนไม่สามารถประมาณค่าคุณลักษณะได้อย่างน่าเชื่อถือ

##### ผลการวิจัย

###### 1) ค่าสถิติพื้นฐาน

- ก. กลุ่มตัวอย่าง จำนวน 51,435 คน ใน 20 ประเทศ
- ข. มีค่าเฉลี่ยของระยะทางน้อยที่สุดตามกฎระยะทางกำลังสองน้อยที่สุด ( $D^2 =$  Mahalanobis distance) ซึ่งแสดงถึงความใกล้เคียงของความรอบรู้ของนักเรียนกับสถานะความรอบรู้ มีค่าน้อยกว่า .05 ซึ่งเป็นค่าที่ต่ำมากเมื่อเทียบกับเกณฑ์จุดตัดที่ตั้งไว้ เท่ากับ 4.5
- ค. มีความสำเร็จในการจำแนกนักเรียนตามแบบแผนการตอบข้อสอบแบบแผนใดแบบแผนหนึ่งได้ ร้อยละ 99.5
- ง. คุณลักษณะที่จำแนกได้ จำนวน 27 คุณลักษณะมีความสามารถในการทำนายค่าความยากได้ดี มีค่า adjusted  $R^2$  เท่ากับ 0.869

###### 2) ผลการเปรียบเทียบระหว่างประเทศ

- ก. แต่ละประเทศมีค่าความคลาดเคลื่อนมาตรฐานใกล้เคียงกัน ประเทศที่มีค่าความคลาดเคลื่อนมาตรฐานต่ำที่สุด ได้แก่ ฮังการีและเกาหลี
- ข. ประเทศที่มีจำนวนกลุ่มตัวอย่างน้อยที่สุด คือ สาธารณรัฐเช็ก (1,700 คน) ประเทศที่มีจำนวนกลุ่มตัวอย่างมากที่สุด คือ สหรัฐ (4,500 คน)
- ค. คุณลักษณะที่ยากมากที่สุด คือ ทักษะการคิดแบบอนุมาน (S6) คุณลักษณะที่ง่ายมากที่สุด ได้แก่ การประเมิน, การพิสูจน์และการตรวจสอบทางเลือก (S5) การใช้รูปภาพและตาราง (S3) และการแปลงข้อความเป็นสูตรหรือพีชคณิต (P1)
- ง. ประเทศที่มีกระบวนการทางพุทธิปัญญาดีที่สุด คือ ญี่ปุ่น มีค่าความน่าจะเป็นในการรอบรู้คุณลักษณะที่ยากสูงสุด 5 คุณลักษณะ ได้แก่ การคิดเชิงอุปมัย (P5) การคิดเชิงวิเคราะห์



(P6) การประยุกต์ใช้มโนทัศน์และทฤษฎี (P3) กระบวนการจัดการและข้อมูล (P9) และทักษะการคิดเชิงอนุมาน (S6)

จ. ประเทศที่มีค่าเฉลี่ยของผลสัมฤทธิ์ในแต่ละคุณลักษณะสูงสุด แสดงในตารางที่ 2 ตารางที่ 2 ประเทศที่มีค่าเฉลี่ยของผลสัมฤทธิ์ในแต่ละคุณลักษณะสูงสุด

ประเทศ	คุณลักษณะที่มีค่าเฉลี่ยสูงสุด
ออสเตรเลีย	การคาดคะเนและการประมาณค่า (S4)
เบลเยียม-ฟรานซิส	มโนทัศน์เบื้องต้นและปฏิบัติการเกี่ยวกับจำนวนและจำนวนเต็ม (C1) มโนทัศน์เบื้องต้นและปฏิบัติการเกี่ยวกับเศษส่วนและทศนิยม (C2) ความเข้าใจคำที่ใช้ในการตั้งคำถาม (S11)
สาธารณรัฐเช็ก	การแก้ปัญหาใหม่หรือปัญหาที่ไม่คุ้นเคย (S8)
ฟินแลนด์	การประยุกต์ใช้คุณสมบัติและความสำคัญของจำนวน, ความหมายของจำนวนเต็มจำนวน (S2)
ฮ่องกง	มโนทัศน์เบื้องต้นและปฏิบัติการเกี่ยวกับพีชคณิตเบื้องต้น (C3) มโนทัศน์เบื้องต้นและปฏิบัติการเกี่ยวกับเรขาคณิต 2 มิติ (C4)
ญี่ปุ่น	ข้อมูล, ความน่าจะเป็นและสถิติพื้นฐาน (C5) แบบแผนและความสัมพันธ์ (ทักษะการคิดเชิงอนุมาน) (S6) การประยุกต์ใช้ความรู้ในการตัดสินใจทางเลขคณิตและเรขาคณิต (P3) เหตุผลเชิงตรรกะ ได้แก่ การรวมเหตุผล การคิดเชิงอนุมาน, ถ้า-แล้ว, ความจำเป็นและความเพียงพอ, การสรุปอ้างอิง (P5) การค้นหาปัญหา, การคิดเชิงวิเคราะห์, การสร้างปัญหาขึ้นใหม่, การคิดเชิงอุปนัย (P6)
เกาหลี	มโนทัศน์เบื้องต้นและปฏิบัติการเกี่ยวกับเรขาคณิต 2 มิติ (C4) การใช้รูปภาพ, ตาราง, แผนภาพและกราฟ (S3) การจัดการและการดำเนินการเกี่ยวกับข้อมูล (P9) การอ่านเชิงปริมาณและการอ่านเชิงตรรกะ (P10)
เนเธอร์แลนด์	การประเมิน, การพิสูจน์และการตรวจสอบทางเลือก (S5) การสร้าง การดูและการอ่านภาพและกราฟ (P7)
สิงคโปร์	การใช้เหตุผลอย่างสมเหตุสมผล (S7) การแปลงหรือการสร้างสูตรและการแก้ปัญหา (P1) การประยุกต์ใช้ความรู้ในการคำนวณทางเลขคณิตและเรขาคณิต (P2)

---

การประยุกต์ใช้กฎในพีชคณิต (P4)

คำถามปลายเปิดที่ไม่ให้คำตอบ (S10)

---

### มโนทัศน์เบื้องต้นเกี่ยวกับโมเดลลำดับชั้นของคุณลักษณะ

โมเดลลำดับชั้นของคุณลักษณะ (attribute hierarchy model) เป็นโมเดลทฤษฎีการตอบสนองข้อสอบทางพุทธิปัญญา (cognitive item response theory model) พัฒนาขึ้นโดย Jacqueline P. Leighton, Mark J. Gierl และ Stephen M. Hunka จาก Centre for research in applied measurement and evaluation, University of Alberta โดยนำเสนอแนวคิดครั้งแรกในเอกสาร Educational Measurement : Issue and Practice ของ National Council on Measurement in Education (NCME) เมื่อปี ค.ศ. 2000 ในชื่อเรื่อง “Exploring the logic of Tatsuoka' s rule space model for test development and analysis” ต่อมานำเสนอในการประชุมประจำปี ของ National Council on Measurement in Education (NCME) เมื่อ วันที่ 2 - 4 เดือนสิงหาคม ปี ค.ศ. 2002 ที่ New Orleans, Louisiana ในชื่อเรื่อง “The Attribute Hierarchy Model for Cognitive Assessment” และได้ตีพิมพ์แนวคิดดังกล่าวใน Journal of Educational Measurement เมื่อปี ค.ศ. 2004 ในชื่อเรื่อง “The attribute hierarchy method for cognitive assessment : A variation on Tatsuoka' s rule-space approach” โมเดลลำดับชั้นของคุณลักษณะ (attribute hierarchy model) เป็นโมเดลทฤษฎีการตอบสนองข้อสอบทางพุทธิปัญญา (cognitive item response theory model) ถูกพัฒนาขึ้นเพื่อแก้ไขข้อจำกัดของ Rule space model โดยมีวัตถุประสงค์เพื่อให้การประเมินเชิงวินิจฉัยมีความชัดเจนเป็นประโยชน์ต่อการพัฒนาการเรียนการสอนและการพัฒนานักเรียนมากขึ้น โดยการกำหนดจำนวนข้อสอบ ลักษณะข้อสอบและแบบแผนการตอบข้อสอบจากลักษณะโครงสร้างที่เป็นลำดับชั้นของความรู้หรือทักษะที่ต้องใช้ในการตอบข้อสอบให้ถูกต้องเรียกว่าคุณลักษณะ (attribute) ซึ่งแสดงให้เห็นถึงประโยชน์ของการเชื่อมโยงระหว่างทฤษฎีทางพุทธิปัญญา (cognitive theory) กับการวัดทางจิต(psychometric practice)

โมเดลลำดับชั้นของคุณลักษณะมีลักษณะสำคัญและขั้นตอนการนำไปใช้ในการพัฒนาแบบทดสอบและการประเมินเชิงวินิจฉัย ดังนี้

#### 1. ลักษณะสำคัญของโมเดลลำดับชั้นของคุณลักษณะ

ลักษณะสำคัญของโมเดลลำดับชั้นของคุณลักษณะ ก็คือ การให้คำจำกัดความและการจำแนกคุณลักษณะ รวมทั้งการกำหนดลักษณะความสัมพันธ์ที่เป็นลำดับชั้นของคุณลักษณะ มีรายละเอียด ดังนี้

### 1.1 การให้คำจำกัดความและการจำแนกคุณลักษณะ

คุณลักษณะเป็นตัวแปรนำเข้าที่สำคัญใน Rule space model เพราะให้การสรุปอ้างอิงถึงทักษะทางด้านพุทธิปัญญาของผู้สอบ แต่จากเอกสารที่เกี่ยวข้องกับการนำ Rule space model ไปใช้ ไม่ได้ให้คำจำกัดความคุณลักษณะไว้อย่างชัดเจน แต่จะมีการให้คำจำกัดความที่แตกต่างกันไป รวมทั้งมีการใช้เทคนิคในการจำแนกที่หลากหลาย นักวิจัยบางคนกล่าวถึง คุณลักษณะในความหมายทั่วไป โดยพิจารณาจากทักษะทางพุทธิปัญญาที่จำเป็นต้องใช้ในการปฏิบัติงาน บางคนกล่าวถึงคุณลักษณะในความหมายเฉพาะในลักษณะการรวมองค์ประกอบย่อยทางพุทธิปัญญาเพื่อใช้ในการแก้ปัญหา ทักษะพื้นฐานทางพุทธิปัญญาที่ต้องใช้เพื่อความรอบรู้ในเนื้อหาเฉพาะหรือชุดของวิธีการหรือการปฏิบัติที่ใช้ในการแก้ปัญหาในเนื้อหาเฉพาะ คุณลักษณะจึงถูกนำไปใช้ทั้งในความหมายเฉพาะและความหมายทั่วไป เช่น Birenbaum Kelly และ Tatsuoka (1993) อ้างถึงใน Gierl, Leighton and Hunka, (2000) ให้คำจำกัดความ คุณลักษณะไว้ว่า “เป็นการบรรยายกระบวนการ ทักษะหรือความรู้ของนักเรียนที่ต้องใช้เพื่อให้สามารถทำงานได้บรรลุเป้าหมาย .....อาจรวมกันได้ แต่ไม่จำกัดเฉพาะความสามารถของนักเรียน ในบางวิธีการคุณลักษณะอาจรวมถึงสิ่งที่นักเรียนใช้ช่วยในการเรียนรู้หรือยุทธวิธีที่ใช้”

สำหรับวิธีการที่ใช้ในการจำแนกคุณลักษณะ ก็ไม่มีเอกสารที่แสดงถึงวิธีการที่ใช้ในการจำแนกไว้อย่างชัดเจนเช่นกัน Gierl, Leighton and Hunka (2000) ยกตัวอย่างวิธีการที่ใช้ในการจำแนกคุณลักษณะ สรุปได้ดังนี้

1) ใช้กรอบมโนทัศน์ เช่น Kate Martinez Sheehan and Tatsuoka (1998) ใช้โมเดลการแก้ปัญหาในการประเมินแบบทดสอบทางสถาปัตยกรรมจำนวน 22 ข้อ โมเดลมี 3 ขั้นตอน ได้แก่ สร้างตัวแทนของข้อสอบขั้นต้น กำหนดเป้าหมายและการปฏิบัติในแต่ละเป้าหมาย ทั้งที่สำเร็จและที่พึงพอใจ คุณลักษณะของข้อสอบได้มาจากกระบวนการทางพุทธิปัญญาที่มีลักษณะเฉพาะในแต่ละชั้นของโมเดล นอกจากนี้ Sheehan, Tatsuoka and Lawis (1992) ยังแนะนำให้ใช้กรอบมโนทัศน์ในลักษณะทั่วไป โดยใช้ทั้งการจัดกลุ่มข้อสอบและโมเดลทางพุทธิปัญญา

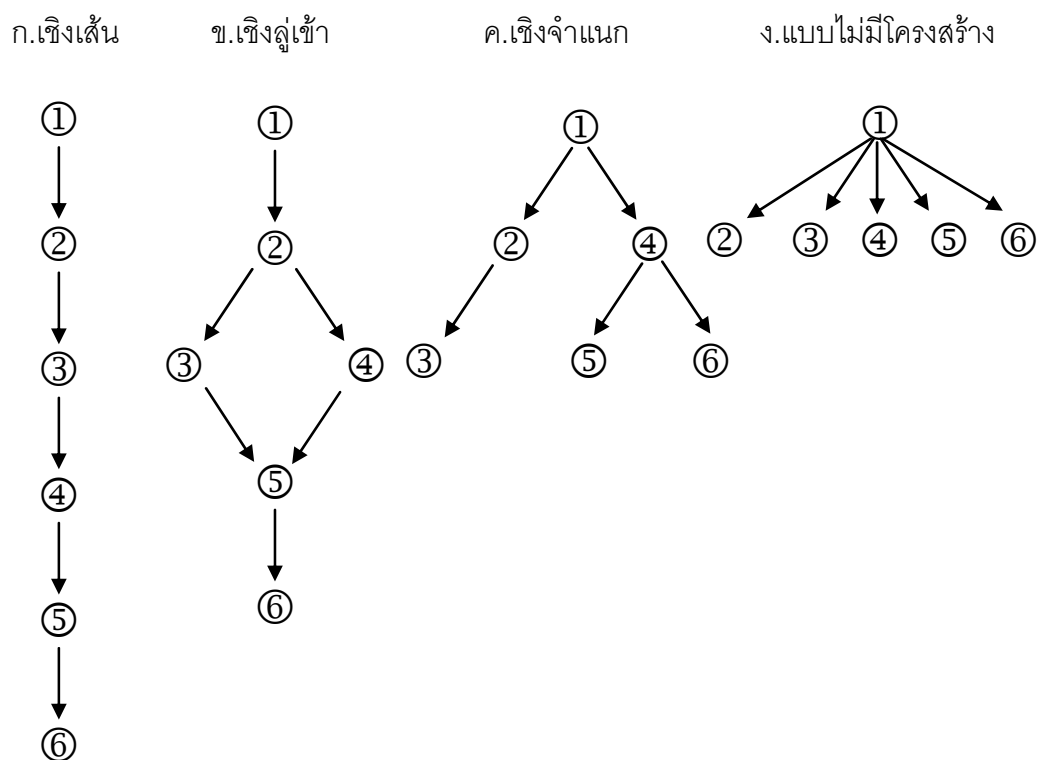
2) ใช้การจับคู่ประโยคและวิธีการที่ใช้ในทฤษฎีภาษา เช่น Tatsuoka, Birenbaum, Lewis และ Sheehan (1992) จำแนกคุณลักษณะที่เป็นเนื้อหาและกระบวนการจากข้อมูลต้นร่าง (protocol data) ของ Scholastic Assessment Test (SAT) โดยจับคู่ประโยคได้ 13 ภาษา แต่ละภาษาจะประกอบด้วยความรู้เบื้องต้น (element) ซึ่งนำมาจำแนกเป็นคุณลักษณะสำหรับแบบทดสอบ

3) ใช้การตัดสินใจของผู้เชี่ยวชาญ Tatsuoka เสนอว่า การวิเคราะห์งานขั้นต้นสามารถใช้ผู้เชี่ยวชาญหรือครูที่เชี่ยวชาญหลาย ๆ คน ถ้าผู้เชี่ยวชาญ จำนวน 2 คน หรือมากกว่ามีวิธีการแก้ปัญหาที่ต่างกันสามารถใช้คุณลักษณะที่ต่างกันในเมทริกซ์ของข้อสอบได้ ในกรณีนี้อาจจะมี incidence matrices ที่ต่างกันหลายเมทริกซ์ และแต่ละเมทริกซ์สามารถทดสอบโดย Rule space model ได้ ทั้งนี้อาจจำแนกคุณลักษณะจากวิธีการหรือใช้แหล่งข้อมูลที่ต่างกัน เช่น ใช้ผู้เชี่ยวชาญด้านเนื้อหาหรือจากต้นร่างเอกสารของนักเรียน เป็นต้น

ความไม่ชัดเจนในการให้คำจำกัดความและการจำแนกคุณลักษณะที่กล่าวมาข้างต้น เป็นเหตุให้คุณลักษณะที่ปรากฏในเอกสารที่เกี่ยวข้องกับการนำ Rule space model ไปใช้ มีข้อจำกัดด้านมิติสัมพันธ์เบื้องต้น เนื่องจากการใช้เทคนิค โมเดลหรือวิธีการที่ต่างกันทำให้คุณลักษณะมีคุณภาพที่แตกต่างกันจึงไม่สามารถเชื่อมโยงพื้นฐานความรู้จากการศึกษาหนึ่งไปยังการศึกษาอื่น ๆ ได้ โดยเฉพาะประโยชน์ในการแปลความหมายจากผลการทดสอบ เพื่อแก้ไขข้อจำกัดดังกล่าว Gierl, Leighton และ Hunka (2000) จึงได้กล่าวถึง คุณลักษณะไว้ว่า “คุณลักษณะเป็นการบรรยายวิธีการหรือการอธิบายความรู้ที่ต้องใช้ในการทำงานที่มีขอบเขตเฉพาะ แม้ว่าคุณลักษณะจะไม่ใช้ยุทธวิธี แต่คุณลักษณะก็จะช่วยสร้างยุทธวิธี ยิ่งไปกว่านั้นชุดของคุณลักษณะจะถูกจัดเข้าสู่ยุทธวิธีเพื่อสนับสนุนบทบาทในการแก้ปัญหาที่เกิดขึ้นทันทีทันใด โดยไม่จำเป็นต้องเป็นกลุ่มยุทธวิธี คุณลักษณะเป็นเอกลักษณ์ที่มีลักษณะเป็นพลวัต ชุดของคุณลักษณะของนักเรียนที่มีความสามารถในเวลาหนึ่งอาจจะไม่เหมือนชุดของคุณลักษณะของนักเรียนที่มีความสามารถในเวลาหนึ่ง ทั้งนี้อาจเป็นพัฒนาการและ/หรือขึ้นอยู่กับการเรียนการสอน หมายความว่าความก้าวหน้าของนักเรียนจากเวลาหนึ่งไปอีกเวลาหนึ่งเกิดขึ้นจากการมีพัฒนาการและ/หรือองค์ประกอบด้านการเรียนการสอน คุณลักษณะสำหรับแบบทดสอบสามารถจำแนกได้โดยใช้วิธีการที่แตกต่างกัน เช่น ความคิดเห็นของผู้เชี่ยวชาญ การวิเคราะห์งาน การเขียนตอบของนักเรียน โดยควรตรวจสอบความตรงของการจำแนกคุณลักษณะ ทั้งจากผู้สอบและข้อสอบเปรียบเทียบกับกลุ่มประชากรเป้าหมาย”

## 1.2 การกำหนดลักษณะความสัมพันธ์ที่เป็นลำดับชั้นของคุณลักษณะ

นอกจากการให้คำจำกัดความและการจำแนกคุณลักษณะได้อย่างถูกต้องแล้ว สิ่งสำคัญประการหนึ่งในการนำโมเดลลำดับชั้นของคุณลักษณะไปใช้ได้อย่างมีประสิทธิภาพ ก็คือ การกำหนดลักษณะความสัมพันธ์ที่เป็นลำดับชั้นที่ถูกต้อง โดยอาจพิจารณาเชิงประจักษ์จากการให้คำจำกัดความที่ชัดเจนหรือจากการวิเคราะห์ต้นร่างเอกสารหรือพิจารณาเชิงทฤษฎีจากจิตวิทยา พัฒนาการ เช่น จากขั้นพัฒนาการของเพียเจต์ เป็นต้น สำหรับโครงสร้างของลำดับชั้น Leighton, Gierl and Hunka (2004) ได้ยกตัวอย่างโครงสร้างของลำดับชั้นอย่างง่ายที่ใช้ในการพัฒนาแบบทดสอบไว้ 4 รูปแบบ ดังนี้



แผนภาพที่ 2 โครงสร้างของลำดับชั้น

1) ลำดับชั้นเชิงเส้น (linear hierarchy) มีลักษณะความสัมพันธ์ระหว่างคุณลักษณะในลักษณะเส้นตรงที่มีจุดสิ้นสุดจุดเดียว เช่น จากแผนภาพที่ 2ก. แสดงให้เห็นว่าคุณลักษณะที่ 1 จะต้องมีมาก่อนคุณลักษณะที่ 2 คุณลักษณะที่ 2 จะต้องมีมาก่อนคุณลักษณะที่ 3 คุณลักษณะที่ 3 จะต้องมีมาก่อนคุณลักษณะที่ 4 คุณลักษณะที่ 4 จะต้องมีมาก่อนคุณลักษณะที่ 5 และคุณลักษณะที่ 5 จะต้องมีมาก่อนคุณลักษณะที่ 6 เป็นต้น

2) ลำดับชั้นเชิงลู่เข้า (hierarchy with a convergent branch) มีลักษณะความสัมพันธ์ที่แยกเป็น 2 ทางแต่มีจุดสิ้นสุดจุดเดียว เช่น จากแผนภาพที่ 2ข. แสดงให้เห็นว่าคุณลักษณะที่ 1 จะต้องมีมาก่อนคุณลักษณะที่ 2 คุณลักษณะที่ 2 จะต้องมีมาก่อนคุณลักษณะที่ 3

และคุณลักษณะที่ 4 คุณลักษณะที่ 3 และคุณลักษณะที่ 4 จะต้องมียาก่อน คุณลักษณะที่ 5 และคุณลักษณะที่ 5 จะต้องมียาก่อน คุณลักษณะที่ 6 เป็นต้น

3) ลำดับชั้นเชิงจำแนก (hierarchy having a divergent branch) มีลักษณะความสัมพันธ์ที่แยกออกเป็น 2 ทาง แต่ไม่ได้มีจุดสิ้นสุดเพียงจุดเดียว เช่น จากแผนภาพที่ 2ค. แสดงให้เห็นว่า คุณลักษณะที่ 1 จะต้องมียาก่อนคุณลักษณะที่ 2 และคุณลักษณะที่ 4 คุณลักษณะที่ 2 จะต้องมียาก่อนคุณลักษณะที่ 3 และคุณลักษณะที่ 4 จะต้องมียาก่อนคุณลักษณะที่ 5 และคุณลักษณะที่ 6 เป็นต้น

4) ลำดับชั้นแบบไม่มีโครงสร้าง (unstructured hierarchy) มีลักษณะความสัมพันธ์ที่ไม่มีความสัมพันธ์เดียวจากจุดเริ่มต้นและไม่มีจุดสิ้นสุดเพียงจุดเดียว เช่น จากแผนภาพที่ 2ง. แสดงให้เห็นว่า คุณลักษณะที่ 1 จะต้องมียาก่อนคุณลักษณะที่ 2 คุณลักษณะที่ 3 คุณลักษณะที่ 4 คุณลักษณะที่ 5 และคุณลักษณะที่ 6 โดยที่คุณลักษณะที่ 2 ถึงคุณลักษณะที่ 6 ไม่มีความสัมพันธ์กัน เป็นต้น

## 2. ขั้นตอนการนำโมเดลลำดับชั้นของคุณลักษณะไปใช้ในการพัฒนาแบบทดสอบและการวินิจฉัย

การนำโมเดลลำดับชั้นของคุณลักษณะไปใช้ในการพัฒนาแบบทดสอบและการวินิจฉัยมีขั้นตอนสำคัญ 7 ขั้นตอน ดังนี้

### 2.1 ระบุลักษณะและความสัมพันธ์เชิงลำดับชั้นที่มีลักษณะเฉพาะ

ในขั้นตอนการระบุลักษณะและความสัมพันธ์เชิงลำดับชั้นที่มีลักษณะเฉพาะ นักพัฒนาแบบทดสอบจะต้องทำความเข้าใจกระบวนการทางพุทธิปัญญาที่จะใช้วัดกลุ่มผู้สอบที่มีลักษณะเฉพาะให้ชัดเจน ให้คำจำกัดความคุณลักษณะอย่างรอบคอบ กำหนดความสัมพันธ์ที่เป็นลำดับชั้นและนำเสนอลำดับชั้นโดยใช้โครงสร้างรูปต้นไม้ (tree structure) ซึ่งบางครั้งเรียกว่าโครงสร้างเครือข่าย (network structure) ความสัมพันธ์ที่เป็นลำดับชั้นในข้อสอบแต่ละข้อจะต้องนำเสนอเพื่อให้ผู้สอบตอบข้อสอบอย่างไม่สับสน แม้ว่าจะมีปฏิสัมพันธ์กันระหว่างคุณลักษณะที่ทำได้กับลำดับที่นำเสนอให้ทำภายในเฉพาะแต่ละข้อ (เช่น ยุทธวิธีในการเสนอ) จะต้องไม่มี 2 มโนทัศน์ในแต่ละข้อ ต้องวิเคราะห์และระบุลักษณะทางพุทธิปัญญาอย่างชัดเจน สมเหตุสมผล และเป็นไปตามหลักจิตวิทยา ทั้งนี้ควรกำหนดคุณลักษณะที่เป็นจุดเริ่มซึ่งเป็นสิ่งที่ต้องมีมาก่อนสำหรับคุณลักษณะทั้งหมดจากจุดเดียวกัน เพื่อช่วยให้นักพัฒนาแบบทดสอบตระหนักถึงคุณลักษณะเริ่มต้นที่ผู้สอบทุกคนต้องมี

## 2.2 กำหนดเมทริกซ์เพื่อใช้ในการพัฒนาแบบทดสอบ

การกำหนดเมทริกซ์เพื่อใช้ในการพัฒนาแบบทดสอบมีขั้นตอน ดังนี้

1) กำหนด binary adjacency matrix (A) เพื่อแสดงความสัมพันธ์ทางตรงระหว่างคุณลักษณะ โดยกำหนดให้ matrix (A) มีขนาด  $k \times k$  เมื่อ  $k$  เป็นจำนวนคุณลักษณะ เลข 1 ณ ตำแหน่ง  $(j,k)$  ในเมทริกซ์จะบ่งบอกถึงการมีความสัมพันธ์ทางตรงระหว่างคุณลักษณะที่  $j$  ในลักษณะการเป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะที่  $k$  เช่นจากแผนภาพที่ 2ค. เขียนเป็น matrix (A) ได้ดังนี้

010100	จากเมทริกซ์แสดงให้เห็นว่า คุณลักษณะที่ 1 เป็น คุณลักษณะที่ต้องมีมาก่อน
001000	คุณลักษณะที่ 2 และ 4 เห็นได้จากตำแหน่งที่ (1,2) และ (1,4) แทนด้วยเลข 1
000000	คุณลักษณะที่ 2 เป็นคุณลักษณะที่ต้องมีมาก่อน คุณลักษณะที่ 3 เห็นได้จาก
000011	ตำแหน่งที่ (2,3) แทนด้วยตัวเลข 1 และคุณลักษณะที่ 4 เป็นคุณลักษณะที่ต้อง
000000	มีมาก่อนคุณลักษณะที่ 5 และ 6 เห็นได้จากตำแหน่งที่ (4,5) และ (4,6) แทน
000000	ด้วยเลข 1

2) กำหนด reachability matrix (R) เพื่อแสดงความสัมพันธ์ทางตรงและความสัมพันธ์ทางอ้อมระหว่างคุณลักษณะ โดยกำหนดให้ matrix (R) มีขนาด  $k \times k$  เมื่อ  $k$  เป็นจำนวนคุณลักษณะ matrix (R) จะบ่งบอกเงื่อนไขในโครงสร้างของลำดับขั้นซึ่งเป็นประโยชน์ในการกำหนดชุดย่อยของข้อคำถาม คำนวณโดยใช้สูตร  $R = (A+I)^n$  เมื่อ  $n$  เป็นจำนวนเต็มและ  $I$  เป็นเมทริกซ์เอกลักษณ์ (identity matrix) จาก matrix (A) ในข้อ 1) กำหนดเป็น matrix (R) ได้ดังนี้

111111	แถวที่ 1 แสดงให้เห็นว่า คุณลักษณะที่ 1 เป็นคุณลักษณะที่ต้องมีมาก่อน
011000	ทุกคุณลักษณะ แถวที่ 2 แสดงให้เห็นว่า คุณลักษณะที่ 2 เป็นคุณลักษณะ
001000	ที่ต้องมีมาก่อนคุณลักษณะที่ 2 และ 3 แถวที่ 3 แสดงให้เห็นว่า คุณลักษณะที่ 3
000111	ไม่ได้เป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะใด แถวที่ 4 แสดงให้เห็นว่า
000010	คุณลักษณะที่ 4 เป็นคุณลักษณะที่ต้องมีมาก่อน คุณลักษณะที่ 4, 5 และ 6
000001	แถวที่ 5 และแถวที่ 6 แสดงให้เห็นว่าคุณลักษณะที่ 5 และที่ 6 ไม่ได้เป็น
	คุณลักษณะที่ต้องมีมาก่อนคุณลักษณะใด

3) กำหนด incidence matrix (Q) เพื่อแสดงชุดของข้อสอบและคุณลักษณะที่ต้องใช้ในการทำข้อสอบในแต่ละข้อให้ถูกต้อง ชุดของข้อสอบที่เป็นไปได้มีขนาดเท่ากับ  $2^k - 1$





### 2.3 พัฒนาแบบทดสอบ

การพัฒนาแบบทดสอบดำเนินการโดยสร้างข้อสอบให้มีการใช้คุณลักษณะตามที่กำหนด

2.4 กำหนดแบบแผนการตอบข้อสอบที่คาดหวัง (expected item response pattern) คะแนนรวม (total score) และคุณลักษณะของผู้สอบในแต่ละแบบแผนการตอบข้อสอบ

จาก matrix (Q) ในข้อ 4) กำหนดแบบแผนการตอบข้อสอบที่คาดหวัง คะแนนรวมและคุณลักษณะของผู้สอบได้ ดังแสดงในตารางที่ 3

ตารางที่ 3 เมทริกซ์การตอบข้อสอบที่คาดหวัง คะแนนรวมและ คุณลักษณะของผู้สอบ

สำหรับชุดของผู้สอบจำนวน 15 คน จากพื้นฐานโครงสร้างของลำดับชั้นในแผนภาพที่ 2ค.

ผู้สอบ	เมทริกซ์การตอบข้อสอบที่คาดหวัง	คะแนนรวม	คุณลักษณะของผู้สอบ
1	100000000000000	1	100000
2	110000000000000	2	110000
3	111000000000000	3	111000
4	100100000000000	2	100100
5	110110000000000	4	110100
6	111111000000000	6	111100
7	100100100000000	3	100110
8	110110110000000	6	110110
9	111111111000000	9	111110
10	100100000100000	3	100101
11	110110000110000	6	110101
12	111111000111000	9	111101
13	100100100100100	5	100111
14	110110110110110	10	110111
15	111111111111111	15	111111

### 2.5 นำแบบทดสอบไปใช้

## 2.6 ประมาณค่าความน่าจะเป็นในการตอบข้อสอบ (estimating probability of item response)

ประมาณค่าความน่าจะเป็นในการตอบข้อสอบในแต่ละข้อจากแบบแผนการตอบข้อสอบที่คาดหวัง โดยใช้ดัชนีความสอดคล้องกับบุคคล (person-fit indices) จากโอกาสในการตอบข้อสอบได้ถูกต้องโดยใช้โมเดลทฤษฎีการตอบสนองข้อสอบ (IRT model) ซึ่งจะทำให้ได้ค่าอำนาจจำแนกของข้อสอบ (a-Parameter) และค่าความยากของข้อสอบ (b-Parameter) กรณีใช้โมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ ทั้งนี้อาจใช้โปรแกรมสำเร็จรูปช่วยในการคำนวณ เช่น โปรแกรม BILOG เป็นต้น

## 2.7 จำแนกแบบแผนการตอบข้อสอบที่สังเกตได้

การจำแนกแบบแผนการตอบข้อสอบที่สังเกตได้ มีขั้นตอน ดังนี้

1) การจำแนกเบื้องต้น (Preliminary classification) วิธีการนี้ใช้การเปรียบเทียบแบบแผนการตอบข้อสอบที่สังเกตได้กับแบบแผนการตอบข้อสอบที่คาดหวังทั้งหมด และคำนวณค่าความเป็นไปได้สูงสุดที่แบบแผนการตอบข้อสอบที่สังเกตได้จะสอดคล้องกับแบบแผนการตอบข้อสอบที่คาดหวัง ถ้ามีความเป็นไปได้สูงสุดที่แบบแผนการตอบข้อสอบที่คาดหวังใด แสดงว่าแบบแผนการตอบข้อสอบที่สังเกตได้นั้นมีคุณลักษณะของผู้สอบตามที่แบบแผนการตอบข้อสอบที่คาดหวังนั้นกำหนด เช่น นาย ก. มีแบบแผนการตอบข้อสอบที่สังเกตได้เป็น 1111000000000000 เมื่อนำไปเปรียบเทียบกับค่าความเป็นไปได้ในแต่ละแบบแผนการตอบข้อสอบที่คาดหวังพบว่า มีค่าความเป็นไปได้สูงสุดเท่ากับ 0.500 ที่ระดับความสามารถเท่ากับ -0.50 ตรงกับแบบแผนการตอบข้อสอบที่คาดหวัง 1110000000000000 ซึ่งมีความคลาดเคลื่อนเกิดขึ้น 1 จุด และมีคุณลักษณะของผู้สอบเป็น 111000 แสดงว่านาย ก. จัดให้อยู่ในกลุ่มที่มีคุณลักษณะที่ 1,2 และ 3 เป็นต้น แบบแผนการตอบข้อสอบที่สังเกตได้ในการจำแนกเบื้องต้น มีรายละเอียดการคำนวณและสูตรที่ใช้ในการคำนวณ ดังนี้

ก. คำนวณโอกาสในการเกิดความคลาดเคลื่อน จาก  $1 \rightarrow 0$  และจาก  $0 \rightarrow 1$  เพื่อคำนวณความเป็นไปได้สูงสุดที่แบบแผนการตอบข้อสอบที่สังเกตได้จะสอดคล้องกับแบบแผนการตอบข้อสอบที่คาดหวัง โดยกำหนดให้

$V_j$  เป็นแบบแผนการตอบข้อสอบที่คาดหวังที่  $j$  สำหรับข้อสอบ  $n$  ข้อ

$X$  เป็นแบบแผนการตอบข้อสอบที่สังเกตได้ซึ่งมีจำนวนข้อเท่ากับ  $V_j$

$d_j = V_j - X$  เป็นความแตกต่างระหว่างแบบแผนการตอบข้อสอบ  
ที่คาดหวังกับแบบแผนการตอบข้อสอบที่สังเกตได้

ถ้า  $d$  เท่ากับ 0 แสดงว่าไม่มีความคลาดเคลื่อน

ถ้า  $d$  เท่ากับ -1 แสดงว่ามีความคลาดเคลื่อน จาก  $0 \rightarrow 1$

ถ้า  $d$  เท่ากับ 1 แสดงว่ามีความคลาดเคลื่อน จาก  $1 \rightarrow 0$

ข. คำนวณค่าความน่าจะเป็นที่แบบแผนการตอบข้อสอบที่สังเกตได้จะ  
จัดอยู่ในแบบแผนการตอบข้อสอบที่คาดหวัง โดยการรวมโอกาสที่จะตอบถูกเมื่อผิดและโอกาสที่  
จะตอบผิดเมื่อถูก ด้วยหลักการของ Boolean Algebra ตามสูตร ต่อไปนี้

กำหนดให้  $P_{jk}(\theta)$  เป็นโอกาสที่จะตอบถูกเมื่อผิด ( $0 \rightarrow 1$ )

$1 - P_{jm}(\theta)$  เป็นโอกาสที่จะตอบผิดเมื่อถูก ( $1 \rightarrow 0$ )

$$P_{j\text{Expected}}(\theta) = \prod_{k=1}^K P_{jk}(\theta) \prod_{m=1}^M [1 - P_{jm}(\theta)]$$

2) การตรวจสอบการจำแนกเบื้องต้น (Verification of preliminary classification)

การจำแนกเบื้องต้นบางครั้งให้ค่าความน่าจะเป็นต่ำมากและจำแนกได้ไม่สอดคล้องกับโครงสร้าง  
ของลำดับชั้น เช่น นาย ก. มีแบบแผนการตอบที่สังเกตเป็น 11111111111101 เมื่อคำนวณ  
ค่าความเป็นไปได้ พบว่า มีค่าความเป็นไปได้สูงสุดเท่ากับ 0.0318 ที่ระดับความสามารถเท่ากับ  
2.37 ตรงกับแบบแผนการตอบข้อสอบที่คาดหวัง 11111111111111 ซึ่งมีความคลาดเคลื่อน  
เกิดขึ้น 1 จุด และมีคุณลักษณะของผู้สอบเป็น 111111 แสดงว่านาย ก. จัดให้อยู่ในกลุ่มที่มี  
คุณลักษณะครบทุกข้อ ซึ่งในความเป็นจริงไม่ได้เป็นเช่นนั้น เนื่องจากนาย ก. ตอบข้อสอบได้  
ไม่ถูกต้องทุกข้อ ทำให้การจำแนกไม่ตรงกับข้อมูลจริง จึงต้องใช้การตรวจสอบการจำแนกเบื้องต้น  
โดยการตรวจสอบความสมเหตุสมผล (logically) ของแบบแผนการตอบข้อสอบที่คาดหวังทั้งหมด  
กับแบบแผนการตอบข้อสอบที่สังเกตได้ ถ้าแบบแผนการตอบข้อสอบที่คาดหวังมีความสมเหตุสมผล  
กับแบบแผนการตอบข้อสอบที่สังเกตได้ แสดงว่าผู้สอบมีคุณลักษณะตามแบบแผน  
การตอบข้อสอบที่คาดหวังนั้น ๆ แต่ถ้าตรวจสอบแล้วพบว่าไม่สมเหตุสมผลกับแบบแผนการตอบ  
ข้อสอบที่คาดหวังใดจึงคำนวณค่าความน่าจะเป็น เช่น แบบแผนการตอบข้อสอบที่สังเกตได้ของ  
นาย ก. เป็น 11111111111101 เมื่อเปรียบเทียบกับแบบแผนการตอบข้อสอบที่คาดหวัง จำนวน  
16 แบบแผน พบว่า มีความคลาดเคลื่อนที่ไม่สมเหตุสมผลเกิดขึ้น 2 แบบแผน ได้แก่ แบบแผนที่  
15 ได้แก่ 110110110110110 และ แบบแผนที่ 16 ได้แก่ 11111111111111 ซึ่งพบว่า เกิด

ความคลาดเคลื่อนจาก  $1 \rightarrow 0$  ในข้อที่ 14 จำนวน 1 จุด เมื่อคำนวณค่าความน่าจะเป็นพบว่า แบบแผนที่ 15 มีค่าความน่าจะเป็น เท่ากับ 0.6835 ที่ระดับความสามารถ 1.54 มีคุณลักษณะของผู้สอบเป็น 110111 ส่วนแบบแผนที่ 16 มีค่าความน่าจะเป็น เท่ากับ 0.0318 ที่ระดับความสามารถ 2.37 มีคุณลักษณะของผู้สอบเป็น 111111 ดังนั้นแบบแผนการตอบข้อสอบที่สังเกตได้ของนาย ก. จึงมีความเป็นไปได้ที่จะสอดคล้องกับแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 15 มากกว่าแบบแผนที่ 16 หมายความว่า นาย ก. มี คุณลักษณะตามแบบแผนการตอบข้อสอบที่คาดหวัง ตั้งแต่แบบแผนที่ 2 ถึงแบบแผนที่ 14 โดยสอดคล้องกับแบบแผนที่ 15 มากที่สุด แต่มีคุณลักษณะไม่ครบตามที่กำหนดในแบบแผนที่ 16

### 3. ตัวอย่างการนำโมเดลลำดับชั้นของคุณลักษณะไปใช้

โมเดลลำดับชั้นของคุณลักษณะเป็นโมเดลที่ได้รับการพัฒนาและนำเสนอแนวคิดให้เห็นเป็นรูปธรรม ตั้งแต่ปี ค.ศ.2002 งานวิจัยที่เกี่ยวข้องดำเนินการจำกัดอยู่เฉพาะใน University of Alberta ซึ่งเป็นสถานที่ทำงานของผู้นำเสนอแนวคิด โดยคณะผู้วิจัย นักศึกษาและอาจารย์อื่น ๆ ในมหาวิทยาลัย ปัจจุบันจึงยังไม่มีผลการวิจัยเผยแพร่อย่างหลากหลาย จากการศึกษางานวิจัยที่เกี่ยวข้องกับการนำโมเดลลำดับชั้นของคุณลักษณะไปใช้แสดงให้เห็นถึงตัวอย่างการนำไปใช้จำนวน 10 ตัวอย่าง มีรายละเอียด ดังนี้

#### 3.1 การใช้โมเดลลำดับชั้นของคุณลักษณะในการพัฒนาแบบทดสอบเพื่อการวินิจฉัย

Leighton, Gierl and Hunka (2004) ได้ทดลองประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะในเนื้อหาเกี่ยวกับการใช้เหตุผลเชิงตรรกะ (syllogistic reasoning) เพื่อแสดงให้เห็นว่าโมเดลลำดับชั้นของคุณลักษณะใช้ในการประเมินความสามารถทางพุทธิพิสัย (cognitive competencies) ที่เกี่ยวข้องกับการคิดระดับสูง (higher-level thinking task) ได้อย่างไร มีรายละเอียด ดังนี้

##### 1) กรอบแนวคิดที่ใช้ในการกำหนดลำดับชั้นของคุณลักษณะ

การกำหนดลำดับชั้นของคุณลักษณะใช้กรอบแนวคิดการใช้เหตุผลเชิงตรรกะตามทฤษฎีโมเดลทางสมองของ Philip Johnson-Laird (Philip Johnson-Laird's theory of mental models) เนื่องจากมีงานวิจัยสนับสนุนแนวคิดที่เชื่อถือได้อย่างแพร่หลาย ทฤษฎีโมเดลทางสมอง

ประกอบด้วยการใช้เหตุผลเชิงตรรกะโดยอาศัยข้อเท็จจริง (categorical syllogism) 2 ลักษณะ ได้แก่ เงื่อนไข (quantified premises) และข้อสรุป (quantified conclusion) ส่วนที่เป็นเงื่อนไข จะสะท้อนให้เห็นถึงความสัมพันธ์ระหว่างสิ่งหนึ่ง (subject : A) กับสิ่งที่เป็นข้อสรุป (predicate : C) โดยผ่านตัวกลาง (middle term : B) ส่วนที่เป็นข้อสรุปจะสะท้อนให้เห็นความสัมพันธ์ระหว่าง สิ่งหนึ่ง (subject : A) กับสิ่งที่เป็นข้อสรุป (predicate : C) เช่น

A ทุกตัวเป็น B (All A are B)

B ทุกตัวเป็น C (All B are C)

เพราะฉะนั้น A ทุกตัวเป็น C (All A are C)

ในเงื่อนไขและข้อสรุปจะประกอบด้วยประโยคที่เป็นการใช้เหตุผลเชิงตรรกะ เช่น A ทุกตัวเป็น B (All A are B) A บางตัวเป็น B (Some A are B) A บางตัวไม่เป็น B (Some A are not B) หรือ ไม่มี A ที่เป็น B (No A are B) การสร้างข้อสรุปให้ถูกต้องต้องยึดหลักการว่า ข้อสรุปจะถูกต้องเมื่อทุกเงื่อนไขเป็นจริง

สำหรับขั้นตอนในการใช้เหตุผลเชิงตรรกะ ประกอบด้วย 3 ขั้นตอน

**ขั้นตอนที่ 1** การสร้างโมเดลเบื้องต้น (construction of an initial model) ซึ่งเป็นการตีความเงื่อนไขที่กำหนดให้ เช่น

กำหนดให้ A บางตัวเป็น B (Some A are B)

ไม่มี B ที่เป็น C (No B are C)

สร้างโมเดลเบื้องต้นได้ดังนี้

A เท่ากับ B  $A = B$  หรือ

A ทุกตัวอาจจะเป็น B ?A = B ตีความจาก A บางตัวเป็น B

B ทุกตัวไม่เป็น C C

A ทุกตัวไม่เป็น C C ตีความจาก ไม่มี B ที่เป็น C

**ขั้นตอนที่ 2** การสร้างข้อสรุป (generation of a conclusion) เป็นการเขียน ข้อสรุปที่ได้จากโมเดลเบื้องต้น เช่น จากโมเดลเบื้องต้นข้างต้น สรุปได้ว่า ไม่มี A ที่เป็น C

**ขั้นตอนที่ 3** สร้างโมเดลที่เป็นไปได้ (construct alternative models) โดยการ พิจารณาโมเดลอื่น ๆ ที่เป็นไปได้ทั้งหมดเพื่อตรวจสอบความถูกต้องของข้อสรุป จากเงื่อนไขใน ขั้นตอนที่ 1 สร้างโมเดลที่เป็นไปได้ได้ดังนี้

**โมเดลที่ 1** มี A อย่างน้อย 1 ตัวที่ไม่เป็น B และไม่เป็น C

A เท่ากับ B  $A = B$

	<u>A อย่างน้อย 1 ตัวไม่เป็น B</u>	<u>?B</u>
เพราะฉะนั้น	B ทุกตัวไม่เป็น C	?A C
	A อย่างน้อย 1 ตัวไม่เป็น C	C

**โมเดลที่ 2** A ทุกตัว ไม่เป็น B ทุกตัว และไม่เป็น C ทุกตัว

	A เท่ากับ B	A = B
	<u>A ทุกตัวไม่เป็น B</u>	<u>?B</u>
เพราะฉะนั้น	B ทุกตัวไม่เป็น C	?A C
	A ทุกตัวไม่เป็น C	?A C

จากโมเดลที่เป็นไปได้ทั้ง 2 โมเดล จะเห็นได้ว่าข้อสรุปจากโมเดลเบื้องต้นที่ว่า ไม่มี A ที่เป็น C เป็นข้อสรุปที่ไม่ถูกต้อง เนื่องจากคำว่า บางตัว (some) หมายความถึง มีอย่างน้อย 1 ตัว A บางตัวที่เป็น B จึงหมายถึงมี A อย่างน้อย 1 ตัวที่เป็น B ไม่ได้หมายถึง A ทุกตัว ต้องเป็น B เสมอไป เมื่อกำหนดให้ ไม่มี B ที่เป็น C จึงไม่ได้หมายถึง ไม่มี A ที่เป็น C ด้วย จากโมเดลที่เป็นไปได้ที่สร้างใหม่ทั้ง 2 โมเดล ซึ่งกำหนดเงื่อนไขว่า มี A อย่างน้อย 1 ตัวที่ไม่เป็น B และไม่เป็น C รวมทั้ง A ทุกตัวไม่เป็น B ทุกตัว และไม่เป็น C ทุกตัว ทำให้สร้างข้อสรุปใหม่ได้ว่า A บางตัวเป็น C (Some A are C) ซึ่งแปลความหมายได้ว่า มี A อย่างน้อย 1 ตัวที่เป็น C หรือ A ทุกตัวเป็น C แต่เมื่อพิจารณาเงื่อนไขที่ว่า A บางตัวเป็น B และไม่มี B ที่เป็น C ข้อสรุปที่สร้างจากโมเดลที่เป็นไปได้ที่สร้างใหม่ทั้ง 2 โมเดลจึงยังเป็นข้อสรุปที่ไม่ถูกต้อง เนื่องจากไม่เป็นจริงในทุกเงื่อนไข ดังนั้นเพื่อแสดงให้เห็นว่า บางตัว ไม่ใช่ ทุกตัว จึงสร้างข้อสรุปใหม่ที่มีความถูกต้องจากทั้ง 3 โมเดล ได้ว่า A บางตัวไม่เป็น C (Some A are not C)

## 2) การดำเนินการ

ก. จำแนกคุณลักษณะที่ใช้ในการใช้เหตุผลเชิงตรรกะ

จากกรอบแนวคิดในข้อ 1) สามารถจำแนกคุณลักษณะในการใช้เหตุผลเชิงตรรกะได้จำนวน 7 คุณลักษณะ ดังนี้

คุณลักษณะที่ 1 (A1) สามารถแปลความจากเงื่อนไขที่กำหนดให้ได้

คุณลักษณะที่ 2 (A2) สามารถสร้างโมเดลเบื้องต้นแทนความหมายจาก  
เงื่อนไขที่กำหนดให้ได้

คุณลักษณะที่ 3 (A3) สามารถสร้างข้อสรุปจากโมเดลเบื้องต้นได้

คุณลักษณะที่ 4 (A4) สามารถสร้างโมเดลที่ 2 จากเงื่อนไขอื่น ๆ ได้

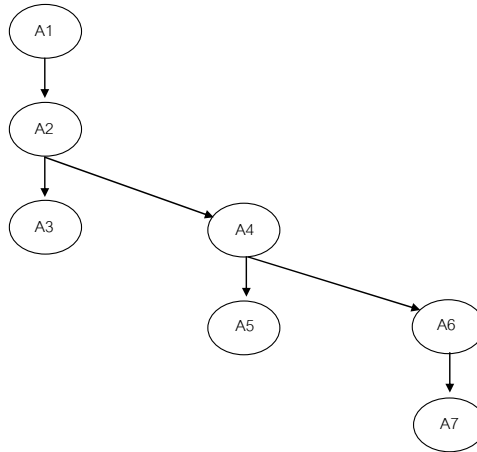
คุณลักษณะที่ 5 (A5) สามารถสร้างข้อสรุปจากโมเดลเบื้องต้นและ

โมเดลที่ 2 ได้

คุณลักษณะที่ 6 (A6) สามารถสร้างโมเดลที่ 3 ได้

คุณลักษณะที่ 7 (A7) สามารถสร้างข้อสรุปโดยการรวมทั้ง 3 โมเดลได้

แสดงเป็นแผนภาพลำดับชั้นของคุณลักษณะ ได้ดังนี้



แผนภาพที่ 3 ลำดับชั้นของคุณลักษณะในการใช้เหตุผลเชิงตรรกะตามทฤษฎีความสามารถทางสมอง

ข. กำหนดเมทริกซ์เพื่อสร้างแบบทดสอบ

จากลำดับชั้นของคุณลักษณะที่แสดงในแผนภาพที่ 2 กำหนดเป็น matrix (A) matrix (R) และ matrix (Q.) ได้ดังนี้

$$A_{Mental Model} = \begin{pmatrix} 0100000 \\ 0011000 \\ 0000000 \\ 0000110 \\ 0000000 \\ 0000001 \\ 0000000 \end{pmatrix}$$

$$R_{Mental Model} = \begin{pmatrix} 1111111 \\ 0111111 \\ 0010000 \\ 0001111 \\ 0000100 \\ 0000011 \end{pmatrix}$$

0000001

$$Q_{r \text{ Mental Model}} = \begin{pmatrix} 111111111111111 \\ 011111111111111 \\ 001010101010101 \\ 000111111111111 \\ 000001100110011 \\ 000000011111111 \\ 000000000001111 \end{pmatrix}$$

ค. สร้างแบบทดสอบ

จาก matrix (Q<sub>r</sub>) กำหนดให้สร้างแบบทดสอบอย่างน้อยจำนวน 15 ข้อ โดยมีคุณลักษณะตามที่กำหนดไว้ในแต่ละคอลัมน์ เช่นคอลัมน์ที่ 4 หรือ ข้อ 4 กำหนดให้ประกอบด้วยคุณลักษณะ A1 A2 และ A4 ในข้อที่ 4 จึงต้องเขียนข้อสอบที่ประกอบด้วยความสามารถดังนี้

- แปลความจากเงื่อนไขที่กำหนดให้ได้
- สร้างโมเดลเบื้องต้นแทนความหมายจากเงื่อนไขที่กำหนดให้ได้
- สร้างโมเดลที่ 2 จากเงื่อนไขอื่น ๆ ได้

ลักษณะของคำถามตามคุณลักษณะดังกล่าวเป็นเพียงคำถามที่ให้ผู้ตอบสร้างโมเดลจากเงื่อนไขที่กำหนดให้ได้โดยไม่ต้องสรุป

ง. กำหนดแบบแผนการตอบข้อสอบที่คาดหวัง

จาก matrix (Q<sub>r</sub>) ข้างต้น กำหนดแบบแผนการตอบข้อสอบที่คาดหวัง คะแนนรวมและคุณลักษณะของผู้สอบได้ ดังแสดงในตารางที่ 4

ตารางที่ 4 เมทริกซ์การตอบข้อสอบที่คาดหวัง คะแนนรวมและคุณลักษณะของผู้สอบสำหรับชุดของผู้สอบจำนวน 15 คน จากลำดับขั้นของคุณลักษณะในแผนภาพที่ 3

ผู้สอบ	เมทริกซ์การตอบข้อสอบที่คาดหวัง	คะแนนรวม	คุณลักษณะของผู้สอบ
1	100000000000000	1	1000000
2	110000000000000	2	1100000
3	111000000000000	3	1110000
4	110100000000000	3	1101000



5	111110000000000	5	1111000
6	110101000000000	4	1101100
7	111111100000000	7	1111100
8	110100010000000	4	1101010
9	111110011000000	7	1111010
10	110101010100000	6	1101110
11	111111111110000	11	1111110
12	110100010001000	5	1101011
13	111110011001100	9	1111011
14	110101010101010	8	1101111
15	111111111111111	15	1111111

จ. นำแบบทดสอบไปใช้และประมาณค่าความน่าจะเป็นในการตอบและ  
จำแนกแบบแผนการตอบข้อสอบ

### 3) ผลการศึกษา

ก. ผลการประมาณค่าพารามิเตอร์ของข้อสอบโดยใช้ โปรแกรม BILOG 3.11  
ปรากฏในตารางที่ 5

ตารางที่ 5 ค่าพารามิเตอร์ของข้อสอบจำนวน 15 ข้อ ซึ่งประมาณค่าจากการใช้แบบแผนการตอบ  
ข้อสอบที่คาดหวังในตารางที่ 4

ข้อที่	ค่าอำนาจจำแนก (a-Parameter)	ค่าความยาก (b-Parameter)
1	3.00	-2.25
2	4.00	-1.44
3	4.00	0.74
4	4.00	-0.60
5	4.00	0.74
6	0.50	1.00
7	1.20	1.70
8	1.20	0.15

9	5.00	1.00
10	0.70	1.60
11	4.00	1.73
12	0.70	1.60
13	2.00	1.80
14	2.00	2.40
15	3.00	2.34

ข. ผลการจำแนกแบบแผนการตอบข้อสอบสำหรับแบบแผนการตอบ (1110000000000000) โดยใช้วิธีการจำแนกเบื้องต้น ดังแสดงในตารางที่ 6

ตารางที่ 6 การจำแนกแบบแผนการตอบข้อสอบ (1110000000000000) โดยใช้วิธีการจำแนกเบื้องต้น

ค่าระดับ ความสามารถ (Theta)	ค่าความน่าจะเป็น $[P_{j\text{Expected}}(\theta)]$	จำนวนที่คลาดเคลื่อน		
		จากแบบแผน การตอบที่คาดหวัง (Slips)	เมทริกซ์การตอบ ข้อสอบที่คาดหวัง	คุณลักษณะ ของผู้สอบ
-3.01	0.0000	3	0000000000000000	0000000
-1.82	0.0000	2	1000000000000000	1000000
-1.07	0.0000	1	1100000000000000	1100000
-0.34	1.0000	0	1110000000000000	1110000
-0.34	0.0001	2	1101000000000000	1101000
0.83	0.0000	2	1111100000000000	1111000
-0.19	0.0001	3	1101010000000000	1101100
0.91	0.0000	4	1111111000000000	1111100

0.17	0.0001	3	110100010000000	1101010
1.18	0.0000	4	111110011000000	1111010
0.45	0.0000	5	110101010100000	1101110
1.88	0.0000	8	111111111110000	1111110
0.37	0.0000	4	110100010001000	1101011
1.51	0.0000	6	111110011001100	1111011
0.68	0.0000	7	110101010101010	1101111
3.03	0.0000	12	111111111111111	1111111

จากตารางที่ 6 แสดงให้เห็นว่าผู้สอบที่มีแบบแผนการตอบข้อสอบ 1110000000000000 มีค่าความน่าจะเป็นสูงสุดที่ระดับความสามารถเท่ากับ -0.34 เป็นผู้ซึ่งมีคุณลักษณะ A1 A2 และ A3 กล่าวคือ เป็นผู้ที่สามารถแปลความจากเงื่อนไขที่กำหนดให้ได้ (A1) สามารถสร้างโมเดลเบื้องต้นแทนความหมายจากเงื่อนไขที่กำหนดให้ได้ (A2) และสามารถสร้างข้อสรุปจากโมเดลเบื้องต้นได้ (A3)

ค. ผลการจำแนกแบบแผนการตอบข้อสอบสำหรับแบบแผนการตอบ (1010000000000000)

ผลการจำแนกแบบแผนการตอบข้อสอบ สำหรับแบบแผนการตอบ (1010000000000000) โดยใช้วิธีการจำแนกเบื้องต้นและการตรวจสอบการจำแนกเบื้องต้น ดังแสดงในตารางที่ 7 และตารางที่ 8 ตามลำดับ

ตารางที่ 7 การจำแนกแบบแผนการตอบข้อสอบ (1010000000000000) โดยใช้วิธีการจำแนกเบื้องต้น

ค่าระดับ ความสามารถ (Theta)	ค่าความน่าจะเป็น เป็น $[P_{jExpected}(\theta)]$	จำนวนที่ คลาดเคลื่อน จากแบบแผน การตอบที่ คาดหวัง(Slips)	เมทริกซ์การตอบ ข้อสอบที่คาดหวัง	คุณลักษณะ ของผู้สอบ
-3.01	0.0000	2	000000000000000	0000000
-1.82	0.0000	1	100000000000000	1000000
-1.07	0.0000	2	110000000000000	1100000
-0.34	0.0008	1	111000000000000	1110000
-0.34	0.0000	3	110100000000000	1101000

0.83	0.0000	3	1111100000000000	1111000
-0.19	0.0000	4	1101010000000000	1101100
0.91	0.0000	5	1111111000000000	1111100
0.17	0.0000	4	1101000100000000	1101010
1.18	0.0000	5	1111100110000000	1111010
0.45	0.0000	6	1101010101000000	1101110
1.88	0.0000	9	1111111111100000	1111110
0.37	0.0000	5	1101000100010000	1101011
1.51	0.0000	7	1111100110011000	1111011
0.68	0.0000	8	1101010101010100	1101111
3.03	0.0000	13	1111111111111111	1111111

จากตารางที่ 7 แสดงให้เห็นว่าค่าความน่าจะเป็นมีค่าต่ำมาก เมื่อพิจารณาค่าความน่าจะเป็นสูงสุดพบว่าแบบแผนการตอบข้อสอบ (1010000000000000) มีค่าความน่าจะเป็นในระดับความสามารถเท่ากับ -0.34 มากที่สุด แต่เมื่อพิจารณาคุณลักษณะของผู้สอบในระดับความสามารถดังกล่าวพบว่าไม่เป็นจริงตามลำดับขั้นของคุณลักษณะ กล่าวคือ ผู้สอบไม่สามารถสร้างโมเดลเบื้องต้นแทนความหมายจากเงื่อนไขที่กำหนดให้ได้ (คุณลักษณะ A2) ซึ่งไม่ควรจะสามารถสร้างข้อสรุปจากโมเดลเบื้องต้นได้ (คุณลักษณะ A3) เนื่องจากความสามารถสร้างโมเดลเบื้องต้นแทนความหมายจากเงื่อนไขที่กำหนดให้ได้เป็นคุณลักษณะที่ต้องมีมาก่อนความสามารถสร้างข้อสรุปจากโมเดลเบื้องต้นได้ การให้การจำแนกเบื้องต้นจึงไม่สามารถจำแนกแบบแผนการตอบข้อสอบ (1010000000000000) ได้ถูกต้อง

ตารางที่ 8 การจำแนกแบบแผนการตอบข้อสอบ (1010000000000000) โดยใช้การตรวจสอบการจำแนกเบื้องต้น

ค่าระดับ ความสามารถ (Theta)	ค่าความน่าจะเป็น $\left[ P_{j\text{Expected}}(\theta) \right]$	จำนวน ที่คลาดเคลื่อน จากแบบแผน การตอบ ที่คาดหวัง (Slips)	เมทริกซ์การตอบ ข้อสอบที่คาดหวัง	คุณลักษณะ ของผู้สอบ
-3.01	-	-	0000000000000000	0000000
-1.82	*	0	1000000000000000	1000000
-1.07	0.0941	1	1100000000000000	1100000

-0.34	0.0008	1	1110000000000000	1110000
-0.34	0.0001	2	1101000000000000	1101000
0.83	0.0000	3	1111100000000000	1111000
-0.19	0.0000	3	1101010000000000	1101100
0.91	0.0000	5	1111111000000000	1111100
0.17	0.0000	3	1101000100000000	1101010
1.18	0.0000	5	1111100110000000	1111010
0.45	0.0000	5	1101010101000000	1101110
1.88	0.0000	9	1111111111110000	1111110
0.37	0.0000	4	1101000100010000	1101011
1.51	0.0000	7	1111100110011000	1111011
0.68	0.0000	7	1101010101010100	1101111
3.03	0.0000	13	1111111111111111	1111111

จากตารางที่ 8 แสดงให้เห็นว่าผู้สอบมีระดับความสามารถมากกว่า -1.82 และมี ความสามารถแปลความจากเงื่อนไขที่กำหนดให้ได้ (คุณลักษณะ A1) สังเกตได้จากเครื่องหมาย \* ซึ่ง แสดงให้เห็นว่าผู้สอบมีคุณลักษณะตามแบบแผนการตอบข้อสอบที่คาดหวังนั้น แต่มีระดับ ความสามารถน้อยกว่า -1.07 ซึ่งไม่สามารถสร้างโมเดลเบื้องต้นแทนความหมายจากเงื่อนไข ที่กำหนดให้ได้ (คุณลักษณะ A2) แสดงว่าผู้สอบมีความรู้ความเข้าใจเรื่องการแปลความเชิงตรรกะ เพียงเล็กน้อยซึ่งไม่เพียงพอที่จะสร้างข้อสรุปที่ถูกต้องได้

### 3.2 การใช้โมเดลลำดับชั้นของคุณลักษณะในการกำหนดมาตรฐาน

Sadesky และ Gushta (2004) ได้ประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะใน การกำหนดมาตรฐานการสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ ซึ่งเป็นแบบทดสอบย่อยในโครงการบ่งชี้ ผลสัมฤทธิ์ของโรงเรียน (School Achievement Indicators Program: SAIP) จัดขึ้นในปี ค.ศ. 1977 เพื่อ แสดงให้เห็นว่าโมเดลลำดับชั้นของคุณลักษณะเป็นโมเดลที่มีประโยชน์ในการจำแนกความสามารถของ นักเรียนในแต่ละระดับมาตรฐานได้อย่างมีประสิทธิภาพ

แบบทดสอบที่ใช้มีทั้งหมด 125 ข้อ จำแนกระดับของข้อสอบเป็น 5 ระดับ ในแต่ละระดับมีจำนวนข้อสอบ 25 ข้อ จำแนกระดับการปฏิบัติจากโมเดลลำดับชั้นของคุณลักษณะ โดยใช้โปรแกรม Mathematica 4.1 ที่เขียนโดย Steve Hunka ในแต่ละระดับกำหนดแบบแผน

การตอบข้อสอบและจำนวนข้อสอบที่ตอบถูกขั้นต่ำ เช่น ในระดับที่ 2 ผู้สอบจะตอบข้อสอบถูกโดยมีจำนวนขั้นต่ำในแต่ละเนื้อหา ดังนี้ 1) 6 ใน 10 ข้อในเรื่องจำนวนและปฏิบัติการเกี่ยวกับจำนวน (number and operation) 2) 1 ใน 3 ข้อ ในเรื่องพีชคณิต (algebra) 3) 5 ใน 8 ข้อในเรื่องการวัดและเรขาคณิต (measurement and geometry) 4) 1 ใน 3 ในเรื่องการจัดการข้อมูลและสถิติ (data management and statistics) และ 5) 15 ใน 25 ข้อจากข้อสอบทั้งหมด

ผลจากการศึกษาพบว่าโมเดลลำดับขั้นของคุณลักษณะเป็นเทคนิคที่มีประโยชน์ในการกำหนดมาตรฐาน ดังนี้ 1) แสดงให้เห็นถึงความรู้หรือทักษะที่ผู้สอบมีในแต่ละระดับของการบรรลุมาตรฐานการปฏิบัติ 2) ลดข้อจำกัดในเรื่องความน่าเชื่อถือของคณะผู้ประเมินในการกำหนดค่าความน่าจะเป็นที่ผู้สอบจะตอบข้อสอบถูกในแต่ละระดับ และ 3) ลดข้อจำกัดในเรื่องการใช้คะแนนรวมเอกมิติ (unidimensional total score)

### **3.3 การศึกษาเชิงสมมติเพื่อประเมินวิธีการจำแนกในโมเดลลำดับขั้นของคุณลักษณะ**

Cui, Leighton และ Zheng (2006) ได้สมมติข้อมูลเพื่อประเมินวิธีการจำแนกที่ใช้ในโมเดลลำดับขั้นของคุณลักษณะจำนวน 2 วิธี ได้แก่ วิธี A และวิธี B ลักษณะของลำดับขั้นที่ใช้ในการสมมติข้อมูลมีจำนวน 4 ลักษณะ ได้แก่ ลำดับขั้นเชิงคู่เข้า ลำดับขั้นเชิงจำแนก ลำดับขั้นเชิงเส้นและลำดับขั้นแบบไม่มีโครงสร้าง แต่ละลักษณะมีจำนวน 7 คุณลักษณะ การประเมินจำแนกเป็น 2 ระดับ ได้แก่ ระดับแบบแผนคุณลักษณะ (the attribute patterns level) และระดับคุณลักษณะของแต่ละบุคคล (the individual attribute level) โดยจะแยกเปรียบเทียบแบบแผนจริงสำหรับแต่ละเวกเตอร์การตอบ ผลการวิจัยพบว่า วิธี B ให้ผลการประเมินดีกว่าวิธี A ทั้งระดับแบบแผนคุณลักษณะและระดับคุณลักษณะของบุคคลและพบว่าการประเมินระดับคุณลักษณะของบุคคลจากทั้ง 2 วิธี มีความสำคัญมากกว่าระดับแบบแผนคุณลักษณะ

### **3.4 การสร้างดัชนีเพื่อประเมินลำดับขั้นของคุณลักษณะ**

Cui และคณะ (2006) ได้สร้างดัชนีเพื่อประเมินความตรงของลำดับขั้นในการให้ข้อมูลย้อนกลับเชิงวินิจฉัยจากความถูกต้องและความเหมาะสมของลำดับขั้นของคุณลักษณะ โดยใช้สถิติความสอดคล้องกับบุคคล (person-fit statistic) เรียกว่า ดัชนีความคงที่ของลำดับขั้น (hierarchy consistency index: HCI) คำนวณโดยใช้สูตร ดังนี้

$$HCI_i = 1 - \frac{2 \sum_{j \in S_{correct_i}} \sum_{g \in S_j} X_{ij}(1 - X_{ij})}{N_{c_i}}$$

- เมื่อ
- $S_{correct_i}$  = ผลรวมจำนวนข้อสอบที่ผู้สอบตอบได้ถูกต้อง
  - $X_{ij}$  = คะแนนของผู้สอบคนที่  $i$  (0,1) สำหรับข้อที่  $j$
  - $S_j$  = ผลรวมข้อข้อที่เป็นส่วนหนึ่งของคุณลักษณะ (subset of attribute) ที่ต้องการวัดในข้อที่  $j$
  - $N_{c_i}$  = จำนวนรวมข้อที่เปรียบเทียบสำหรับข้อที่ตอบถูกต้องทั้งหมด โดยผู้สอบคนที่  $i$

ดัชนี HCI มีค่าตั้งแต่ -1 ถึง 1 ถ้ามีความสอดคล้องแบบสมบูรณ์ (fit perfectly) กล่าวคือเวกเตอร์การตอบของผู้สอบจับคู่แบบหนึ่งต่อหนึ่งกับแบบแผนการตอบข้อสอบที่คาดหวัง โดยไม่มีความผิดพลาด (slip จาก 0 เป็น 1 หรือ จาก 1 เป็น 0) ดัชนีจะมีค่าเท่ากับ 0 ถ้าดัชนีมีค่าเท่ากับ 1 แสดงว่าไม่มีความสอดคล้องแบบสมบูรณ์ (completely misfit) เช่น ผู้สอบตอบข้อสอบถูกต้อง แต่ตอบข้อสอบที่เป็นคุณลักษณะที่ต้องมีมาก่อนข้อนั้นผิด เป็นต้น ถ้ามีค่าดัชนีใกล้ -1 แสดงว่าผู้สอบใช้ทักษะทางพุทธิปัญญาที่แตกต่างจากลำดับขั้นของคุณลักษณะที่กำหนด สำหรับค่าความคลาดเคลื่อนมาตรฐานของดัชนี HCI ใช้เพื่อแสดงให้เห็นระดับความสอดคล้อง กล่าวคือ ถ้ามีค่าเฉลี่ยสูงแต่มีค่าความคลาดเคลื่อนมาตรฐานต่ำแสดงว่าเวกเตอร์การตอบมีความสอดคล้องกับโมเดลลำดับขั้นของคุณลักษณะ

การทดลองใช้ดัชนี HCI ในการประเมินลำดับขั้นของคุณลักษณะจากแบบทดสอบวัดผลสัมฤทธิ์มาตรฐานที่ใช้ในการทดสอบระดับกว้าง (large-scale standardized achievement test) ซึ่งประกอบด้วยข้อสอบคณิตศาสตร์จำนวน 54 ข้อ จำแนกคุณลักษณะได้ 8 คุณลักษณะ โดยการสมมติข้อมูลแบบแผนการตอบข้อสอบจำนวน 5,000 แบบแผน พบว่ามีแบบแผนการตอบข้อสอบจำนวน 4,923 แบบแผนที่สอดคล้องกับลำดับขั้นของคุณลักษณะ โดยมีค่าความสอดคล้องของโมเดลกับข้อมูลโดยรวม (overall model data fit) เท่ากับ 0.8230 แสดงว่าลำดับขั้นของคุณลักษณะมีความสอดคล้องกับข้อมูลดี

ข้อเสนอแนะในการวิจัยเกี่ยวกับดัชนี HCI มี 2 ประเด็น ประเด็นแรก เสนอแนะให้ทดลองใช้ดัชนี HCI กับลำดับขั้นของคุณลักษณะรูปแบบอื่น ๆ ประเด็นที่สอง เสนอแนะให้เพิ่มเงื่อนไขในการประเมิน เช่น จำนวนข้อสอบ จำนวนคุณลักษณะ โครงสร้างของลำดับขั้น และความซับซ้อนของลำดับขั้น เพื่อพัฒนาเกณฑ์ที่มียกทั่วไปสำหรับดัชนี HCI

### 3.5 การประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะในแบบทดสอบการอ่าน ภาษาต่างประเทศ

Wang, Gierl และ Leighton (2006) ได้ประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะเพื่อศึกษาลำดับชั้นของคุณลักษณะของข้อสอบที่เป็นแบบเลือกตอบในเรื่องความเข้าใจในการอ่านในแบบทดสอบภาษาอังกฤษระดับวิทยาลัย (College English Test) ซึ่งจัดโดยมณฑลในประเทศจีน และประเมินเชิงวิจักษ์จากลำดับชั้นของคุณลักษณะนั้น มีการดำเนินการ 2 ขั้นตอน ขั้นตอนแรกเป็นการวิเคราะห์เนื้อหา (substantive analysis) ขั้นตอนที่สองเป็นการวิเคราะห์ผลการตอบข้อสอบของนักเรียน การวิเคราะห์เนื้อหา ประกอบด้วย การกำหนดลักษณะเฉพาะของลำดับชั้นของคุณลักษณะ (specification of attribute hierarchy) ได้แก่ จำนวนคุณลักษณะทางพุทธิปัญญาในเรื่องความเข้าใจในการอ่านของนักเรียนที่เรียนภาษาที่สองและภาษาต่างประเทศ จำแนกความสัมพันธ์ระหว่างคุณลักษณะโดยการศึกษาเอกสารที่เกี่ยวข้องกับความเข้าใจในการอ่านภาษาที่สองและภาษาต่างประเทศได้จำนวน 8 คุณลักษณะ 2 รูปแบบลำดับชั้น ดังนี้

คุณลักษณะที่ใช้ในเรื่องความเข้าใจในการอ่านภาษาที่สองและภาษาต่างประเทศ ประกอบด้วย

1) ความรู้พื้นฐานทางภาษา (BA) เช่น การจำคำ (word recognition) และความรู้พื้นฐานในการสร้างประโยค (basic syntactic)

2) ความเข้าใจเนื้อหา รูปแบบและโครงสร้างของประโยค (US)

3) ความเข้าใจเนื้อหา รูปแบบและโครงสร้างของบทอ่านที่ยาวขึ้น (UT)

4) การวิเคราะห์วัตถุประสงค์ เป้าหมายและยุทธวิธีของผู้เขียน (PGS)

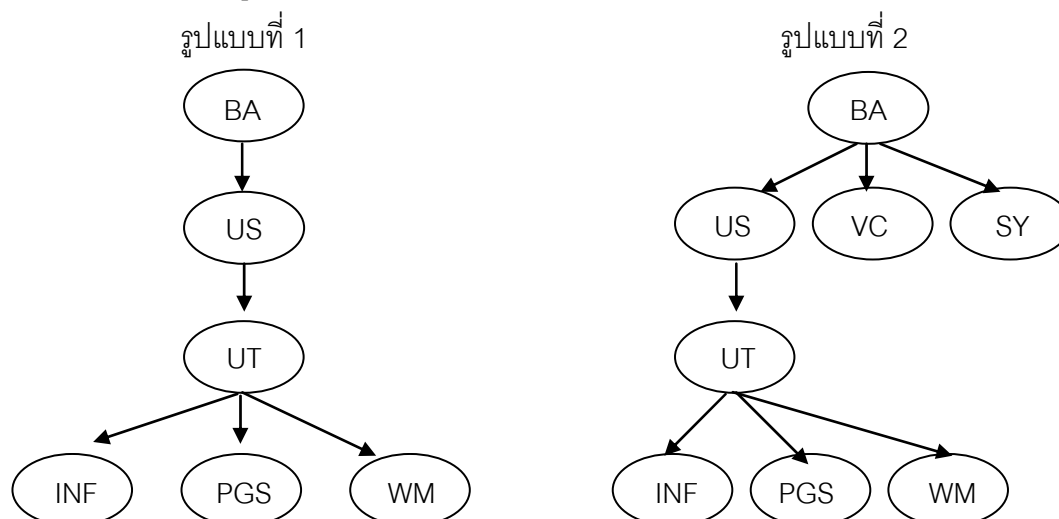
5) การกำหนดความหมายของคำในบริบท (WM)

6) การสรุปอ้างอิงโดยใช้พื้นฐานความรู้เดิม (INF)

7) ความเข้าใจบทอ่านที่ประกอบด้วยคำศัพท์ยาก (VC)

8) ความเข้าใจบทอ่านที่มีโครงสร้างของประโยคที่ซับซ้อน (SY)

มีรูปแบบลำดับชั้น ดังนี้





การกำหนดคุณลักษณะ (coding the attribute) ดำเนินการโดยให้ผู้เชี่ยวชาญที่มีประสบการณ์ด้านการสอนภาษาอังกฤษเป็นภาษาต่างประเทศและเป็นผู้ออกข้อสอบในการสมัครรับเข้า จำนวน 2 คน เป็นผู้กำหนดคุณลักษณะในข้อสอบแต่ละข้อ โดยกำหนดได้มากกว่า 1 คุณลักษณะในแต่ละข้อ มีข้อสอบจำนวน 20 ข้อ แต่นำมาใช้ในการกำหนดคุณลักษณะเพียง 16 ข้อ เนื่องจากข้อสอบจำนวน 4 ข้อเป็นข้อสอบที่มีอำนาจจำแนกเป็นลบ ได้จำนวนข้อสอบที่ผู้เชี่ยวชาญมีความคิดเห็นสอดคล้องกันจำนวน 9 ข้อ ซึ่งใช้ในการวิเคราะห์ผลการตอบข้อสอบของนักเรียน

การวิเคราะห์ผลการตอบข้อสอบของนักเรียน ดำเนินการโดยกำหนดแบบแผนการตอบข้อสอบที่คาดหวัง (expected examinee response patterns) สำหรับแต่ละรูปแบบลำดับขั้นของคุณลักษณะ รูปแบบที่ 1 ได้จำนวนแบบแผนการตอบข้อสอบที่คาดหวัง จำนวน 10 แบบแผน ส่วนรูปแบบที่ 2 ได้จำนวนแบบแผนการตอบข้อสอบที่คาดหวังจำนวน 37 แบบแผน แต่มีแบบแผนที่สอดคล้องกับลำดับขั้นของคุณลักษณะ จำนวน 25 แบบแผน ผลของการจำแนกแบบแผนการตอบข้อสอบที่สังเกตได้ พบว่า ทั้งสองรูปแบบสามารถจำแนกระดับความสามารถที่แตกต่างกันของผู้สอบได้ โดยสามารถแสดงให้เห็นถึงการรอบรู้หรือไม่รอบรู้ที่แตกต่างกันของผู้สอบที่มีระดับความสามารถเดียวกันหรือใกล้เคียงกัน สำหรับผลของการประเมินลำดับขั้นของคุณลักษณะโดยใช้ดัชนี HCI (hierarchy consistency index) พบว่า รูปแบบที่ 2 สะท้อนให้เห็นถึงโมเดลทางพุทธิพิสัยจากการตอบข้อสอบของผู้สอบที่ถูกต้องมากกว่ารูปแบบที่ 2 แสดงให้เห็นจากการมีค่าเฉลี่ยสูงกว่า แต่มีค่าความคลาดเคลื่อนมาตรฐานต่ำกว่า รูปแบบที่ 2 จึงสามารถให้สารสนเทศเชิงวินิจฉัยแก่ผู้สอบได้มากกว่า

ผลจากการวิจัยแสดงให้เห็นว่าโมเดลลำดับขั้นของคุณลักษณะให้ประโยชน์มากกว่าวิธีการให้คะแนนโดยการวัดทางจิตแบบดั้งเดิม (traditional psychometric scoring method) จำนวน 3 ประการ ประการแรกโมเดลลำดับขั้นของคุณลักษณะให้ผลการประเมินความสามารถในภาพรวมและสารสนเทศเชิงวินิจฉัยเกี่ยวกับความรอบรู้หรือไม่รอบรู้ของผู้สอบได้มากกว่าซึ่งสามารถใช้เป็นข้อมูลพื้นฐานเพื่อการปรับปรุงนักเรียนและการจัดการเรียนรู้อีก ประการที่สองโมเดลลำดับขั้นของคุณลักษณะให้สารสนเทศเกี่ยวกับความตรงเชิงโครงสร้าง โดยการวิเคราะห์เนื้อหาเพื่อให้ได้สารสนเทศเกี่ยวกับสิ่งที่ต้องการวัด ซึ่งชี้ให้เห็นระดับความตรงเชิงโครงสร้างได้โดยใช้ดัชนี HCI ประการที่สาม

โมเดลลำดับชั้นของคุณลักษณะเป็นกลไกในการให้ข้อมูลย้อนกลับเกี่ยวกับทฤษฎีทางพุทธิปัญญาในโมเดลเนื้อหา เช่น จากผลการวิจัยพบว่า รูปแบบที่ 2 มีความสอดคล้องกับข้อมูลเชิงประจักษ์มากกว่า แสดงให้เห็นว่าความเข้าใจบทอ่านที่ประกอบด้วยคำศัพท์ยาก (VC) และความเข้าใจบทอ่านที่มีโครงสร้างของประโยคที่ซับซ้อน (SY) เป็นคุณลักษณะทางพุทธิปัญญาที่สำคัญต่อความสามารถในการอ่านภาษาที่สองหรือภาษาต่างประเทศซึ่งไม่ควรจะละเลย

### **3.6 การประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะในการสืบสวนคุณลักษณะทางพุทธิปัญญาจากการทำแบบทดสอบย่อย เรื่องการอ่านเชิงวิพากษ์ ในการสอบ SAT**

Wang และ Gierl (2007) ได้ประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะในการทดสอบย่อย เรื่องการอ่านเชิงวิพากษ์ในการสอบ SAT เพื่อแสดงให้เห็นถึงประโยชน์ของโมเดลลำดับชั้นของคุณลักษณะในการเชื่อมโยงกันระหว่างผลการทำแบบทดสอบของนักเรียนกับทฤษฎีทางพุทธิปัญญา ซึ่งให้สารสนเทศการวินิจฉัยทางพุทธิปัญญาทั้งคะแนนสอบและแนวทางการพัฒนาการเรียนและการสอนที่เป็นไปได้ การวิจัยมีการดำเนินการ 2 ขั้นตอน ขั้นตอนที่ 1 เป็นการวิเคราะห์แบบทดสอบย่อย SAT เรื่องการอ่านเชิงวิพากษ์เพื่อกำหนดลำดับชั้นของคุณลักษณะจำแนกคุณลักษณะทางพุทธิปัญญาโดยนักเรียน ตรวจสอบความถูกต้องของลำดับชั้นของคุณลักษณะใช้การรายงานทางภาษาของนักเรียน (student verbal report) และดัชนี HCI (hierarchy consistency index) ขั้นตอนที่ 2 สร้างแบบแผนของคุณลักษณะและแบบแผนการตอบข้อสอบที่คาดหวัง คำนวณความน่าจะเป็นของคุณลักษณะสำหรับผู้สอบและข้อมูลย้อนกลับด้านพุทธิปัญญา ผลการวิจัยพบว่าโมเดลลำดับชั้นของคุณลักษณะให้ผลดีกว่าการให้คะแนนการวัดทางจิตแบบดั้งเดิม 3 ข้อ ข้อที่ 1 ให้ทั้งคะแนนรวมและข้อมูลย้อนกลับเชิงวินิจฉัยที่ช่วยในการพัฒนาการเรียนรู้นักเรียน ข้อที่ 2 ให้ข้อมูลที่เป็นความจริงโครงสร้างของแบบทดสอบ และข้อที่ 3 ให้ข้อมูลที่เป็นประโยชน์ต่อนักวิจัยในการพัฒนาทฤษฎีทางพุทธิปัญญา

### **3.7 การประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะในการประเมินทางวิศวกรรม โดยใช้โมเดลความเกี่ยวข้อง (Connectionist model) เพื่อประเมินแบบแผนการตอบข้อสอบ**

Gierl, Cui และ Hunka (2007) ได้ประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะในการประเมินทางวิศวกรรม ซึ่งเป็นนวัตกรรมใหม่ในการวัด เมื่อวิศวกรรมเป็นหลักการที่ใช้เพื่อออกแบบการวิเคราะห์ในการให้คะแนนและการรายงานผลการประเมิน การประเมินทางวิศวกรรมประกอบด้วยขั้นตอนสำคัญ 3 ขั้นตอน ขั้นตอนที่ 1 เป็นการสร้างโมเดลการปฏิบัติงานที่มีลักษณะเฉพาะ ขั้นตอนที่ 2 สร้างแบบแผน (templates) การปฏิบัติงานตามกรอบของโมเดลและ

ขั้นตอนที่ 3 ใช้โมเดลการวัดทางจิตในการเก็บรวบรวมข้อมูลการตอบของนักเรียนโดยใช้แผนแบบที่สร้างขึ้นในการให้คะแนน การวิจัยครั้งนี้มีเป้าหมายเพื่อบรรยายกระบวนการวิเคราะห์ในขั้นตอนการจำแนกแบบแผนการตอบของผู้สอบ โดยใช้ neural network ในการคำนวณความน่าจะเป็นของคุณลักษณะ กลุ่มตัวอย่าง ได้แก่ นักเรียนที่เขียนตอบแบบทดสอบพีชคณิต ในการสอบ SAT การดำเนินการวิจัยกำหนดการวิเคราะห์เป็น 2 ลักษณะ ลักษณะแรกกำหนดให้คำนวณเครือข่ายงานโดยไม่มีผลจากภายนอก ลักษณะที่ 2 กำหนดให้คำนวณเครือข่ายงานโดยมีข้อมูลภายนอกซึ่งเป็นการประมาณค่าความสามารถสำหรับแต่ละแบบแผนการตอบ ผลการวิจัยพบว่า ลักษณะที่ 2 มี root mean square น้อยกว่าลักษณะที่ 1 แสดงให้เห็นว่าข้อมูลจากภายนอกช่วยเพิ่มความสามารถในการอธิบายได้ของเครือข่ายการแก้ปัญหา (network solution)

### **3.8 การใช้ดัชนี HCI เพื่อประเมิน person fit สำหรับการประเมินเชิงวินิจฉัยทางพุทธิปัญญา**

Cui และ Leighton (2007) ได้ใช้ดัชนี HCI ในการตรวจสอบ misfitting item - response vectors สำหรับการประเมินเชิงวินิจฉัยทางพุทธิปัญญา การวิจัยใช้ข้อมูลเชิงสมมติเพื่อตรวจสอบอำนาจในการตรวจสอบของดัชนี HCI เมื่อประเภทของ misfitting item-response vectors แตกต่างกัน พบว่า ดัชนี HCI มีอำนาจในการตรวจสอบมากเมื่อข้อสอบมีอำนาจจำแนกมาก

### **3.9 การประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะและดัชนี HCI ในการทดสอบโมเดลทางพุทธิปัญญาที่ให้นักเรียนเป็นฐานและผู้เชี่ยวชาญเป็นฐาน**

Leighton, Cui และ Cor (2008) การประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะและดัชนี HCI ในการทดสอบโมเดลทางพุทธิปัญญาที่ให้นักเรียนเป็นฐานและผู้เชี่ยวชาญเป็นฐาน โดยมีวัตถุประสงค์เพื่อเปรียบเทียบความเหมาะสมของโมเดลทางพุทธิปัญญาที่สร้างจากนักเรียนและที่สร้างจากผู้เชี่ยวชาญในการทำนายการปฏิบัติงานของนักเรียนจากการทำข้อสอบพีชคณิต 1 และ 2 ในการสอบ SAT เดือนมีนาคม ปีค.ศ. 2005 ผลการวิจัย พบว่า โมเดลของนักเรียนทำนายการปฏิบัติงานของนักเรียนได้ดีสำหรับกลุ่มที่มีความสามารถปานกลางถึงสูง ส่วนโมเดลของนักเรียนทำนายการปฏิบัติงานของนักเรียนได้ดีสำหรับกลุ่มที่มีความสามารถปานกลางถึงสูงและกลุ่มที่มีความสามารถปานกลางหรือเท่ากับค่าเฉลี่ย แสดงให้เห็นว่า โมเดลของผู้เชี่ยวชาญทำนาย

กระบวนการตอบโดยทั่วไปได้มากกว่าเมื่อผู้สอบมีระดับความสามารถหลากหลาย ส่วนโมเดลของนักเรียนจะดีกว่าเมื่อผู้สอบมีระดับความสามารถเฉพาะ

### 3.10 การประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะในจำแนกและตีความแตกต่างของทักษะทางพุทธิปัญญาระหว่างกลุ่ม

Gierl, Zheng และ Cui (2008) ได้ประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะในการจำแนกและตีความความแตกต่างทางพุทธิปัญญาระหว่างกลุ่ม กลุ่มตัวอย่างเป็นนักเรียนที่เขียนตอบแบบทดสอบพีชคณิต ในการสอบ SAT ในเดือนมีนาคม ปี ค.ศ.2005 จำนวน 3016 คน เพศหญิงและเพศชาย เพศละ 1508 คน ขั้นตอนการทดสอบการทำหน้าที่ต่างกันของระดับคุณลักษณะจำแนกเป็น 4 ขั้นตอน ขั้นตอนที่ 1 ตั้งสมมติฐานเฉพาะของฟังก์ชันการทำหน้าที่ต่างกันของระดับคุณลักษณะ (Attribute-level differential functioning) ขั้นตอนที่ 2 ประเมินค่าความน่าจะเป็นสำหรับการศึกษาคุณลักษณะ ขั้นตอนที่ 3 ให้คำจำกัดความการจับคู่คุณลักษณะและขั้นตอนที่ 4 ทดสอบฟังก์ชันการทำหน้าที่ต่างกันของระดับคุณลักษณะ กำหนดสมมติฐานการทดสอบจากลำดับชั้นย่อยของแผนภาพลำดับชั้นของคุณลักษณะในการปฏิบัติงานพีชคณิต ซึ่งมีคุณลักษณะที่ 1 เป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะที่ 6 และคุณลักษณะที่ 6 เป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะที่ 7 จำแนกสมมติฐานเป็น 2 ข้อ ข้อแรกเป็นการทดสอบความแตกต่างระหว่างกลุ่มสำหรับคุณลักษณะที่ 6 เรื่อง สมการเชิงเส้นตรง (linear equation) และข้อที่ 2 เป็นการทดสอบความแตกต่างระหว่างกลุ่มสำหรับคุณลักษณะที่ 7 เรื่อง สมการกำลังสอง (quadratic equation) ผลการวิจัยพบว่า ไม่มีความแตกต่างระหว่างเพศสำหรับคุณลักษณะที่ 1 และคุณลักษณะที่ 6 แต่มีความแตกต่างระหว่างเพศ สำหรับคุณลักษณะที่ 7 แสดงให้เห็นว่า เพศหญิงและเพศชายมีทักษะพื้นฐานทางคณิตศาสตร์สามารถเทียบเคียงกันได้และสามารถนำไปดำเนินการแก้สมการเชิงเส้นตรงได้เหมือนกัน แต่ต่างกันที่การนำไปแก้สมการกำลังสอง

#### ตอนที่ 3 การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

##### มโนทัศน์เบื้องต้นเกี่ยวกับการทดสอบแบบปรับเหมาะ

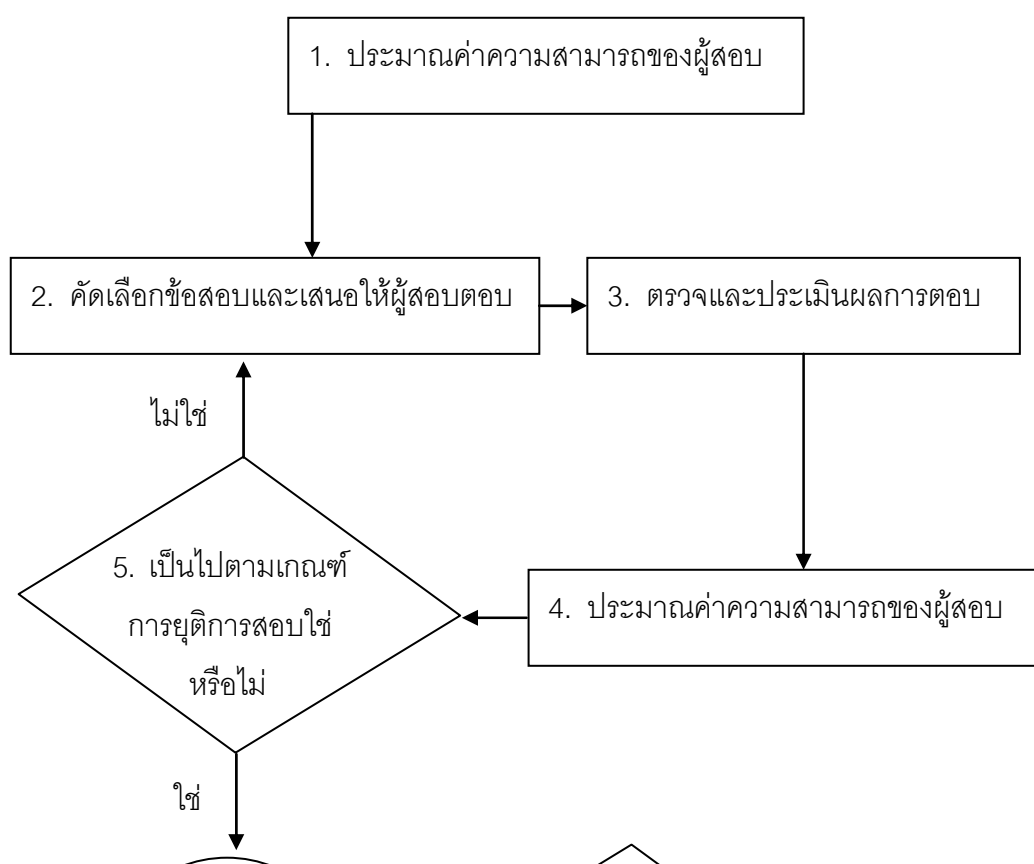
การทดสอบแบบปรับเหมาะ (adaptive testing or tailored testing) เป็นการทดสอบที่มุ่งวัดคุณลักษณะที่ต้องการ โดยการคัดเลือกข้อสอบตามความสามารถของผู้สอบในระหว่างกระบวนการทดสอบ การคัดเลือกข้อสอบจะพิจารณาจากผลการตอบข้อสอบหรือข้อคำถามที่มีมาก่อน ถ้าตอบข้อสอบข้อนั้นได้ถูกต้อง ข้อสอบข้อต่อไปก็จะยากขึ้นแต่ถ้าตอบข้อสอบข้อนั้นผิด

ข้อสอบข้อต่อไปก็จะง่ายลง (Weiss,1983; Weiss and Schleisman, 1999) แบบทดสอบที่ใช้ในการทดสอบแบบปรับเหมาะจึงเป็นแบบทดสอบที่ประกอบไปด้วยชุดของข้อสอบที่แตกต่างกันสำหรับการสอบของแต่ละบุคคลซึ่งเป็นลักษณะที่ตรงกันข้ามกับแบบทดสอบแบบประเพณีนิยม (conventional tests) ที่ผู้สอบทุกคนจะถูกสอบด้วยแบบทดสอบชุดเดียวกัน Weiss (1988,1990)

แบบทดสอบแบบปรับเหมาะฉบับแรก ได้แก่ แบบทดสอบเชาว์ปัญญาของบินเนท์ (Binet intelligence test) พัฒนาขึ้นโดย Alfred Binet นักจิตวิทยาชาวฝรั่งเศส แบบทดสอบเชาว์ปัญญาของบินเนท์ใช้การเลือกชุดของข้อสอบเบื้องต้นตามระดับอายุ (ความยาก) ข้อสอบจะถูกให้คะแนนและถูกเลือกให้ผู้สอบทำตามทางเลือกที่กำหนดไว้คงที่ (fixed branching) โดยผู้บริหารการสอบ เมื่อผู้สอบทำข้อสอบในระดับอายุจริงได้ไม่ถูกต้องก็จะถูกให้ทำข้อสอบในระดับอายุที่ต่ำกว่า แต่ถ้าทำได้ถูกต้องก็จะถูกให้ทำข้อสอบในระดับอายุที่สูงกว่า จนกระทั่งถึงระดับเพดาน (Ceiling level) เมื่อผู้สอบทำข้อสอบทุกข้อไม่ถูกต้องหรือระดับพื้นฐาน (basal level) เมื่อผู้สอบทำข้อสอบทุกข้อได้ถูกต้อง เกณฑ์ที่ใช้ในการยุติการสอบจะแตกต่างกันตามระดับความสามารถของแต่ละบุคคล เมื่อสามารถระบุระดับเพดานและระดับพื้นฐานของผู้สอบได้ (Weiss, 1983; Weiss, 1988,1990) แม้ว่าแบบทดสอบเชาว์ปัญญาของบินเนท์จะถูกนำไปใช้อย่างแพร่หลายและเป็นต้นฉบับในการกำหนดหลักการของการทดสอบแบบปรับเหมาะในปัจจุบัน แต่แบบทดสอบเชาว์ปัญญาของบินเนท์ก็มีข้อจำกัดที่ผู้สอบอาจแสดงผลการตอบไม่คงเส้นคงวาระหว่างการทดสอบ ผู้สอบบางคนอาจทำข้อสอบบางข้อได้ดีแตกต่างกันเนื่องจากอิทธิพลของเชื้อชาติ เพศ วัฒนธรรม รวมทั้งยังมีค่าใช้จ่ายในการทดสอบสูงและใช้เวลาในการสอบมาก (ศิริชัย กาญจนวาสี, 2538, 2545) แบบทดสอบแบบปรับเหมาะอื่น ๆ เริ่มพัฒนาขึ้นปี ค.ศ. 1950 โดยเริ่มตั้งแต่การเขียนตอบลงบนกระดาษคำตอบ (paper-and-pencil) จนถึงการใช้คอมพิวเตอร์ช่วยในการทดสอบ (computer assisted testing) ซึ่งมีการพัฒนาวิธีเลือกข้อสอบเริ่มต้นและกำหนดเกณฑ์ในการยุติการสอบที่แตกต่างกัน แต่ยังคงใช้หลักการที่เน้นความเหมาะสมกับความสามารถของผู้สอบเช่นเดียวกับแบบทดสอบเชาว์ปัญญาของบินเนท์ และมักจะเกี่ยวข้องกับคำว่า เหมาะสม (tailored) ความไม่แน่นอนในการตอบ (response-contingent) โปรแกรม (programmed) คอมพิวเตอร์ (computerized) อัตโนมัติ (automated) แต่ละบุคคล (individualized) ทางแยก (branched) และลำดับขั้นของการทดสอบ (sequential testing) (Weiss, 1988, 1990)

การดำเนินการทดสอบแบบปรับเหมาะจะเกี่ยวข้องกับคำถามสำคัญ 3 คำถาม ได้แก่ จะเลือกข้อสอบที่ใช้ในการเริ่มต้นการสอบได้อย่างไร (How to START) จะเลือกข้อสอบข้อถัดไปได้อย่างไร (How to CONTINUE) และรู้ได้อย่างไรว่าจะยุติการสอบ (How to STOP) (Thissen

and Mislevy, 1990; Wainer, 1990) คำถามดังกล่าวจะไม่พบในการทดสอบแบบประเพณีนิยม เนื่องจากในการเริ่มต้นทำแบบทดสอบมักจะเริ่มจากการทำข้อสอบข้อที่ 1 ข้อต่อ ๆ ไป จนถึงข้อสุดท้าย แต่ในการทดสอบแบบปรับเหมาะจะการเลือกข้อสอบที่ใช้ในการประมาณค่าความสามารถเบื้องต้นหรือเป็นขั้นตอนของการเริ่มการทดสอบเป็นที่สำคัญและมีผลต่อการประมาณค่าความสามารถของผู้สอบเพื่อการเลือกข้อสอบข้อต่อไป รวมถึงการยุติการสอบและการประมาณค่าความสามารถสุดท้ายของผู้สอบ การเริ่มต้นการสอบ การดำเนินการสอบและการยุติการสอบจึงเป็นสิ่งที่ต้องพิจารณาเพื่อให้การทดสอบแบบปรับเหมาะมีประสิทธิภาพสูงสุดและมักจะเป็นประเด็นที่ใช้ในการศึกษาวิจัยพัฒนาเกี่ยวกับการทดสอบแบบปรับเหมาะ โดยทั่วไปการทดสอบแบบปรับเหมาะมักดำเนินการโดยเริ่มจากการประมาณความสามารถของผู้สอบ คัดเลือกข้อสอบที่เหมาะสมกับความสามารถของผู้สอบและเสนอให้ผู้สอบทำ หลังจากนั้นจึงประมาณค่าความสามารถของผู้สอบอีกครั้งหนึ่ง ตรวจสอบและประเมินคำตอบ พิจารณาว่าเป็นไปตามเกณฑ์ที่ใช้ในการยุติการสอบใช่หรือไม่ ถ้าไม่ใช่ก็ดำเนินการคัดเลือกข้อสอบข้อต่อไป ถ้าใช่ก็สิ้นสุดการสอบสำหรับแบบทดสอบฉบับนั้นและพิจารณาว่าเสร็จสิ้นการทดสอบทุกฉบับในแบบทดสอบชุดนั้นแล้วใช่หรือไม่ ถ้าไม่ใช่ก็ดำเนินการสอบแบบทดสอบฉบับต่อไป ถ้าใช่ก็ยุติการสอบ (Thissen and Mislevy, 1990) ดังแสดงในแผนภูมิที่ 1



## แผนภูมิที่ 1 ขั้นตอนการทดสอบแบบปรับเหมาะ

### ประเภทของการทดสอบแบบปรับเหมาะ

การจำแนกประเภทของการทดสอบแบบปรับเหมาะพิจารณาได้จากเครื่องมือที่ใช้ในการทดสอบและยุทธวิธีที่ใช้ในการทดสอบ (ศิริชัย กาญจนวาสี, 2538, 2545) มีรายละเอียด ดังนี้

#### 1. การทดสอบแบบปรับเหมาะจำแนกตามเครื่องมือที่ใช้

การทดสอบแบบปรับเหมาะจำแนกตามเครื่องมือที่ใช้ได้ 2 ประเภท ประเภทแรกเป็นการทดสอบแบบปรับเหมาะโดยไม่ใช้คอมพิวเตอร์และประเภทที่ 2 เป็นการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ การทดสอบแบบปรับเหมาะโดยไม่ใช้คอมพิวเตอร์ ประกอบด้วย แบบทดสอบเซวาน์ปัญญาของบิเนท์ แบบทดสอบเลือกคำตอบเป็นกลุ่มแบบสองขั้นตอน (Two-Stage Test) แบบทดสอบปรับระดับความสามารถที่ยืดหยุ่น (Flexilevel Test) และแบบทดสอบกำหนดตาม

การแยกทาง (Branching Test) ส่วนการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ซึ่งเป็นการนำคอมพิวเตอร์มาใช้เพื่อความสะดวกในการคัดเลือกข้อสอบและประมาณค่าความสามารถของผู้สอบ เริ่มพัฒนาขึ้นประมาณปี ค.ศ. 1970 (Weiss, 1983)

## 2. การทดสอบแบบปรับเหมาะจำแนกตามยุทธวิธีที่ใช้ในการทดสอบ

การทดสอบแบบปรับเหมาะจำแนกตามยุทธวิธีที่ใช้ในการทดสอบได้ 2 ประเภท ประเภทแรกเป็นการทดสอบแบบปรับเหมาะโดยใช้ยุทธวิธีสองขั้นตอน (Two-stage Strategies) และประเภทที่ 2 เป็นการทดสอบแบบปรับเหมาะโดยใช้ยุทธวิธีหลายขั้นตอน (Multi-stage Strategies) ประกอบด้วยยุทธวิธีหลายขั้นตอนแบบทางแยกคงที่ (Fixed-Branching) และยุทธวิธีหลายขั้นตอนแบบทางแยกแปรผัน (Variable Branching)

### ยุทธวิธีที่ใช้ในการทดสอบแบบปรับเหมาะ

ยุทธวิธีที่ใช้ในการทดสอบแบบปรับเหมาะประกอบด้วย ยุทธวิธีสองขั้นตอน (Two-stage Strategies) และยุทธวิธีหลายขั้นตอน (Multi-stage Strategies) มีรายละเอียด ดังนี้

#### 1. ยุทธวิธีสองขั้นตอน (Two-stage Strategies)

ในปี ค.ศ. 1970 และ ปี ค.ศ. 1980 นักวิจัยได้มีความพยายามในการพัฒนาแบบทดสอบแบบปรับเหมาะโดยเน้นไปที่ขั้นตอนการคัดเลือกข้อสอบ เพื่อให้ได้จำนวนข้อสอบที่น้อยที่สุดในการประมาณค่าความสามารถของผู้ตอบได้อย่างถูกต้อง ยุทธวิธีสองขั้นตอนจึงได้รับการพัฒนาขึ้นเพื่อตอบสนองวัตถุประสงค์ดังกล่าว โดยใช้แบบทดสอบสองฉบับในการทดสอบสองขั้นตอน แบบทดสอบแต่ละฉบับในแต่ละขั้นตอนจะประกอบไปด้วยกลุ่มของข้อสอบที่มุ่งวัดเนื้อหาใดเนื้อหาหนึ่งซึ่งกำหนดทิศทางคำตอบสำหรับผู้สอบไว้ เรียกว่า testlets (Thissen and Mislevy, 1990) แบบทดสอบในแต่ละขั้นตอนมีรายละเอียด ดังนี้

##### 1.1 แบบทดสอบขั้นตอนแรกหรือแบบทดสอบเพื่อกำหนดทิศทาง (first-stage test or routing test)

แบบทดสอบขั้นตอนแรกหรือแบบทดสอบเพื่อกำหนดทิศทางเป็นแบบทดสอบที่ค่อนข้างสั้นประกอบด้วยข้อสอบที่มีระดับความยากกระจายในช่วงกว้างหรือมีระดับความยากง่ายปานกลาง เพื่อให้ผู้สอบทำในขั้นตอนแรกและตรวจข้อสอบทันทีเพื่อใช้ผลการทดสอบเป็นเกณฑ์ในการคัดเลือกแบบทดสอบในขั้นที่สองโดยพิจารณาตามค่าความยากง่ายของแบบทดสอบ



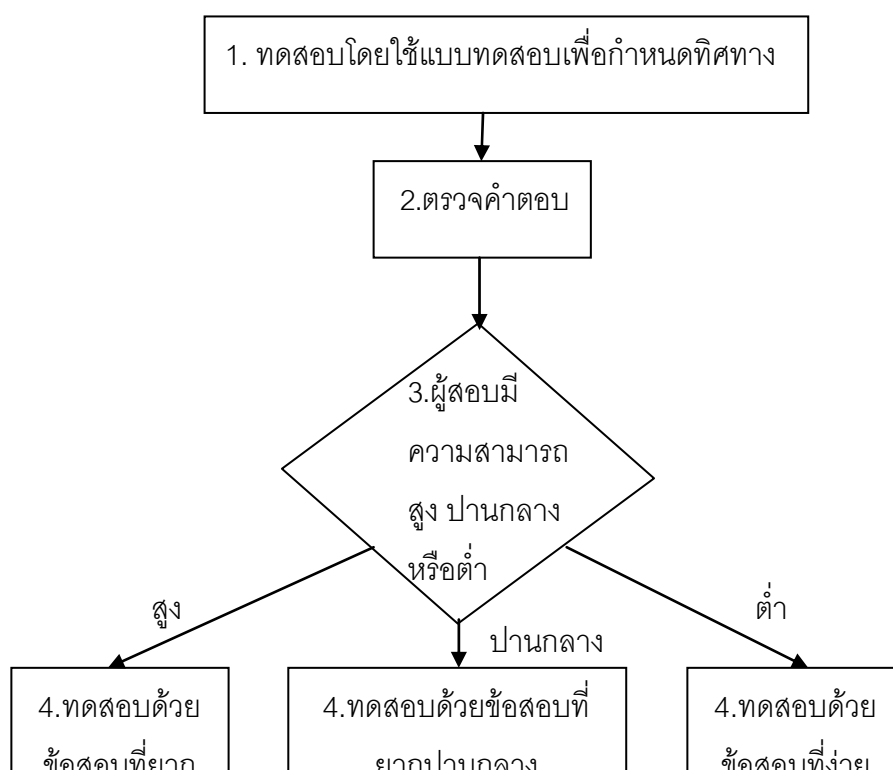
ที่เหมาะสมกับระดับความสามารถของผู้สอบซึ่งประมาณได้จากการทดสอบแบบทดสอบฉบับแรก (นันทิยา พึ่งคำ, 2531; ศิริชัย กาญจนวาสี, 2538, 2545)

## 1.2 แบบทดสอบขั้นตอนที่สอง แบบทดสอบหลักหรือแบบทดสอบวัดผล (second-stage test, main test or measurement test)

แบบทดสอบขั้นตอนที่สอง แบบทดสอบหลักหรือแบบทดสอบวัดผลเป็นแบบทดสอบที่ถูกคัดเลือกให้เหมาะสมกับความสามารถของผู้สอบซึ่งประมาณได้จากการทดสอบในขั้นแรกมักเป็นแบบทดสอบที่ยาวกว่าแบบทดสอบในขั้นแรกและประกอบด้วยแบบทดสอบย่อยประมาณ 3-5 ฉบับ แยกตามระดับความยาก ในแต่ละฉบับจะมีข้อสอบจำนวน 20-30 ข้อ (นันทิยา พึ่งคำ, 2531; สายชล อบทม, 2539) คะแนนที่ได้จากแบบทดสอบแต่ละฉบับไม่สามารถเปรียบเทียบกันได้โดยตรง เนื่องจากข้อสอบแต่ละฉบับมีความยากง่ายต่างกันวิธีนำคะแนนมาเปรียบเทียบกันอาจทำได้โดยการนำคะแนนของแต่ละฉบับมาถ่วงด้วยค่าความยากเฉลี่ยของแต่ละฉบับเพื่อปรับค่าความยากให้เท่ากัน (ศิริชัย กาญจนวาสี, 2538, 2545)

การคัดเลือกข้อสอบและจัดชุดของข้อสอบในการทดสอบทั้งสองขั้นตอนดังกล่าว ศิริชัย กาญจนวาสี (2538, 2545) ได้เสนอไว้ว่าสามารถนำทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) มาประยุกต์ใช้ได้เพื่อให้การประมาณค่าความสามารถของผู้สอบไม่ขึ้นกับระดับความยากของแบบทดสอบ

ขั้นตอนการทดสอบโดยใช้ยุทธวิธีสองขั้นตอน แสดงได้ในแผนภูมิที่ 2



## แผนภูมิตี่ 2 ขั้นตอนการทดสอบแบบปรับเหมาะโดยใช้ยุทธวิธีสองขั้นตอน

สำหรับข้อดีและข้อจำกัดของการทดสอบโดยใช้ยุทธวิธีสองขั้นตอน ศิริชัย กาญจนวาสี (2538, 2545) ได้กล่าวไว้ว่า มีข้อดีที่ช่วยลดจำนวนข้อสอบที่แต่ละคนจะต้องทำ โดยไม่ทำให้ความแม่นยำในการวัดลดลง แต่มีข้อจำกัดกรณีทดสอบกับคนจำนวนมากในเวลาเดียวกัน การตรวจข้อสอบกำหนดทิศทางจะต้องรีบเร่งตรวจและใช้คนจำนวนมาก รวมทั้งการปรับเหมาะของกระบวนการทดสอบเกิดขึ้นเพียงครั้งเดียวหลังจากการทดสอบในขั้นแรก ซึ่งถ้าเกิดความคลาดเคลื่อนในขั้นนี้อาจทำให้เกิดความคลาดเคลื่อนในการการทดสอบในขั้นที่สองได้ ความคลาดเคลื่อนในการจำแนกกลุ่มความสามารถของผู้สอบอาจเกิดขึ้นได้ประมาณร้อยละ 20 ตลอดจนแบบทดสอบที่มีช่วงค่าความยากที่แคบจะทำให้ได้สารสนเทศค่อนข้างน้อย แต่ถ้าสร้างแบบทดสอบวัดผลในขั้นที่สองให้มีช่วงความยากที่ค่อนข้างกว้างก็จะขัดแย้งกับหลักการของการทดสอบแบบปรับเหมาะ ทั้งนี้ได้เสนอแนะการแก้ปัญหาดังกล่าวไว้ว่าสามารถดำเนินการได้โดยใช้คอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ

## 2. ยุทธวิธีหลายขั้นตอน (Multi-stage Strategies)

การทดสอบแบบปรับเหมาะโดยใช้ยุทธวิธีหลายขั้นตอนเป็นการใช้การทดสอบที่มากกว่า 2 ขั้นตอน ในการกำหนดทางแยก (branch) ระหว่างการทดสอบ จำแนกได้เป็น รูปแบบการทดสอบแบบทางแยกคงที่และรูปแบบการทดสอบแบบทางแยกแปรผัน (Hambleton and Swaminathan, 1985) มีรายละเอียดดังนี้

### **รูปแบบทางแยกคงที่ (Fixed-branching model)**

รูปแบบทางแยกคงที่ที่ใช้ในยุทธวิธีหลายขั้นตอนเหมือนยุทธวิธีสองขั้นตอน ในส่วนของกำหนัดทางแยกที่คงที่ แต่ต่างกันในส่วนของการตัดสินใจแยกทางซึ่งในการใช้

ยุทธวิธีสองขั้นตอนกำหนดการตัดสินใจเพียงครั้งเดียวจากการทำแบบทดสอบกำหนดทิศทางในขั้นแรกไปการทำแบบทดสอบวัดผลในขั้นที่สอง ส่วนรูปแบบทางแยกคงที่ในการใช้ยุทธวิธีหลายขั้นตอนจะกำหนดการตัดสินใจหลังจากที่ผู้สอบทำข้อสอบในแต่ละข้อตามโครงสร้างของแบบทดสอบที่กำหนดไว้ล่วงหน้า โดยคำนวณค่าพารามิเตอร์ของข้อสอบไว้แล้ว ผู้สอบแต่ละคนจะได้ทำแบบทดสอบที่มีโครงสร้างของแบบทดสอบเดียวกันแต่จะได้ทำข้อสอบในข้อที่แตกต่างกันตามกฎการคัดเลือกข้อสอบที่กำหนดไว้ รูปแบบในการกำหนดโครงสร้างของแบบทดสอบ สรุปได้ดังนี้

1) รูปแบบปิรามิดหรือรูปแบบโครงสร้างต้นไม้ (Pyramidal model) เป็นรูปแบบที่มีการวิจัยและการนำไปใช้จำนวนมาก มีการจัดโครงสร้างของข้อสอบเหมือนรูปทรงของปิรามิดหรือต้นไม้ โดยมีรูปแบบการกำหนดค่าความยาก วิธีการแยกทางและจำนวนข้อสอบที่แตกต่างกันประกอบด้วย แบบขนาดความยากคงที่ (constant step size) แบบขนาดความยากแปรผัน (variable step size) แบบปิรามิดข้างตัด (truncated pyramid) แบบปิรามิดที่มีหลายข้อในแต่ละชั้น (multiple pyramids) และแบบปิรามิดที่ให้ค่าน้ำหนักตัวเลือกเพื่อแยกทาง (differential response option branching) การทดสอบจะดำเนินการตามโครงสร้างของแบบทดสอบและข้อกำหนดในการแยกทางที่กำหนดไว้

2) รูปแบบปรับระดับแบบยืดหยุ่นหรือรูปแบบเฟล็กซ์เลเวล (Flexilevel model) เป็นรูปแบบที่ Frederic Lord พัฒนาขึ้นเมื่อปี ค.ศ. 1971 ในแบบทดสอบชื่อ self-scoring flexilevel test (Thissen and Mislevy, 1990) ซึ่งเป็นแบบทดสอบที่ประกอบด้วยข้อสอบที่มีช่วงห่างของค่าความยากเท่ากัน โดยจัดเรียงข้อสอบจากข้อที่ง่ายที่สุดไปข้อที่ยากที่สุด การทดสอบจะเริ่มจากการให้ผู้สอบทำข้อสอบที่มีค่าความยากปานกลาง ถ้าผู้สอบตอบข้อสอบได้ถูกต้อง ข้อสอบข้อต่อไปจะยากขึ้น แต่ถ้าผู้สอบตอบข้อสอบได้ไม่ถูกต้องข้อต่อไปจะเป็นข้อที่ง่ายลง การทดสอบจะดำเนินการไปจนกระทั่งผู้สอบทำแบบทดสอบได้จำนวนครึ่งหนึ่งของแบบทดสอบก็จะยุติการทดสอบ ในระยะแรกที่มีการพัฒนา ดำเนินการทดสอบโดยใช้การตอบบนกระดาษคำตอบ (paper-and-pencil) การเตรียมแบบทดสอบดำเนินการโดยการเรียงข้อสอบจากข้อที่ง่ายที่สุดไปข้อที่ยากที่สุดและแบ่งข้อสอบออกเป็นสองกลุ่ม ได้แก่ กลุ่มข้อสอบง่ายและกลุ่มข้อสอบยาก โดยมีข้อสอบที่มีค่าความยากง่ายปานกลางจำนวนหนึ่งข้อ ข้อสอบในแต่ละกลุ่มจะถูกจัดพิมพ์แยกในแต่ละส่วนของหน้ากระดาษ (กลุ่มข้อสอบง่ายอยู่ด้านซ้าย กลุ่มข้อสอบยากอยู่ด้านขวา) ในการทดสอบจะทำการทำสัญลักษณ์ (marking bubbles) สำหรับข้อที่ตอบถูกและข้อที่ตอบผิดแตกต่างกันโดยการขีด (scratch-off) ข้อที่ตอบถูกจะทำเครื่องหมายจุดด้วยสีน้ำเงิน ส่วนข้อที่ตอบผิดจะทำเครื่องหมายจุดด้วยสีแดง การทดสอบเริ่มจากการให้ผู้สอบทำข้อสอบที่มีความ

ยากง่ายปานกลางถ้าตอบผิดก็เลือกข้อสอบข้อต่อไปในกลุ่มของข้อสอบง่ายซึ่งอยู่ด้านซ้ายของกระดาษ ถ้าตอบข้อดังกล่าวถูกต้องก็ทำเครื่องหมายจุดสีน้ำเงินและเลือกข้อสอบข้อต่อไปในกลุ่มข้อสอบยากซึ่งอยู่ด้านขวาของกระดาษ แต่ถ้าตอบผิดก็ทำเครื่องหมายจุดสีแดงในข้อนั้นและเลือกข้อสอบข้อต่อไปที่อยู่ในกลุ่มข้อสอบง่ายซึ่งอยู่ด้านซ้ายของกระดาษ ดำเนินการจนกระทั่งทำข้อสอบได้ครึ่งหนึ่งของแบบทดสอบจึงยุติการสอบ

3) รูปแบบปรับระดับขั้น (Stradaptive model) เป็นรูปแบบที่ David J. Weiss พัฒนาขึ้นในปี ค.ศ. 1973 ในแบบทดสอบชื่อ Stratified adaptive (stradaptive) test โดยมีแนวคิดว่าคุณคนมีความสามารถที่แปรผันได้อย่างเป็นอิสระในการทำแบบทดสอบฉบับหนึ่ง ๆ ความคงเส้นคงวาในการทำแบบทดสอบ (consistency) จึงเป็นสิ่งสำคัญที่ควรพิจารณา (Weiss, 1983) แบบทดสอบแบบปรับระดับขั้นพัฒนามาจากแบบทดสอบแบบปิรามิด แต่แตกต่างกันที่แบบทดสอบแบบปิรามิดมีข้อสอบที่มีค่าความยากระดับเดียว ส่วนแบบทดสอบแบบปรับระดับขั้นมีข้อสอบที่มีค่าความยากหลายระดับ (Hambleton and Swaminathan, 1985) ค่าความยากของข้อสอบจะถูกแบ่งเป็นระดับขั้น (stratified) แต่ระดับขั้นประกอบด้วยข้อสอบที่มีค่าความยากอยู่ในช่วงที่กำหนดเดียวกัน ในการทดสอบมีการดำเนินการคล้ายกับการทดสอบโดยใช้แบบทดสอบเซวาน์บัญญัติของบิเนท์ โดยผู้สอบจะเริ่มทำข้อสอบข้อแรกที่แตกต่างกันตามระดับความสามารถเบื้องต้น หลังจากนั้นก็จะประมาณค่าความสามารถของผู้สอบจากข้อสอบข้อแรกเพื่อคัดเลือกข้อสอบข้อต่อไป ถ้าผู้สอบตอบข้อสอบข้อแรกถูกข้อสอบข้อต่อไปก็จะเป็นข้อที่ยากขึ้น แต่ถ้าตอบผิดข้อต่อไปจะเป็นข้อที่ง่ายลง การทดสอบจะดำเนินการไปจนกระทั่งผู้สอบทำข้อสอบในระดับนั้นได้ไม่ถูกต้องทุกข้อหรือจนถึงระดับเพดาน (ceiling stratum) หรือทำคะแนนได้ไม่เกินคะแนนที่ได้จากการเดา (chance score)

### **รูปแบบทางแยกแปรผัน (Variable-branching model)**

รูปแบบทางแยกแปรผันเป็นรูปแบบการทดสอบที่ไม่ได้กำหนดโครงสร้างและไม่ได้กำหนดข้อสอบไว้ล่วงหน้า ว่าถ้าผู้สอบตอบถูกจะต้องไปทำข้อสอบข้อใดหรือถ้าตอบผิดจะต้องไปทำข้อสอบข้อใดแต่จะดำเนินการโดยใช้ข้อสอบที่กำหนดค่าฟังก์ชันสารสนเทศของข้อสอบไว้แล้ว การคัดเลือกข้อสอบเพื่อให้เหมาะสมกับความสามารถของผู้สอบจะดำเนินการโดยใช้วิธีการทางสถิติในการประมาณค่าความสามารถซึ่งวิธีการประมาณค่าความสามารถที่นิยมใช้มี 2 วิธี ได้แก่ การประมาณค่าแบบเบย์ (Bayesian Estimation) และการประมาณค่าแบบความเป็นไปได้สูงสุด (Maximum Likelihood Estimation)

## การให้คะแนนการทดสอบแบบปรับเหมาะ

Weiss (1974 อ้างถึงใน สุพัฒน์ สุขมลสันต์, 2536) ได้กล่าวถึงวิธีการให้คะแนนการทดสอบแบบปรับเหมาะไว้หลายวิธี สรุปได้ดังนี้

1. ให้คะแนนตามค่าความยากของข้อสอบที่ยากที่สุดที่ตอบถูก (The most difficult of items answered correctly) เป็นการให้คะแนนโดยดูจากข้อสอบทั้งหมดที่ผู้สอบสามารถตอบได้ถูกต้อง หากข้อใดที่มีค่าความยากสูงสุดที่ผู้สอบตอบถูก ถือว่าเป็นค่าความสามารถของผู้สอบคนนั้น ค่าความยากดังกล่าวควรเป็นค่าความยากที่ได้จากการวิเคราะห์ข้อทดสอบตามทฤษฎีการตอบสนองของข้อสอบ

2. ให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ทำถูก (average difficulty of all items answered correctly) เป็นการให้คะแนนโดยการนำค่าความยากของข้อสอบทุกข้อที่ผู้สอบตอบถูก มาหาค่าเฉลี่ย ค่าเฉลี่ยความยากนี้ถือเป็นค่าความสามารถของผู้สอบคนนั้น

3. ให้คะแนนตามค่าเฉลี่ยความยากของทุกข้อที่ทำ (average difficulty) วิธีการให้คะแนนแบบนี้ Lord เป็นผู้เสนอ ในปี ค.ศ. 1970 โดยการนำค่าความยากของข้อสอบทุกข้อที่ผู้สอบได้ตอบ โดยไม่คำนึงว่าจะตอบถูกหรือไม่มาหาค่าเฉลี่ย ค่าเฉลี่ยนี้ถือเป็นค่าความสามารถของผู้สอบ

4. ให้คะแนนตามค่าความยากของข้อสุดท้ายที่ตอบ (difficulty of the final item) นั่นคือ ไม่คำนึงว่าขั้นสุดท้ายจะตอบถูกหรือไม่ก็ตาม จะถือว่าค่าความยากของข้อสอบข้อสุดท้ายที่ผู้สอบตอบคือค่าความสามารถของผู้สอบผู้นั้น

5. ให้คะแนนตามค่าความยากของข้อสอบที่ต่อจากขั้นสุดท้ายที่ตอบ (difficulty of (N+1)<sup>th</sup> item) ซึ่งมีลักษณะคล้ายกับวิธีที่ 4 แต่เป็นการตรวจสอบความสามารถต่อไปอีกขั้นหนึ่ง กล่าวคือ ถ้าขั้นสุดท้ายที่ผู้สอบสามารถตอบถูก คะแนนในขั้นต่อไปก็จะเพิ่มขึ้น แต่ถ้าขั้นสุดท้ายตอบผิดคะแนนขั้นต่อไปก็จะลดลง

6. ให้คะแนนตามกลวิธีของเบย์ (Bayesian strategies) ซึ่งมีหลายวิธีแต่ที่นิยมใช้กันมากคือ วิธีที่เรียกว่า “วิธีของเบย์ที่ปรับปรุงให้ทันสมัย” (Bayesian updating) ซึ่งสามารถประมาณค่าความสามารถของผู้สอบได้ค่อนข้างคงที่

7. ให้คะแนนโดยวิธีการประมาณค่าความเป็นไปได้สูงสุด (maximum likelihood) ซึ่งมีหลายวิธีแต่ที่นิยมใช้มากคือ วิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข (conditional maximum likelihood) การประมาณค่าวิธีนี้มีข้อจำกัดกรณีที่ผู้สอบตอบข้อสอบถูกต้องหมดหรือผิดหมดจะไม่สามารถประมาณค่าได้ ขั้นตอนการให้คะแนนมี ดังนี้ (Hambleton and Swaminathan, 1985 อ้างถึงใน สุพัฒน์ สุขมลสันต์)

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

$$\theta_0 = \ln[r_a / (k - r_a)]$$

เมื่อ  $r_a = \sum a_1 U_1$ ;  $(U_1 = 1)$  เมื่อตอบถูก  
 $(U_1 = 0)$  เมื่อตอบผิด  
 $a_1$  คือ อำนาจจำแนกของข้อสอบข้อที่  $i$   
 $K$  คือ จำนวนข้อทดสอบทั้งหมดที่ผู้สอบตอบ

ขั้นที่ 2 หาค่า  $P_1(\theta_m)$  และ  $Q_1(\theta_m)$  โดยใช้สูตร ดังนี้

$$P_1(\theta_m) = c_1 + (1 - c_1) \frac{e^{Da_1(\theta_m - b_1)}}{1 + e^{Da_1(\theta_m - b_1)}}$$

$$Q_1(\theta_m) = 1 - P_1(\theta_m)$$

$$= \frac{1 - c_1}{1 + e^{Da_1(\theta_m - b_1)}}$$

เมื่อ  $\theta_m$  คือ ความสามารถของผู้สอบที่ประมาณค่าได้ในครั้งที่  $m$

$a_1$  คือ ค่าอำนาจจำแนกของข้อทดสอบข้อที่  $i$

$b_1$  คือ ค่าความยากของข้อสอบข้อที่  $i$

$c_1$  คือ ค่าการเดาของข้อสอบข้อที่  $t$

$D$  คือ ค่าคงที่ ซึ่งมีค่าเท่ากับ 1.7

$e$  คือ ค่าคงที่ ซึ่งมีค่าเท่ากับ 2.7182

ขั้นที่ 3 หาค่าปรับแก้ ( $h_m$ ) โดยใช้สูตร

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

ขั้นที่ 4 ประมวลค่าความสามารถของผู้สอบใหม่ ( $\theta_m + 1$ ) โดยใช้สูตร

$$\theta_m + 1 = \theta_m - h_m$$

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

### การประยุกต์ใช้คอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ

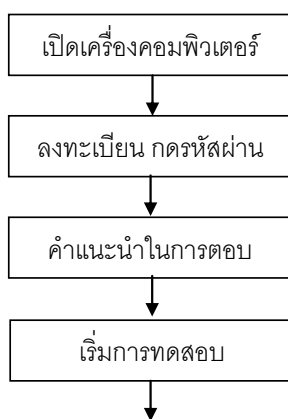
การนำคอมพิวเตอร์มาใช้ในการทดสอบ (Computer Assisted Testing) เพื่อเพิ่มประสิทธิภาพในการตรวจข้อสอบ การวิเคราะห์ข้อสอบ การตัดเกรด การรายงานผลการสอบและการสร้างคลังข้อสอบมีการดำเนินการอย่างแพร่หลายระหว่างปี ค.ศ. 1970 ถึง ปี ค.ศ. 1980 (ศิริชัย กาญจนวาสี, 2545) สำหรับการทดสอบแบบปรับเหมาะเริ่มมีการประยุกต์คอมพิวเตอร์

มาใช้ในโครงการวิจัยการทดสอบแบบปรับเหมาะโดยได้รับการสนับสนุนจาก U.S. Office of Naval Research ตั้งแต่ปลายปี ค.ศ. 1960 ถึงต้นปี ค.ศ. 1970 (Weiss, 1983, 2004)

คอมพิวเตอร์เป็นเทคโนโลยีที่มีหน่วยความจำขนาดใหญ่สามารถเก็บรวบรวมข้อมูลทั้งตัวหนังสือและรูปภาพได้จำนวนมาก สามารถประมวลผลได้อย่างรวดเร็ว รวมทั้งมีปฏิสัมพันธ์กับผู้สอบได้อย่างมีประสิทธิภาพ คอมพิวเตอร์จึงมีความเหมาะสมและประโยชน์ยิ่งต่อการนำมาประยุกต์ใช้ในทดสอบแบบปรับเหมาะที่ต้องมีระบบการส่งผ่านซึ่งสามารถนำข้อสอบออกจากคลังข้อสอบและนำเสนอให้ผู้สอบตอบได้อย่างรวดเร็ว โดยต้องมีหลักเกณฑ์ในการคัดเลือกข้อสอบเริ่มต้นการประมาณค่าความสามารถของผู้สอบเพื่อคัดเลือกข้อสอบระหว่างกระบวนการทดสอบและการยุติการสอบที่ชัดเจน (van der Linden, 1999)

การประยุกต์ใช้คอมพิวเตอร์ในการทดสอบแบบปรับเหมาะเกิดขึ้นในทุกกระบวนการของการทดสอบตั้งแต่การสร้างคลังข้อสอบ การวิเคราะห์สารสนเทศของข้อสอบ การคัดเลือกข้อสอบ การประมาณค่าความสามารถของผู้สอบ การยุติการสอบ ข้อสอบจะถูกเก็บไว้ในคอมพิวเตอร์และแสดงออกทางหน้าจอ (monitor) เพื่อให้ผู้สอบตอบข้อสอบโดยผ่านทางแป้นพิมพ์ (keyboard) หรือ เมาส์ (mouse) โปรแกรมคอมพิวเตอร์จะทำหน้าที่เสมือนนักจิตวิทยาที่ดำเนินการสอบโดยใช้แบบทดสอบเซาว์ปัญญาของบิเนตซึ่งจะช่วยกำหนดข้อสอบเริ่มต้น ข้อสอบที่เหมาะสมกับความสามารถของผู้สอบในระหว่างกระบวนการทดสอบ และตัดสินใจยุติการสอบ (Weiss, 2004)

ขั้นตอนในการดำเนินการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์จะเริ่มจากการลงทะเบียนสอบและการกรอกรหัสผ่านบนแป้นพิมพ์ของคอมพิวเตอร์ ก่อนเริ่มการทดสอบคอมพิวเตอร์จะแนะนำวิธีการตอบข้อสอบ มีการชักจูงและทดลองการตอบ เมื่อผู้สอบมีความเข้าใจดีแล้ว การสอบก็จะเริ่มขึ้น โดยการสุ่มข้อสอบข้อแรก (initial item) จากคลังข้อสอบซึ่งตามหลักการทั่วไปจะเป็นข้อสอบที่มีความยากปานกลางหรือเป็นข้อสอบที่คัดเลือกให้เหมาะสมกับผู้สอบตามสารสนเทศที่ได้รับ หลังจากตรวจข้อสอบจะมีการประมาณค่าความสามารถเบื้องต้นของผู้สอบ ถ้าตอบถูกข้อถัดไปจะยากขึ้น แต่ถ้าตอบผิดข้อถัดไปจะง่ายลง กระบวนการทดสอบจะดำเนินการเช่นนี้ซ้ำไปเรื่อย ๆ จนกว่าจะถึงเกณฑ์การยุติการสอบที่กำหนดไว้ โปรแกรมจะยุติการทำงานบันทึกผลการสอบ รายงานผลความสามารถที่แท้จริงของผู้สอบบนหน้าจอคอมพิวเตอร์และพิมพ์รายงานไปยังผู้ที่เกี่ยวข้อง (ศิริชัย กาญจนวาสี, 2538, 2545) ดังแสดงในแผนภูมิที่ 3



แผนภูมิที่ 3 ขั้นตอนการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์  
องค์ประกอบของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

องค์ประกอบของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ มีรายละเอียด ดังนี้

1. จุดเริ่มต้นการทดสอบ (starting point)

การทดสอบควรเริ่มต้นด้วยข้อสอบที่ตรงกับความสามารถของผู้สอบหรือใกล้เคียงกับ  
ความสามารถของผู้สอบมากที่สุด ซึ่งจะช่วยให้จำนวนข้อสอบที่ใช้ในการทดสอบสำหรับแต่ละ  
บุคคลมีจำนวนลดลง (นันทิยา พึ่งคำ, 2531) การเลือกข้อสอบข้อแรกอาจพิจารณาจากผลสัมฤทธิ์



ที่ผ่านมาของผู้สอบ (Hambleton, zaal และ Pieters (2000 อ้างถึงใน Latu and Chapman, 2002) โดยทั่วไปนิยมเลือกข้อสอบที่มีความยากปานกลาง โดยอาจดำเนินการได้ใน 2 กรณี กรณีแรก กลุ่มผู้สอบมีความสามารถใกล้เคียงกัน (homogeneous) หรือไม่มีข้อมูลสำหรับผลสัมฤทธิ์ที่ผ่าน มาควรเริ่มต้นด้วยข้อสอบที่มีค่าความยากปานกลางเท่ากันสำหรับทุกคน กรณีที่สอง ผู้สอบมี ความสามารถแตกต่างกัน (heterogeneous) หรือมีข้อมูลเกี่ยวกับการศึกษาของผู้สอบบ้าง ควรใช้ ค่าความยากปานกลางที่เหมาะสมสำหรับผู้สอบแต่ละกลุ่มความสามารถ โดยพิจารณาจากข้อมูล ที่มีอยู่ เช่น กลุ่มอายุ ระดับชั้นที่ศึกษา เป็นต้น (ศิริชัย กาญจนวาสี, 2545)

ความสำคัญของจุดเริ่มต้นยังคงเป็นที่โต้แย้งกันในหมู่นักวิจัย บางคนกล่าวว่า จุดเริ่มต้นไม่มีผลกระทบต่อ การสอบถ้ามีจำนวนข้อสอบมากพอ เช่น Lord (1977) อ้างถึงใน ศิริชัย, 2545) ได้ทำการทดลองด้วยสถานการณ์จำลอง โดยเริ่มต้นด้วยข้อสอบที่มีค่าความยาก ระดับต่าง ๆ พบว่าเกือบไม่มีผลต่อความถูกต้องแม่นยำในการประมาณค่าความสามารถของ ผู้สอบเลย โดยเฉพาะกรณีที่มีข้อสอบในคลังข้อสอบจำนวนมาก แต่บางคนกล่าวว่าจุดเริ่มต้นที่ไม่ เหมาะสมจะทำให้ผู้สอบเกิดความวิตกกังวลและรู้สึกผิดหวังซึ่งมีผลกระทบต่อ การทำแบบทดสอบ (Hambleton, zaal และ Pieters 2000 อ้างถึงใน Latu and Chapman, 2002)

## 2. การคัดเลือกข้อสอบและการประมาณค่าความสามารถ (item selection and ability estimate)

การคัดเลือกข้อสอบและการประมาณค่าความสามารถของผู้สอบมีประสิทธิภาพมาก ขึ้นเมื่อนำทฤษฎีการตอบสนองข้อสอบ (item response theory) มาประยุกต์ใช้ โดยทั่วไปนิยมใช้ วิธีที่สอดคล้องกัน เพื่อความสะดวกรวดเร็วในการคำนวณ วิธีที่นิยมนำมาใช้ในการประมาณค่า ความสามารถมี 2 วิธี ได้แก่ วิธีการประมาณค่าความเป็นไปได้สูงสุด (Maximum Likelihood Estimate) และการประมาณค่าแบบเบย์ (Bayesian Estimate) วิธีการคัดเลือกข้อสอบที่ใช้สำหรับ การประมาณค่าความเป็นไปได้สูงสุด ประกอบด้วย การคัดเลือกข้อสอบที่มีค่าความยากสอดคล้อง กับค่าความสามารถที่ประมาณได้ การคัดเลือกข้อสอบที่มีตำแหน่งสารสนเทศ (ค่าความยาก ค่าอำนาจจำแนกและค่าการเดา) สูงสุดสอดคล้องกับความสามารถที่ประมาณได้และการคัดเลือก ข้อสอบที่ให้สารสนเทศสูงสุดที่ตำแหน่งความสามารถนั้น ส่วนวิธีการคัดเลือกข้อสอบที่ใช้สำหรับ การประมาณค่าแบบเบย์ใช้การคัดเลือกข้อสอบข้อที่ยังไม่ได้นำมาใช้ในการทดสอบนั้นที่จะให้ค่า ความแปรปรวนของค่าความสามารถที่คาดหวังมีค่าต่ำสุด (นันทิยา พึ่งคำ, 2531; สายชล อบทม, 2539; ศิริชัย กาญจนวาสี, 2545)

### 3. เกณฑ์การยุติการสอบ (termination criteria or Stopping criteria)

ปัจจุบันมีการพัฒนาวิธีที่ใช้ในการยุติการสอบหลายวิธี แต่วิธีที่นิยมใช้มี 2 วิธี วิธีแรก คือ การทดสอบจนกระทั่งได้ระดับความคงที่ในการวัด (level of measurement stability or consistency) โดยการกำหนดระดับความคลาดเคลื่อนที่ยอมรับได้ วิธีที่สอง คือ การกำหนดจำนวนข้อสอบให้คงที่ วิธีการกำหนดความยาวของแบบสอบให้คงที่ช่วยให้แน่ใจได้ว่าผู้สอบได้ทำข้อสอบในจำนวนข้อสอบที่เท่ากันและช่วยให้แบบทดสอบไม่ยาวเกินไป แต่ให้คุณภาพของการวัดผลที่แตกต่างกัน เนื่องจากมีระดับความยากของข้อสอบแตกต่างกัน การใช้การยุติการสอบโดยพิจารณาจากค่าความคลาดเคลื่อนในการวัดจึงเป็นวิธีที่ดีกว่า (ศิริชัย กาญจนวาสี, 2545; Hambleton, zaal และ Pieters 2000 อ้างถึงใน Latu and Chapman, 2002)

### 4. จำนวนและเนื้อหาของข้อสอบในคลังข้อสอบ (Size and content of item banks)

ความตรงของแบบทดสอบขึ้นอยู่กับความเกี่ยวข้องและความเหมาะสมของข้อสอบที่ใช้ แต่สำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ความตรงของแบบทดสอบจะขึ้นอยู่กับทั้งจำนวนและเนื้อหาของข้อสอบในคลังข้อสอบ นอกจากข้อสอบจะต้องมีจำนวนมากเพียงพอในการประมาณค่าระดับความสามารถของผู้สอบซึ่งมีความแตกต่างหลากหลายกันได้อย่างถูกต้องเหมาะสมแล้ว ข้อควรคำนึงที่สำคัญประการหนึ่ง ก็คือ เนื้อหาของข้อสอบ เนื่องจากการมีข้อสอบในคลังจำนวนมากทำให้มีโอกาสที่จะมีข้อสอบที่ไม่ดีได้มาก ซึ่งข้อสอบเหล่านั้นจะมีผลกระทบอย่างมากต่อการให้คะแนนสุดท้ายในการทดสอบแบบปรับเหมาะ เพราะในการทดสอบแบบปรับเหมาะมีการตรวจให้คะแนนทันที ข้อสอบที่ถูกเลือกให้ผู้สอบทำจึงมีผลต่อความถูกต้องในการประมาณค่าความสามารถของผู้สอบและการให้คะแนนสุดท้ายโดยไม่มีโอกาสแก้ไขหรือย้อนกลับได้ (Hambleton, zaal และ Pieters 2000 อ้างถึงใน Latu and Chapman, 2002)

#### ประโยชน์ของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

Green (1983 อ้างถึงใน Wainer, 1990) และ Latu และ Chapman (2002) ได้กล่าวถึงประโยชน์ของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ไว้ สรุปได้ดังนี้

1. ช่วยเพิ่มประสิทธิภาพในการทดสอบ โดยเฉพาะภาระงานการทดสอบต่าง ๆ ได้แก่ การดำเนินการทดสอบ การกำหนดเกณฑ์และการให้คะแนน
2. แบบทดสอบที่เก็บไว้ในคลังข้อสอบในคอมพิวเตอร์จะมีความปลอดภัย (security) ต่อการคัดลอกมากกว่า

3. ไม่มีข้อจำกัดเรื่องเวลา ผู้สอบสามารถทำแบบทดสอบในเวลาที่แตกต่างกันตามความพร้อมของแต่ละบุคคลภายในช่วงเวลาที่กำหนด
4. ผู้สอบไม่เกิดความท้อแท้ในการทำข้อสอบ เนื่องจากการคัดเลือกข้อสอบที่เหมาะสมกับความสามารถของแต่ละบุคคล
5. การตรวจข้อสอบมีความชัดเจนขึ้น เนื่องจากไม่มีปัญหาที่เกิดจากกระดาษคำตอบ เช่น การทำเครื่องหมายไม่ชัดเจน ความไม่ชัดเจนในข้อที่เลือกเนื่องจากการลบเมื่อต้องการเปลี่ยนคำตอบ เป็นต้น
6. ข้อสอบสามารถตรวจให้คะแนนและให้ผลย้อนกลับ (feedback) กับผู้สอบได้ทันที
7. สามารถคัดเลือกข้อสอบจากการตอบข้อสอบเบื้องต้นได้อย่างรวดเร็ว
8. ผู้สอบได้ทำข้อสอบที่มีความยาวเหมาะสมกับระดับความสามารถ
9. มีการเสนอข้อสอบให้ผู้สอบตอบในรูปแบบที่หลากหลายทั้งภาพนิ่ง ภาพเคลื่อนไหว และเสียงประกอบ

### **ข้อจำกัดของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์**

Latu และ Chapman (2002) กล่าวถึงข้อจำกัดของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ไว้ ดังนี้

1. ความปลอดภัยของแบบทดสอบ (Test security) แม้ว่าจะมีนักวิจัยหลาย ๆ คน แสดงให้เห็นว่าการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มีความปลอดภัยของแบบทดสอบมากกว่าการทำแบบทดสอบบนโต๊ะ (desk drawer) แต่ก็มีนักวิจัยบางท่านโต้แย้งว่าปัจจุบันมีวิธีการขโมยข้อสอบได้หลายวิธี เช่น pager กล้องขนาดเล็ก (miniature cameras) กล้องวิดีโอ (video transmitters) เครื่องบันทึกวิดีโอขนาดเล็ก (micro video recorders) นอกจากนี้ผู้สอบสามารถคัดลอกข้อสอบจากเครื่องคอมพิวเตอร์ได้ง่ายและรวดเร็วกว่าการคัดลอกข้อสอบจากกระดาษคำตอบอีกด้วย

2. ปัญหาเกี่ยวกับผู้สอบ (Examinee issues) แม้ว่าการทดสอบแบบปรับเหมาะจะมีประโยชน์ต่อผู้สอบมาก แต่ยังคงมีข้อจำกัดเกี่ยวกับผู้สอบ เช่น ผู้สอบหลายคนไม่เคยชินกับการใช้คอมพิวเตอร์ ทำให้เกิดความวิตกกังวล ผู้สอบบางคนแสดงให้เห็นว่าการนำเสนอโดยใช้คอมพิวเตอร์เป็นสิ่งที่ยุ่งยาก (difficult or fatiguing) โดยเฉพาะการตอบโดยใช้เมาส์หรือคีย์บอร์ด

3. ค่าใช้จ่ายในการดำเนินการ (Financial realities) การทดสอบแบบปรับเหมาะมีค่าใช้จ่ายในการดำเนินการสร้างแบบทดสอบการเตรียมการและการดำเนินการสอบที่ค่อนข้างสูง เช่น Educational Testing Service (ETS) ศึกษาค่าใช้จ่ายในการทดสอบแบบปรับเหมาะโดยใช้

คอมพิวเตอร์พบว่าค่าใช้จ่ายที่ต้องจ่ายจริงมากกว่าวิธีการทดสอบแบบประเมินนิยม ดังจะเห็นได้จากในสหรัฐอเมริกาผู้สอบจะต้องเสียค่าใช้จ่ายมากกว่า 100 ดอลลาร์ เพื่อสอบ TOFEL ในระบบการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ค่าใช้จ่ายที่แพงมากดังกล่าวมาจากการพัฒนาคลังข้อสอบ การบริหารการทดสอบและการพัฒนาโปรแกรมสำเร็จรูป ซึ่งมีค่าใช้จ่ายในการพัฒนาและการนำไปใช้ในระบบที่สมบูรณ์ระหว่าง 3,000 -5,000 ดอลลาร์

### งานวิจัยที่เกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ในประเทศไทย

จิราพร ไกรสรศิวเวท (2529) ศึกษาความสัมพันธ์ระหว่างรูปแบบบางประการของแบบทดสอบรูปปิรามิดกับความสามารถทางการเรียน ในวิชาคณิตศาสตร์ โดยมีวัตถุประสงค์ที่สำคัญ 3 ประการ ประการแรก เพื่อศึกษาความสัมพันธ์ระหว่างผลการสอบที่ใช้แบบทดสอบรูปปิรามิดจำนวน 3 ฉบับ คือ แบบทดสอบรูปปิรามิดที่ใช้ขนาดชั้นคงที่ แบบที่ใช้ขนาดชั้นแปรผัน และแบบข้างตัดกับความสามารถทางการเรียนของนักเรียน ประการที่ 2 เพื่อศึกษาความสัมพันธ์ระหว่างวิธีการตรวจให้คะแนน 3 วิธี คือ วิธีการตรวจให้คะแนนตามค่าความยากเฉลี่ยของข้อสอบที่ทำ วิธีการตรวจให้คะแนนตามค่าเฉลี่ยของข้อสอบที่ตอบถูกและวิธีการตรวจให้คะแนนข้อสอบทุกข้อ ประการที่ 3 เพื่อศึกษาเปรียบเทียบค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบรูปปิรามิดแต่ละฉบับกับความสามารถทางการเรียนและเปรียบเทียบค่าสัมประสิทธิ์สหสัมพันธ์ภายในระหว่างวิธีการตรวจให้คะแนนทั้ง 3 วิธี กลุ่มตัวอย่าง ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 2 ปีการศึกษา 2528 สังกัดกรมสามัญศึกษา จังหวัดกำแพงเพชร จำนวน 1,740 คน ได้จากการสุ่มกลุ่มตัวอย่างโดยใช้การสุ่มอย่างง่าย ผลการวิจัยพบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผลการสอบที่ใช้รูปแบบปิรามิดแต่ละฉบับซึ่งตรวจให้คะแนนในแต่ละวิธีกับความสามารถทางการเรียนมีค่าเป็นบวกอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และ .05 เมื่อทดสอบระหว่างคู่พบว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบรูปปิรามิดที่ใช้ขนาดชั้นคงที่และแบบที่ใช้ขนาดชั้นแปรผันกับความสามารถทางการเรียนมีค่าสูงกว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบรูปปิรามิดแบบข้างตัดขวาง เมื่อใช้วิธีการตรวจให้คะแนนตามค่าความยากเฉลี่ยของข้อสอบที่ทำ ส่วนวิธีการตรวจให้คะแนนตามค่าความยากเฉลี่ยของข้อสอบที่ตอบถูกพบว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบรูปปิรามิดที่ใช้ขนาดชั้นคงที่และแบบข้างตัดกับความสามารถทางการเรียนมีค่าสูงกว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบรูปปิรามิดแบบที่ใช้ขนาดชั้นแปรผัน สำหรับวิธีการตรวจให้คะแนนข้อสอบทุกข้อ พบว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบรูปปิรามิดแบบที่ใช้ขนาดชั้นแปรผันกับความสามารถทางการเรียนสูงกว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบรูปปิรามิดที่ใช้ขนาดชั้นคงที่และแบบข้างตัด ผลการทดสอบ

ค่าสัมประสิทธิ์สหสัมพันธ์ภายในระหว่างวิธีการตรวจให้คะแนน 3 วิธี โดยแยกผลการวิเคราะห์ตามแบบทดสอบรูปปริมาตรแต่ละฉบับ มีค่าเป็นบวกอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกค่า และผลการทดสอบค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคู่ พบว่า แบบทดสอบรูปปริมาตรแต่ละฉบับส่วนใหญ่มีค่าสัมประสิทธิ์ภายในระหว่างการตรวจให้คะแนน 3 วิธีแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และ .05

ชวรงค์ชัย คงเสน่ห์ (2530) สร้างแบบทดสอบสองขั้นตอน โดยใช้ทฤษฎีการตอบข้อคำถามวิชาฟิสิกส์ เรื่องการเคลื่อนที่และกฎการเคลื่อนที่ เพื่อใช้วัดผลสัมฤทธิ์ของนักเรียนต่างระดับความสามารถ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 ปีการศึกษา 2529 ของโรงเรียนสังกัดกรมสามัญศึกษาในจังหวัดสุราษฎร์ธานี จำนวน 916 คน วิเคราะห์ข้อสอบรายข้อตามทฤษฎีการตอบข้อคำถาม โดยใช้โปรแกรมคอมพิวเตอร์สำเร็จรูปโลจิส 5 ได้แบบทดสอบแยกกลุ่ม จำนวน 20 ข้อ และแบบทดสอบหลัก จำนวน 4 ฉบับ ๆ ละ 30 ข้อ การทดสอบตามระดับความสามารถของนักเรียน ใช้วิธีการประมาณค่า 3 วิธี คือ คำนวณโดยโปรแกรมโลจิส 5 ใช้ค่าความยากสูงสุดของข้อที่ตอบถูกและใช้ค่าความยากเฉลี่ยจากข้อที่นักเรียนตอบถูก ศึกษาความสัมพันธ์ระหว่างวิธีประมาณค่าสามวิธี โดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์และสร้างมาตราเทียบความสามารถของนักเรียน ผลการวิจัย พบว่าแบบทดสอบแยกกลุ่มมีค่าความยากอยู่ระหว่าง -0.6194 ถึง 0.6812 มีค่าอินฟอर्मเมชันของแบบทดสอบสูงสุดที่ความสามารถ 0.2000 แบบทดสอบหลักฉบับที่ 1 มีค่าความยากอยู่ระหว่าง -2.4533 ถึง 0.8765 มีค่าอินฟอर्मเมชันของแบบทดสอบสูงสุดที่ความสามารถ -0.7500 แบบทดสอบหลักฉบับที่ 2 มีค่าความยากอยู่ระหว่าง -0.7906 ถึง 1.7110 มีค่าอินฟอर्मเมชันของแบบทดสอบสูงสุดที่ความสามารถ 1.3500 แบบทดสอบหลักฉบับที่ 3 มีค่าความยากอยู่ระหว่าง 1.4614 ถึง 2.2523 มีค่าอินฟอर्मเมชันของแบบทดสอบสูงสุดที่ความสามารถ 2.0000 แบบทดสอบหลักฉบับที่ 4 มีค่าความยากอยู่ระหว่าง 1.9383 ถึง 2.9984 มีค่าอินฟอर्मเมชันของแบบทดสอบสูงสุดที่ความสามารถ 2.6500 นอกจากนี้ยังพบว่าแบบทดสอบหลักฉบับที่ 4 มีค่าเฉลี่ยความสามารถนักเรียนจากการประมาณค่าทั้ง 3 วิธีสูงสุด โดยนักเรียนที่ทำแบบทดสอบหลักฉบับที่ยากกว่ามีระดับความสามารถสูงกว่านักเรียนที่ทำแบบทดสอบหลักฉบับที่ง่ายกว่าซึ่งมีค่าเฉลี่ยของข้อสอบที่ใช้ร่วมกันต่ำด้วย สำหรับการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างการประมาณค่าความสามารถพบว่าการใช้ค่าความยากสูงสุดของข้อที่ตอบถูกกับการใช้ค่าความยากเฉลี่ยของข้อที่ตอบถูกและการใช้โปรแกรมสำเร็จรูปโลจิส 5 กับการใช้ค่าความยากเฉลี่ยของข้อที่ตอบถูกสัมพันธ์กันอย่างมีนัยสำคัญที่ระดับ .01 ทุกฉบับ ส่วนการใช้โปรแกรมสำเร็จรูปโลจิส 5 กับการใช้ค่าความยากสูงสุดของข้อที่ตอบถูกสัมพันธ์กันอย่างมีนัยสำคัญที่ระดับ .01 เพียง 1 ฉบับ

นันทิยา พึ่งคำ (2531) เปรียบเทียบคุณภาพการทดสอบแบบ ซี เอ ที และแบบประเพณีนิยมในการวัดความสามารถด้านคำศัพท์ภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในด้านความตรงเชิงเกณฑ์สัมพัทธ์และความคลาดเคลื่อนมาตรฐานในการประมาณค่า ผลการวิจัยพบว่าการทดสอบ ซี เอ ที มีความตรงเชิงเกณฑ์สัมพัทธ์ไม่แตกต่างจากการทดสอบแบบประเพณีนิยมแม้ว่าการทดสอบแบบ ซี เอ ที จะลดความยาวของแบบสอบถามมากกว่าร้อยละ 50 ส่วนค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพบว่าการทดสอบแบบ ซี เอ ที มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่าการทดสอบแบบประเพณีนิยม

เสรี ปรมชวลิตโรจน์ (2531) ศึกษาความเที่ยงตรงในการจำแนกความรอบรู้และการประมาณค่าความสามารถของแบบทดสอบเพลิกซีเลเวลและแบบทดสอบรูปปริมาตรที่ใช้ขนาดชั้นแปรผันวิชาคณิตศาสตร์ เรื่องสมการ กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2531 ของโรงเรียนสังกัดสำนักงานสามัญศึกษา จังหวัดเชียงใหม่ จำนวน 5,987 คน จากการสุ่มตัวอย่างอย่างง่าย ประมาณค่าความสามารถผู้สอบโดยให้คะแนนแบบแมกซิมันไลลิสซูด ผลการวิจัยพบว่าแบบทดสอบเพลิกซีเลเวลและแบบทดสอบรูปปริมาตรที่ใช้ขนาดชั้นแปรผันมีค่าความเที่ยงตรงในการจำแนกความรอบรู้เป็นบวกอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และ .05 นอกจากนี้ยังพบว่าความเที่ยงตรงในการจำแนกความรอบรู้จากแบบทดสอบเพลิกซีเลเวลมีค่าสูงกว่าแบบแบบทดสอบรูปปริมาตรที่ใช้ขนาดชั้นแปรผันทั้ง 2 ฉบับ

สงวค์ บุญปลูก (2533) ศึกษาความเที่ยงตรงในการจำแนกความรอบรู้และประมาณค่าความสามารถในวิชาคณิตศาสตร์ เรื่องอัตราส่วนและร้อยละ จากแบบทดสอบอิงเกณฑ์แบบทดสอบเทเลอร์ชนิดแยกทางและแบบทดสอบเทเลอร์ชนิดทดสอบตามสภาพ ผลการวิจัยพบว่าแบบทดสอบอิงเกณฑ์มีค่าความเที่ยงตรงในการจำแนกความรอบรู้ เท่ากับ .8324 แบบทดสอบเทเลอร์ชนิดแยกทางมีค่าความเที่ยงตรงในการจำแนกความรอบรู้ เท่ากับ .8557 แบบทดสอบเทเลอร์ชนิดทดสอบตามสภาพมีค่าความเที่ยงตรงในการจำแนกความรอบรู้ เท่ากับ .9052 ส่วนผลการประมาณค่าความสามารถพบว่าแบบทดสอบทั้ง 3 ประเภท ให้ค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ต่าย เชี่ยวฉวี (2534) ศึกษาเปรียบเทียบประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบจากการทดสอบเทเลอร์รูปปริมาตรที่มีรูปแบบ จำนวนชั้นและวิธีการให้คะแนนที่แตกต่างกันโดยใช้วิธีมอนติ คาร์โล โดยใช้คะแนนการตรวจสอบคุณภาพการศึกษาทั่วประเทศ ปีการศึกษา 2531 วิชาภาษาไทย ชั้นประถมศึกษาปีที่ 6 จำนวน 106,637 คน มาแปลงเป็นคะแนนมาตรฐานซี แล้วสุ่มคะแนนมาตรฐานซีมาเป็นกลุ่มตัวอย่าง จำนวน 32 กลุ่ม ๆ ละ 1,000 คน

ใช้รูปแบบการทดลองแบบคอมพลีทีลี่ แรนดอมไมซ์แฟคทอเรียล ดีไซน์ (CRF 2×4× 4) มีทั้งหมด 32 กลุ่มทดลอง ในแต่ละกลุ่มทดลองได้ทำการทดลองซ้ำ 40 ครั้ง หาค่าโอกาสในการตอบถูกโดยใช้ 1-พารามิเตอร์ แบบโลจิสติกโมเดล ผลการวิจัยพบว่าการให้คะแนนตามค่าความยากของข้อที่ยากที่สุดที่ทำถูกประมาณค่าความสามารถของผู้สอบสูงกว่าความสามารถจริงอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ในทุกกลุ่มและทุกครั้งของการทดลอง การให้คะแนนตามค่าความยากของข้อที่ทำถูกประมาณค่าความสามารถของผู้สอบต่ำกว่าความสามารถจริงอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ในทุกกลุ่มและทุกครั้งของการทดลอง การให้คะแนนโดยใช้วิธีของเบส์และวิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไขในแต่ละกลุ่มการทดลองจากการทดลองซ้ำ 40 ครั้ง มีจำนวนครั้งไม่ต่ำกว่า 34 ครั้ง ที่ความสามารถจริงกับความสามารถที่ได้จากการประมาณค่าไม่แตกต่างกัน สำหรับจำนวนขั้นที่ใช้พบว่าเมื่อมีจำนวนขั้นเพิ่มขึ้น มีแนวโน้มจะให้ประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงขึ้น เมื่อให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ทำถูก วิธีของเบส์และวิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข ซึ่งถ้าใช้จำนวนขั้นตั้งแต่ 10 ขั้นขึ้นไป การให้คะแนนโดยใช้วิธีของเบส์จะให้ประสิทธิภาพในการประมาณค่าความสามารถของผู้สอบสูงที่สุด โดยเฉพาะแบบขนาดขั้นคงที่ที่มี 16 ขั้น รองลงมาคือวิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไขและวิธีให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ทำถูก นอกจากนี้ยังพบว่าการทดสอบเทเลอรรูปปิรามิดที่ใช้รูปแบบที่ต่างกัน จำนวนขั้นที่ต่างกันหรือวิธีการให้คะแนนที่ต่างกันจะมีผลต่อค่าความคลาดเคลื่อน ค่าความคลาดเคลื่อนมาตรฐานและค่าความเที่ยงตรงในการประมาณค่าความสามารถของผู้สอบ

ตรึงใจ พูนผลอำนวย (2534) พัฒนาแบบทดสอบเฉพาะบุคคล วิชาวิทยาศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ดำเนินการสร้างแบบทดสอบจากรูปแบบฟาเซท ตรวจสอบความตรงตามเนื้อหาโดยพิจารณาจากการประเมินของผู้เชี่ยวชาญ วิเคราะห์ค่าพารามิเตอร์ของข้อสอบโดยใช้โปรแกรม Logist 5 จัดกลุ่มข้อสอบตามค่าความยากโดยใช้โครงสร้างรูปแบบปิรามิดขนาดขั้นคงที่ จำนวน 10 ขั้น ได้ 19 ขั้น ผลการวิจัยพบว่าแบบทดสอบเฉพาะบุคคลที่สร้างขึ้นมีความตรงตามเนื้อหา โดยมีค่าเฉลี่ยของการตัดสินความสอดคล้องของผู้เชี่ยวชาญอยู่ระหว่าง .75 ถึง 1.00 มีความตรงตามสภาพ โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนความสามารถของผู้สอบกับเกรดเฉลี่ยวิชาวิทยาศาสตร์เท่ากับ .7517 และมีค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากแบบทดสอบเฉพาะบุคคลจำนวน 10 ข้อกับแบบทดสอบผลสัมฤทธิ์ทางวิทยาศาสตร์ที่ใช้การเขียนตอบจำนวน 55 ข้อ เท่ากับ .7818 มีความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของแบบทดสอบเฉพาะบุคคลเท่ากับ .443

สุพัฒน์ สุกมลสันต์ (2536) สร้างโปรแกรมคอมพิวเตอร์และศึกษาคุณภาพและประสิทธิภาพของการทดสอบปรับเปลี่ยนโยงเนื้อหา โดยใช้เครื่องคอมพิวเตอร์ของแบบทดสอบอิงบริเขต วิชาภาษาอังกฤษ แบบทดสอบแบบปรับเปลี่ยนโยงเนื้อหา 5 บริเขต ๆ ละ 21 ข้อ สร้างเป็นแบบปิรามิด 5 ชั้น เปรียบเทียบผลการทดสอบหลาย ๆ ครั้ง โดย z-test และ t-test ผลการวิจัยพบว่าโปรแกรมที่จัดสร้างขึ้นทำงานได้ดีเป็นที่น่าพอใจและการทดสอบปรับเปลี่ยนโยงเนื้อหาคุณภาพและประสิทธิภาพดีกว่าแบบประเพณีนิยม โดยมีความตรงเชิงพยากรณ์สูงกว่า มีความเชื่อถือได้ไม่ต่างกัน ใช้ข้อสอบน้อยกว่าร้อยละ 50 มีความคลาดเคลื่อนในการวัดน้อยกว่าใช้เวลาในการทดสอบน้อยกว่าประมาณ 3 เท่า และทำให้ผู้สอบมีเจตคติในการทดสอบดีกว่า

ชูเกียรติ ละอองแก้ว (2538) สร้างแบบทดสอบสองชั้นตอนด้วยคอมพิวเตอร์ ในวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 5 เรื่องบทประยุกต์ วิเคราะห์ค่าพารามิเตอร์ของข้อสอบด้วยโปรแกรม BILOG VERSION 3.0 แบบทดสอบสองชั้นตอนประกอบด้วย แบบทดสอบเพื่อแยกทางจำนวน 10 ข้อ และแบบทดสอบเพื่อวัดผล จำนวน 4 ฉบับ ๆ ละ 20 ข้อ หาค่าความเชื่อมั่นของแบบทดสอบโดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากการสอบ 2 ครั้ง ผลการวิจัยพบว่าแบบทดสอบสองชั้นตอนที่ดำเนินการสอบด้วยคอมพิวเตอร์ มีค่าความเชื่อมั่นเท่ากับ .8344 ส่วนความคิดเห็นของนักเรียนที่เป็นกลุ่มตัวอย่างเห็นว่าแบบทดสอบแบบสองชั้นตอนที่ดำเนินการทดสอบด้วยคอมพิวเตอร์มีประสิทธิภาพในการนำไปใช้ค่อนข้างสูง ใช้ดำเนินการในชั้นเรียนได้ดี โดยผู้สอบไม่ต้องการการอธิบายเพิ่มเติม ผู้สอนมีแรงจูงใจอยากทำข้อสอบมากขึ้น ไม่มีความวิตกกังวลในขณะสอบและอยากให้มีการสอบแบบนี้ในวิชาอื่น ๆ อีก สำหรับผลการประเมินประสิทธิภาพของโปรแกรมคอมพิวเตอร์ พบว่า ไม่มีข้อผิดพลาดทางเทคนิค และไม่มีปัญหาที่เกิดจากผู้ใช้

ชูใจ ทรงเมฆ (2538) ศึกษาเปรียบเทียบผลการประมาณความสามารถด้วยวิธีการทดสอบแบบเทเลอร์และวิธีการทดสอบแบบดั้งเดิม ในการทดสอบผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ เรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 5 แบบทดสอบแบบเทเลอร์จัดเรียงข้อสอบตามโครงสร้างของการตอบเป็นรูปปิรามิดจำนวนชั้นคงที่ชนิด 10 ชั้น ใช้เทคนิค PROX ประมาณค่าความสามารถจากผลการตอบทั้ง 2 วิธี แล้วนำมาทดสอบความแตกต่างด้วยสถิติทดสอบที ผลการวิจัยพบว่า การทดสอบแบบเทเลอร์และการทดสอบแบบดั้งเดิมให้ผลการประมาณค่าความสามารถในด้านผลสัมฤทธิ์ทางการเรียนไม่แตกต่างกัน

วสันต์ ทองไทย (2539) ศึกษาเปรียบเทียบผลการประมาณค่าความสามารถด้วยวิธีการทดสอบแบบเทเลอร์รูปพีระมิดขนาดชั้นคงที่และรูปแบบพีระมิดข้างตัด ในการวัดทักษะ



กระบวนการทางวิทยาศาสตร์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ประมาณค่าความสามารถตามหลักการของเบส์ ทดสอบความแตกต่างของผลการประมาณค่าความสามารถด้วยสถิติ One Way Repeated Measures Analysis of Variance และทดสอบความแตกต่างของค่าฟังก์ชันสารสนเทศและค่าความเที่ยงในการประมาณค่าความสามารถโดยใช้ดัชนีร้อยละของความแตกต่าง ผลการวิจัยพบว่าการทดสอบทั้ง 3 วิธี ให้ค่าประมาณความสามารถ ค่าฟังก์ชันสารสนเทศและค่าความเที่ยงในการประมาณค่าความสามารถไม่แตกต่างกัน

สายชล อบทม (2539) พัฒนาโปรแกรมสำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ โดยใช้ชุดคำสั่งภาษา PASCAL และใช้ Delphi Version 2.0 เป็นตัวแปรภาษาประเมินคุณภาพของโปรแกรมโดยครูสังกัดต่าง ๆ จำนวน 17 คนและนักเรียน จำนวน 122 คน ผลการวิจัยทำให้ได้โปรแกรมคอมพิวเตอร์สำเร็จรูปสำหรับการทดสอบแบบปรับเหมาะที่ใช้รูปแบบทางแยกแปรผัน ประมาณค่าความสามารถของผู้สอบโดยใช้วิธีของเบส์ ยุติการทดสอบเมื่อความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบต่ำกว่า 0.3 ประกอบด้วยโปรแกรมย่อย 2 โปรแกรม คือ โปรแกรมย่อยพิมพ์ข้อสอบและโปรแกรมย่อยการทดสอบที่ปฏิบัติการบนระบบวินโดว์ 95 ผลการจากประเมินโปรแกรม พบว่า โปรแกรมมีประสิทธิภาพและเป็นประโยชน์ต่อการวัดและประเมินผล รวมทั้งสามารถนำไปใช้ในสถานศึกษาได้

อำไพ วรณเกษม (2539) สร้างแบบทดสอบเฉพาะบุคคลแบบ ซี เอ ที่ รูปแบบทางแยกแปรผัน โดยใช้กลยุทธ์แบบเบส์ ประมาณค่าความสามารถของผู้สอบโดยวิธีเบเซียน อัพเดทตั้งและหาคุณภาพของแบบทดสอบที่สร้างขึ้น ผลการวิจัยพบว่า แบบทดสอบที่สร้างขึ้นมีความตรงตามเนื้อหาและความถูกต้องทางภาษาตามผลการประเมินความสอดคล้องของผู้เชี่ยวชาญแบบทดสอบเฉพาะบุคคล จำนวน 120 ข้อ จัดกลุ่มโดยเรียงตามค่าความยาก 9 ระดับ ๆ ละ 5-30 ข้อ มีค่าความยากช่วงละประมาณ .9 ใช้กฎการเลือกข้อสอบแบบเพิ่ม 1 /ลด 1 มีค่าอำนาจจำแนกอยู่ระหว่าง .370 ถึง .887 ค่าความยากกระจายอยู่ระหว่าง -3.673 ถึง 3.700 ค่าการเดากระจายอยู่ระหว่าง .084 ถึง .300 ผู้สอบทำข้อสอบจำนวน 5-20 ข้อ ใช้เวลานาน .55 ถึง 18.04 นาที มีค่าความสามารถระหว่าง -1.93 ถึง 1.28 และค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของแบบทดสอบเท่ากับ .34

รังสรรค์ มณีเล็ก (2540) ศึกษาผลของตัวแปรบางตัวต่อความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ วิชาคณิตศาสตร์ เรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 ผลการวิจัยพบว่าความยากง่ายของข้อสอบข้อแรกที่แตกต่างกันมีผลทำให้ค่าความเที่ยงตรงเชิงสภาพไม่แตกต่างกัน แต่ทำให้จำนวนข้อสอบที่ใช้ในการทดสอบแตกต่างกัน โดยความยากง่ายของข้อสอบข้อแรกในระดับปานกลางมีผลทำให้

ใช้ข้อสอบในการทดสอบน้อยที่สุด เกณฑ์การคัดเลือกข้อสอบจากคลังข้อสอบที่แตกต่างกันมีผลทำให้ค่าความเที่ยงตรงเชิงสภาพไม่แตกต่างกัน แต่ทำให้จำนวนข้อสอบที่ใช้ในการทดสอบแตกต่างกัน วิธีใช้ข้อสอบที่มีค่าฟังก์ชันสารสนเทศสูงสุดใช้จำนวนข้อสอบน้อยที่สุด วิธีการประมาณค่าความสามารถของผู้สอบที่แตกต่างกันมีผลทำให้ค่าความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการทดสอบแตกต่างกัน วิธีเบสท์ที่ปรับใหม่ให้ค่าความเที่ยงตรงเชิงสภาพสูงกว่าและใช้จำนวนข้อสอบมากกว่า เกณฑ์ยุติการทดสอบซึ่งกำหนดให้ค่าความคลาดเคลื่อนในการประมาณค่าความสามารถน้อยกว่า .30 ให้ค่าความเที่ยงตรงเชิงสภาพสูงสุดและใช้ข้อสอบมากที่สุด ระดับความสามารถทางคณิตศาสตร์ของผู้สอบที่แตกต่างกันมีผลทำให้ค่าความเที่ยงตรงเชิงสภาพและจำนวนข้อสอบที่ใช้ในการทดสอบแตกต่างกัน ระดับความสามารถสูงใช้ข้อสอบน้อยกว่าระดับความสามารถต่ำ

ละมัย พรหมศร (2540) สร้างแบบทดสอบปิรามิดสำหรับใช้ทดสอบด้วยคอมพิวเตอร์ วิชาคณิตศาสตร์ เรื่อง ภาคตัดกรวย สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 4 วิเคราะห์ค่าพารามิเตอร์ของข้อสอบโดยใช้โปรแกรม IRT Analysis By Bay Version 2.1 สร้างแบบทดสอบแบบปิรามิดชนิด 12 ชั้น จำนวน 78 ข้อ โดยใช้โปรแกรมสำเร็จรูป Authorware Professional 2.0 หาความตรงเชิงเกณฑ์สัมพันธ์โดยค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนสอบจากแบบทดสอบแบบปิรามิดจำนวน 2 ครั้งกับเกรดวิชาคณิตศาสตร์ ค 011 ผลการวิจัยพบว่าแบบทดสอบแบบปิรามิดที่สร้างขึ้นมีค่าความตรงตามเกณฑ์สัมพันธ์เท่ากับ .6875

สุวิทยา พรหมอ่อน (2540) สร้างแบบทดสอบเทเลอรรูปปิรามิดขนาดชั้นคงที่ วิชาคณิตศาสตร์ เรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 ประมาณค่าความสามารถของผู้สอบด้วยวิธีของเบสท์ แบบทดสอบแบบปิรามิดมีจำนวนชั้น 16 ชั้น จำนวนข้อสอบ 136 ข้อ ใช้กฎการแยกเส้นทางแบบเพิ่ม 1/ ลด 1 ผลการวิจัยพบว่า การเลือกข้อสอบในชั้นที่สูงขึ้นค่าความยากที่เพิ่มขึ้นหรือลดลงคงที่เท่ากับ .12 ค่าความยากของจุดกึ่งกลางที่ฐานโครงสร้างของแบบทดสอบมีค่าเท่ากับ .20 และมีค่าความตรงเชิงสัมพันธ์กับเกณฑ์เท่ากับ .8218

นุจรี สุภิสุทธิ (2541) สร้างชุดทดสอบแบบเทเลอร์และดำเนินการทดสอบโดยใช้ไมโครคอมพิวเตอร์ในวิชาคณิตศาสตร์ เรื่องสมการ ชั้นมัธยมศึกษาปีที่ 1 เพื่อศึกษาผลการใช้โปรแกรมและความคิดเห็นที่มีต่อการใช้โปรแกรม ผลการวิจัยพบว่า การติดตั้งโปรแกรมการทดสอบสามารถทำได้ง่าย ลักษณะโดยทั่วไปของคู่มือการใช้และลักษณะทั่วไปของชุดทดสอบอยู่ในเกณฑ์ดี ผู้สอบใช้จำนวนข้อสอบในการทดสอบโดยเฉลี่ยประมาณ 15 ข้อ เวลาที่ใช้โดยเฉลี่ยประมาณ 15 นาที ผลการทดสอบค่าความแตกต่างระหว่างค่าเฉลี่ยของคะแนนสอบจากการสอบ

2 ครั้ง พบว่าไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 นักเรียนมีความคิดเห็นต่อการทดสอบไปในทางบวก

ปราโมทย์ แส่นหาญ (2542) สร้างแบบทดสอบปรับระดับขั้นที่ดำเนินการสอบด้วยคอมพิวเตอร์ วิชาคณิตศาสตร์ เรื่องจำนวนเชิงซ้อน ชั้นมัธยมศึกษาปีที่ 5 ทดสอบความเป็นเอกมิติของแบบทดสอบด้วยโปรแกรม SPSS for WINDOWS หาค่าพารามิเตอร์ของแบบสอบโดยใช้โปรแกรม BILOG 3 สร้างแบบทดสอบปรับระดับขั้นโดยใช้โปรแกรม ToolBook 5 สร้างแบบทดสอบปรับระดับขั้นจำนวน 9 ระดับขั้น จำนวน 142 ข้อ หาค่าความเชื่อมั่นของแบบทดสอบโดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบปรับระดับขั้น จำนวน 2 ครั้งและหาค่าความตรงเชิงเกณฑ์สัมพัทธ์ของแบบทดสอบโดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบปรับระดับขั้นกับแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนที่ทดสอบด้วยวิธีปกติ ผลการวิจัยพบว่าแบบทดสอบปรับระดับขั้นที่ประมาณค่าความสามารถของผู้สอบด้วยค่าความยากสูงสุดของข้อสอบที่ตอบถูก มีค่าความเชื่อมั่นเท่ากับ .87 แบบทดสอบปรับระดับขั้นที่ประมาณค่าความสามารถของผู้สอบด้วยค่าความยากเฉลี่ยของข้อสอบที่ตอบถูก มีค่าความเชื่อมั่นเท่ากับ .92 สำหรับค่าความเที่ยงตรงเชิงเกณฑ์สัมพัทธ์พบว่าแบบทดสอบปรับระดับขั้นที่ประมาณค่าความสามารถของผู้สอบด้วยค่าความยากสูงสุดของข้อสอบที่ตอบถูก มีค่าความเที่ยงตรงเชิงเกณฑ์สัมพัทธ์เท่ากับ .58 แบบทดสอบปรับระดับขั้นที่ประมาณค่าความสามารถของผู้สอบด้วยค่าความยากเฉลี่ยของข้อสอบที่ตอบถูก มีค่าความเที่ยงตรงเชิงเกณฑ์สัมพัทธ์เท่ากับ .74 นอกจากนี้ยังพบว่าแบบทดสอบที่สร้างขึ้นมีประสิทธิภาพเชิงปฏิบัติการ ได้แก่ มีความเหมาะสม ความชัดเจนของคำชี้แจงและขั้นตอนในการดำเนินการทดสอบ ความง่ายในการปฏิบัติและการจูงใจในการสอบอยู่ในเกณฑ์ดีทุกด้าน

ปราริชญ์ นามคุณ (2541) เปรียบเทียบความเที่ยงตรงในการจำแนกความรู้และการประมาณค่าความสามารถของผู้สอบในวิชาคณิตศาสตร์ เรื่องสมการและการแก้สมการของแบบทดสอบอิงเกณฑ์ แบบทดสอบปิรามิดขนาดขั้นคงที่และแบบทดสอบปิรามิดขนาดขั้นแปรผัน ประมาณค่าพารามิเตอร์ของข้อสอบตามวิธีของเบย์ โดยใช้โปรแกรมสำเร็จรูป IRT Analysis by Bayes version 2.1 ผลการวิจัยพบว่าแบบทดสอบแบบอิงเกณฑ์มีความเที่ยงตรงในการจำแนกความรู้สูงสุด รองลงมาคือแบบทดสอบปิรามิดขนาดขั้นแปรผันและแบบทดสอบปิรามิดขนาดขั้นคงที่ แบบทดสอบปิรามิดขนาดขั้นคงที่และแบบทดสอบปิรามิดขนาดขั้นแปรผันมีผลการจำแนกความรู้มากกว่าแบบทดสอบอิงเกณฑ์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 แต่มีค่าประมาณความสามารถของนักเรียนไม่แตกต่างกัน

ศลิษา สัตถาผล (2542) ศึกษาเปรียบเทียบแบบทดสอบสองขั้นตอนและแบบทดสอบสามขั้นตอนที่ดำเนินการสอบโดยใช้คอมพิวเตอร์ ในวิชาภาษาอังกฤษ ชั้นประถมศึกษาปีที่ 6 วิเคราะห์ค่าพารามิเตอร์ของข้อสอบด้วยโปรแกรม IRT (BAY 1.0) แบบทดสอบสองขั้นตอน ประกอบด้วยแบบทดสอบแยกทาง จำนวน 10 ข้อ และแบบทดสอบวัดผลจำนวน 5 ฉบับ ๆ ละ 30 ข้อ แบบทดสอบสามขั้นตอนประกอบด้วย แบบทดสอบแยกทางครั้งที่ 1 จำนวน 10 ข้อ แบบทดสอบแยกทางครั้งที่ 2 จำนวน 3 ฉบับ ๆ ละ 10 ข้อ และแบบทดสอบวัดผล จำนวน 6 ฉบับ ๆ ละ 20 ข้อ ผลการวิจัยพบว่าในระยะพัฒนาและทดลองใช้โปรแกรมไม่มีปัญหาทางเทคนิคและไม่มีปัญหาที่เกิดจากผู้ใช้ สำหรับผลการประเมินความเชื่อถือได้จากแบบทดสอบพบว่าแบบทดสอบสองขั้นตอนและแบบทดสอบแบบประเพณีนิยมมีค่าความเชื่อมั่นไม่ต่างจากเกณฑ์ที่กำหนด (.80) โดยมีค่าความเชื่อมั่นเท่ากับ .858 และ .801 ส่วนแบบทดสอบสามขั้นตอนมีค่าความเชื่อมั่นสูงกว่าเกณฑ์ที่กำหนด โดยมีค่าความเชื่อมั่นเท่ากับ .979 ผลการประเมินทางจิตสังคมพบว่าแบบทดสอบที่ดำเนินการสอบโดยใช้คอมพิวเตอร์มีประสิทธิภาพด้านการนำไปใช้ นักเรียนมีทัศนคติที่ดีต่อการทำแบบทดสอบ

อนุชิต กอศักดิ์ (2542) สร้างแบบทดสอบเทเลอร์ ชนิดยุทธวิธีสองขั้นตอน วิชาคณิตศาสตร์ เรื่องฟังก์ชันเอกซ์โพเนนเชียลและฟังก์ชันลอการิทึม ชั้นมัธยมศึกษาปีที่ 5 วิเคราะห์ค่าพารามิเตอร์ของข้อสอบด้วยวิธีของเบส์ ทดสอบความเป็นเอกมิติของแบบทดสอบโดยการวิเคราะห์ตัวประกอบโดยใช้โปรแกรมคอมพิวเตอร์ SPSS FOR WINDOWS แบบทดสอบ 2 ขั้นตอนประกอบด้วยแบบทดสอบแยกกลุ่ม จำนวน 10 ข้อ แบบทดสอบหลัก 4 ฉบับ ๆ ละ 30 ข้อ ประเมินค่าความสามารถของนักเรียนด้วยวิธีของเบส์ วิธีใช้ค่าความยากของข้อที่ยากที่สุดที่นักเรียนตอบถูกและวิธีใช้ค่าความยากเฉลี่ยของทุกข้อที่นักเรียนตอบถูก ผลการวิจัยพบว่านักเรียนที่ทำแบบทดสอบหลักฉบับที่ยากกว่ามีความสามารถสูงกว่านักเรียนที่ทำแบบทดสอบหลักฉบับที่ง่ายกว่าการประมาณค่าความสามารถด้วยวิธีใช้ค่าความยากของข้อที่ยากที่สุดที่นักเรียนตอบถูกและวิธีใช้ค่าความยากเฉลี่ยของทุกข้อที่นักเรียนตอบถูกสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกฉบับ นอกนั้นมีความสัมพันธ์กันบางฉบับ ซึ่งทั้ง 3 วิธีมีความสัมพันธ์กับคะแนนความสามารถตามทฤษฎีดั้งเดิมอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกฉบับ ยกเว้นวิธีใช้ค่าความยากเฉลี่ยของทุกข้อที่นักเรียนตอบถูกสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติเฉพาะในฉบับที่ 1

กนกวรรณ รัตนธ (2544) ศึกษาเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกัน กลุ่มตัวอย่าง คือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2543 จำนวน 319 คน

เครื่องมือที่ใช้ คือ แบบทดสอบวัดความสามารถทางภาษาอังกฤษ คลังข้อสอบวิชาภาษาอังกฤษ และโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Micro CAT) วิเคราะห์ข้อมูลโดยใช้ค่าสถิติเบื้องต้น ค่าสหสัมพันธ์ ค่าไคสแควร์และการวิเคราะห์ความแปรปรวนแบบวัดซ้ำ ผลการวิจัยพบว่าระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกันทำให้ความตรงตามสภาพ ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แตกต่างกัน ผู้สอบที่มีระดับความสามารถสูงและมีอัตราการใช้ข้อสอบซ้ำสูงขึ้นส่งผลให้ความตรงตามสภาพลดลง ส่วนความคงที่ในการประมาณค่าความสามารถ พบว่าระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกันทำให้ความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์แตกต่างกัน ผู้สอบที่มีระดับความสามารถสูงและมีอัตราการใช้ข้อสอบซ้ำสูงขึ้นส่งผลให้ความคงที่ในการประมาณค่าความสามารถลดลง

พงษ์พิชิต พรหมสิทธิ์ (2544) ดำเนินการสอบแบบทดสอบเทเลอร์ ตามยุทธวิธีของเบส์ในวิชาคณิตศาสตร์ (ค.014) ชั้นมัธยมศึกษาปีที่ 5 โดยเว็บเพจ วิเคราะห์ค่าพารามิเตอร์ของข้อสอบด้วยโปรแกรม IRT (BAY) 1.0 สร้างฐานข้อมูลกลุ่มข้อสอบด้วย Microsoft Access 97 พัฒนาเว็บเพจด้วย ASP 3.0 (Active Server Pages) ผลการวิจัยพบว่าแบบทดสอบเทเลอร์มีค่าสัมประสิทธิ์ความเชื่อมั่นเท่ากับ .9481 ซึ่งแตกต่างจากเกณฑ์ (.80) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ผลการประเมินเชิงจิตสังคมพบว่านักเรียนมีความคิดเห็นเกี่ยวกับการดำเนินการสอบอยู่ในระดับดี

เกียรติศักดิ์ สองแสง (2547) พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์และศึกษาประสิทธิภาพของโปรแกรมการทดสอบ โดยเปรียบเทียบเวลาที่ใช้ในการทดสอบ จำนวนข้อสอบของแบบทดสอบและค่าฟังก์ชันสารสนเทศของแบบทดสอบ เมื่อมีการจัดชุดข้อสอบให้ตอบในแต่ละชั้นต่างกัน กำหนดเวลาที่ใช้ในการตอบข้อสอบในแต่ละข้อต่างกันและมีการจัดรูปแบบข้อสอบต่างกัน ผลการวิจัยพบว่าโปรแกรมการทดสอบสามารถทำงานได้อย่างมีประสิทธิภาพ ผู้สอบมีความพึงพอใจในการทดสอบ ในด้านการจัดชุดข้อสอบพบว่าการจัดชุดข้อสอบให้ตอบชั้นละ 3 ข้อให้ค่าฟังก์ชันสารสนเทศของแบบทดสอบสูงที่สุด สำหรับการกำหนดเวลาที่ใช้ในการทดสอบข้อสอบแต่ละข้อพบว่าการกำหนดเวลาให้ตอบเฉลี่ยเท่ากันทุกข้อหรือการกำหนดเวลาให้ตอบข้อสอบในแต่ละข้อตามค่าความยากของข้อสอบจะทำให้ผู้สอบใช้เวลาในการทดสอบน้อยลง ส่วนการจัดชุดข้อสอบที่มีสื่อประสมประกอบการทดสอบพบว่าทำให้ผู้สอบใช้เวลาในการทดสอบน้อยลงและผู้สอบมีความพึงพอใจในการทดสอบมากขึ้น

จากการศึกษางานวิจัยที่เกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ในประเทศไทยดังกล่าวข้างต้น จะเห็นได้ว่ามีการใช้ทั้งคำว่า การทดสอบแบบปรับเหมาะ (adaptive testing) และการทดสอบแบบเทเลอร์ (tailored testing) ในการศึกษาซึ่งส่วนใหญ่เป็นการสร้างหรือพัฒนาแบบทดสอบและการศึกษาประสิทธิภาพของแบบทดสอบในด้านความเที่ยงและความตรง หรือเปรียบเทียบประสิทธิภาพของแบบทดสอบที่มีประเภทของแบบทดสอบแบบปรับเหมาะที่แตกต่างกัน มีงานวิจัยจำนวนน้อยมากที่ศึกษาความคิดเห็นที่มีต่อการทดสอบ สรุปสาระสำคัญที่ได้จากการศึกษาได้ ดังนี้

1. ความคิดเห็นที่มีต่อการทดสอบพบว่าผู้มีส่วนร่วมในการทดสอบมีความคิดเห็นต่อการทดสอบเป็นไปในทางบวกหรือมีทัศนคติที่ดีต่อการทดสอบ ไม่มีความวิตกกังวลในการทำข้อสอบ มีแรงจูงใจอยากได้รับการทดสอบ ขั้นตอนการทดสอบง่ายต่อการปฏิบัติ คำชี้แจงมีความเหมาะสมและชัดเจน รวมทั้งมีความคิดเห็นว่าเป็นการทดสอบที่สามารถนำไปใช้ในสถานศึกษาได้จริง

2. ประเภทของการทดสอบกับประสิทธิภาพการทดสอบ พบว่า

แบบทดสอบปรับเหมาะแบบปิรามิดขนาดขั้นแปรผันมีความเที่ยงมากกว่าแบบขั้นคงที่

แบบทดสอบปรับเหมาะแบบปิรามิดขนาดขั้นคงที่และแบบทดสอบแบบปิรามิดข้างตัดให้สารสนเทศและค่าความเที่ยงไม่แตกต่างกัน

แบบทดสอบปรับเหมาะแบบเฟล็กซิเลเวลมีค่าความเที่ยงสูงกว่าแบบทดสอบปรับเหมาะแบบปิรามิดขนาดขั้นแปรผัน

แบบทดสอบแบบอิงเกณฑ์ แบบทดสอบแบบประเพณีนิยมและแบบทดสอบแบบปรับเหมาะแบบปิรามิดทั้งขนาดขั้นแปรผันและขนาดขั้นคงที่ให้ค่าประมาณความสามารถไม่แตกต่างกัน

แบบทดสอบแบบปรับเหมาะที่ใช้ยุทธวิธีสองขั้นตอนมีค่าความเชื่อมั่นไม่แตกต่างจากแบบทดสอบแบบประเพณีนิยม

แบบทดสอบแบบปรับเหมาะที่ใช้ยุทธวิธีสามขั้นตอนมีค่าความเชื่อมั่นสูงกว่าแบบทดสอบที่ใช้ยุทธวิธีสองขั้นตอนและแบบทดสอบแบบประเพณีนิยม

3. อัตราการใช้ข้อสอบซ้ำกับระดับความสามารถ พบว่า

อัตราการใช้ข้อสอบซ้ำสูงในกลุ่มผู้สอบที่มีระดับความสามารถสูงทำให้ความตรงตามสภาพลดลง ความคงที่ในการประมาณค่าลดลง แนวโน้มการประมาณค่าความสามารถสูงขึ้น

อัตราการใช้ข้อสอบซ้ำต่างกันในกลุ่มผู้สอบที่มีระดับความสามารถต่างกันทำให้ความคงที่ในการประมาณค่าต่างกัน

อัตราการใช้ข้อสอบซ้ำโดยมีจำนวนครั้งเพิ่มขึ้นแนวโน้มการประมาณค่าความสามารถสูงขึ้น

4. ความยากของข้อสอบข้อแรก พบว่า การใช้ความยากของข้อสอบข้อแรกที่มีระดับความยากต่างกันทำให้ค่าความเที่ยงไม่ต่างกัน แต่จำนวนข้อสอบที่ใช้ต่างกัน โดยใช้ข้อสอบข้อแรกที่มีระดับความยากปานกลางจะใช้ข้อสอบจำนวนน้อยที่สุด

5. เกณฑ์การคัดเลือกข้อสอบจากคลังข้อสอบ พบว่า การใช้วิธีจับคู่ระหว่างค่าความยากกับระดับความสามารถ ค่าระดับความสามารถสูงสุดกับค่าระดับความสามารถและวิธีการใช้ค่าฟังก์ชันสารสนเทศสูงสุด ทำให้ค่าความตรงไม่แตกต่างกัน แต่จำนวนข้อสอบที่ใช้ต่างกัน โดยวิธีใช้ค่าฟังก์ชันสารสนเทศสูงสุดใช้จำนวนข้อสอบน้อยที่สุด

6. วิธีการประมาณค่าความสามารถของผู้สอบ พบว่า การใช้วิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข (conditional maximum likelihood) และวิธีเบย์ที่ปรับใหม่ (Bayesian updating) ทำให้ค่าความตรงต่างกัน วิธีเบย์ที่ปรับใหม่ให้ค่าความตรงสูงกว่าวิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข และทำให้จำนวนข้อสอบที่ใช้ต่างกัน วิธีเบย์ที่ปรับใหม่ใช้ข้อสอบจำนวนมากกว่าวิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข

7. เกณฑ์ยุติการสอบ พบว่า การใช้เกณฑ์ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าน้อยกว่า .30 ให้ค่าความตรงสูงกว่าและใช้จำนวนข้อสอบมากกว่าการใช้เกณฑ์ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าน้อยกว่า .40 และ .50

8. การให้คะแนน พบว่า การให้คะแนนโดยใช้วิธีของเบย์มีความสัมพันธ์กับการให้คะแนนตามระดับความยากของข้อที่ยากที่สุดที่ตอบถูกและมีความสัมพันธ์กับการให้คะแนนตามระดับความยากเฉลี่ยของข้อที่ตอบถูกและการให้คะแนนตามระดับความยากของข้อที่ยากที่สุดที่ตอบถูกมีความสัมพันธ์กับการให้คะแนนตามระดับความยากเฉลี่ยของข้อที่ตอบถูก นอกจากนี้ยังพบว่า การให้คะแนนตามระดับความยากของข้อที่ยากที่สุดที่ตอบถูกมีแนวโน้มการให้คะแนนสูงกว่าความเป็นจริง ส่วนการให้คะแนนตามระดับความยากเฉลี่ยของข้อที่ตอบถูกมีแนวโน้มการให้คะแนนต่ำกว่าความเป็นจริง

9. ระดับความสามารถของผู้สอบ พบว่า ผู้สอบที่มีระดับความสามารถต่างกันให้ค่าความตรงต่างกัน ผู้สอบที่มีระดับความสามารถสูงใช้ข้อสอบจำนวนน้อยกว่าผู้สอบที่มีระดับความสามารถต่ำ

10. การจัดชุดข้อสอบให้ตอบในแต่ละชั้น พบว่า การจัดชุดข้อสอบให้ตอบชั้นละ 3 ข้อให้ค่าฟังก์ชันสารสนเทศสูงกว่าการจัดชุดข้อสอบให้ตอบชั้นละ 2 ข้อและชั้นละ 1 ข้อ

11. การกำหนดเวลาที่ใช้ในการตอบ พบว่า การกำหนดเวลาแต่ละข้อเฉลี่ยเท่ากันและการกำหนดเวลาตามค่าความยากของข้อสอบทำให้ใช้เวลาในการสอบน้อยกว่าการไม่กำหนดเวลาในการสอบ

12. รูปแบบของข้อสอบ พบว่า การใช้ข้อสอบที่มีสื่อประสมภาพกราฟฟิก ภาพวีดิโอและเสียงประกอบการทดสอบทำให้ใช้เวลาในการทดสอบน้อยลงและผู้สอบมีความพึงพอใจในการทดสอบมากขึ้น

#### ตอนที่ 4 งานวิจัยเกี่ยวข้อง

##### งานวิจัยเกี่ยวกับลำดับชั้นของคุณลักษณะ

Chen; Khoo และ Yan (2002) ได้พัฒนาวิธีการในการเข้าถึงแบบแผนความต้องการของลูกค้าและทดลองใช้ใน wood golf club โดยใช้เทคนิคขั้นบันได (laddering) ที่เรียกว่า ลำดับชั้นคุณลักษณะของลูกค้า (customer attribute hierarchy: CAM) เพื่อค้นหาความต้องการของลูกค้า (customer requirements elicitation: CRE) และวิเคราะห์ความต้องการของลูกค้าและโอกาสทางการตลาด (customer/marketing analysis: CMA) โดยใช้ ART2 neural network ผลการวิจัยพบว่าสามารถวิเคราะห์ความต้องการของลูกค้าได้ละเอียดและชัดเจนขึ้น

Bělohávek; Vladimír และ Jirí (2004) ได้ศึกษาการจัดแบบแผนมโนทัศน์ของข้อมูลนำเข้าโดยใช้การเรียงมโนทัศน์ตามลำดับชั้นของคุณลักษณะ แทนวิธีการเรียงลำดับชั้นจากมโนทัศน์ย่อยไปมโนทัศน์ใหญ่ (subconcept superconcept hierarchy) ผลการวิจัยพบว่าการเรียงลำดับมโนทัศน์ตามลำดับชั้นของคุณลักษณะช่วยให้สามารถเข้าใจแบบแผนมโนทัศน์ได้ดีขึ้น

Han (2004) ได้พัฒนาเครื่องมือการควบคุมการตัดสินใจรายคู่ (pairwise dominance) ในโปรแกรมคอมพิวเตอร์ ซึ่งจำแนกเป็นการควบคุมแบบเข้ม (strict dominance) และแบบไม่เข้ม (weak dominance) โดยใช้การแบ่งส่วนข้อมูลเชิงเส้นและลำดับชั้นของคุณลักษณะรูปต้นไม้ซึ่งแยกเป็นต้นไม้ย่อย (sub tree) ผลการวิจัยพบว่าสามารถกำหนดน้ำหนักของคุณลักษณะ (attribute weight) และคุณค่าที่เป็นไปได้ (alternative value) ได้ชัดเจนขึ้น

กมล ภูประเสริฐ (2520) ได้เปรียบเทียบผลการทดสอบความเที่ยงตรงของลำดับชั้นการเรียนรู้ ตามวิธีของวอลเบสเซอร์-ไอเซนเบิร์กกับวิธีของไวท์-คลาร์ค จำแนกตามการเป็นลำดับชั้นที่ต่อเนื่องและไม่ต่อเนื่องของพฤติกรรมคู่หนึ่ง ความยากง่ายของข้อสอบ จำนวนข้อสอบที่ใช้และจำนวนกลุ่มตัวอย่างที่ใช้ กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 2 จำนวน 412 คน ผลการวิจัยพบว่าวิธีการทั้งสองให้ผลไม่แตกต่างกันเมื่อพฤติกรรมเป็นลำดับชั้นที่ต่อเนื่อง แต่ต่างกันเมื่อพฤติกรรมเป็นลำดับชั้นที่ไม่ต่อเนื่อง โดยวิธีของวอลเบสเซอร์-ไอเซนเบิร์กมีความแหลมคม (power) สูงกว่า ด้านความยากง่ายของข้อสอบและจำนวนข้อสอบที่ใช้พบว่าเมื่อใช้ข้อสอบที่ง่ายกว่าและจำนวนข้อมากกว่า วิธีของวอลเบสเซอร์-ไอเซนเบิร์กมีความเที่ยงตรงของค่าสถิติสูงกว่า



ความสัมพันธ์เชิงลำดับชั้นน้อยกว่าเมื่อพฤติกรรมเป็นลำดับชั้นต่อเนื่อง เมื่อใช้ข้อสอบที่ยากกว่า และจำนวนข้อน้อยกว่า วิธีของไวท์-คลาร์คมีความเที่ยงตรงของค่าสถิติสูงกว่า แต่วิธีของวอลเบเชอร์-ไอเซนเบิร์กมีความแหลมคมสูงกว่า ด้านจำนวนกลุ่มตัวอย่างที่ใช้พบว่าวิธีของไวท์-คลาร์คให้ผลเปลี่ยนแปลงเมื่อใช้กลุ่มตัวอย่างต่างกันมาก ทั้งสองวิธีให้ผลไม่แตกต่างกันเมื่อใช้กลุ่มตัวอย่างมาก แต่จะให้ผลแตกต่างกันเมื่อใช้กลุ่มตัวอย่างน้อย โดยวิธีของวอลเบเชอร์-ไอเซนเบิร์กมีความแหลมคมสูงกว่า

จินดา ลีเมถาวรศิริพงศ์ (2525) ได้ศึกษาการสร้างลำดับชั้นเนื้อหาวิชาจากการวินิจฉัยการเรียนคณิตศาสตร์เรื่องจำนวนเต็ม โดยยึดหลักเหตุผลตามลักษณะวิชาและหลักการสอน ตรวจสอบความตรงของลำดับชั้นเนื้อหาโดยผู้เชี่ยวชาญ จำนวน 7 คน พัฒนาแบบทดสอบวินิจฉัยการเรียนตามลำดับชั้นเนื้อหาและนำไปทดสอบกับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ในเขตกรุงเทพมหานคร จำนวน 396 คน ผลการวิจัยพบว่าได้ลำดับชั้นเนื้อหาวิชา 5 ตอน จำนวน 34 ชั้น ได้แก่ ความรู้พื้นฐาน จำนวน 7 ลำดับชั้น การบวกจำนวนเต็ม จำนวน 7 ลำดับชั้น การลบจำนวนเต็ม จำนวน 5 ลำดับชั้น การคูณจำนวนเต็ม จำนวน 9 ลำดับชั้นและการหารจำนวนเต็ม จำนวน 6 ชั้น เมื่อวิเคราะห์ข้อบกพร่องในการเรียนคณิตศาสตร์ พบว่าในเรื่องความรู้พื้นฐาน นักเรียนไม่เข้าใจการเขียนจำนวนเต็มบวกลงบนเส้นจำนวน เรื่องการบวกจำนวนเต็ม นักเรียนไม่เข้าใจเรื่องคุณสมบัติการจัดหมวดหมู่สำหรับการบวก เรื่องการลบจำนวนเต็ม นักเรียนไม่เข้าใจเรื่องเต็มบวก-จำนวนเต็มลบ เรื่องการคูณจำนวนเต็ม นักเรียนไม่เข้าใจเรื่องคุณสมบัติการกระจาย และเรื่องการหารจำนวนเต็ม นักเรียนไม่เข้าใจเรื่องคุณสมบัติของศูนย์เกี่ยวกับการหาร

วรรณวิภา จัตตชัย (2529) ได้ศึกษารูปแบบลำดับชั้นการเรียนรู้ด้านพุทธิพิสัยแบบ simple linear ในการอธิบายลำดับชั้นการเรียนรู้วิชาคณิตศาสตร์ของนักเรียนมัธยมศึกษาตอนต้นในกรุงเทพมหานคร โดยศึกษาและเปรียบเทียบการอธิบายลำดับชั้นการเรียนรู้ใน 3 รูปแบบ ได้แก่ simple linear model, Miller's model และ Wright's model จำแนกความสามารถทางพุทธิพิสัยเป็น 6 ด้าน ได้แก่ ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์และการประเมินค่า กลุ่มตัวอย่างที่ใช้ ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ในเขตกรุงเทพมหานคร จำนวน 642 คน เครื่องมือที่ใช้ ได้แก่ แบบสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ระดับมัธยมศึกษาตอนต้น วิเคราะห์ข้อมูลโดยใช้การวิเคราะห์เส้นทาง ประมาณค่าพารามิเตอร์แบบ Partial Least Squares ในการเปรียบเทียบรูปแบบโครงสร้างของลำดับชั้นการเรียนรู้ด้านพุทธิพิสัยแต่ละรูปแบบกับข้อมูลเชิงประจักษ์และรูปแบบโครงสร้างของลำดับชั้นการเรียนรู้ด้านพุทธิพิสัยแบบเต็มรูป ผลการวิจัยพบว่ารูปแบบโครงสร้างลำดับชั้นของการเรียนรู้ด้านพุทธิพิสัย แบบ simple linear model, แบบ

Miller's model และแบบ Wright's model สามารถอธิบายลำดับชั้นการเรียนรู้ของข้อมูลเชิงประจักษ์ได้ประมาณร้อยละ 54, 55 และ 57 ตามลำดับ นอกจากนี้ยังพบว่ารูปแบบลำดับชั้นของการเรียนรู้ที่มีการแยกแขนงของทิศทางสามารถอธิบายข้อมูลเชิงประจักษ์ได้ประมาณร้อยละ 73 หรือคิดเป็นร้อยละ 98 ของรูปแบบเต็มรูป

สุวัฒนา อุตย์รัตน์, สุชาวดี เขียมอรพรรณ และธรรณ พิษณานนท์ (2541) ได้วิเคราะห์ลำดับชั้นโครงสร้างเนื้อหาวิชาคณิตศาสตร์ในหลักสูตรมัธยมศึกษาตอนต้นเพื่อพัฒนาแบบฝึกหัดเสริมการเรียนรู้ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2, 3 และ 4 สังกัดกรมสามัญศึกษา จำนวน 2,192 คน ผลการวิจัยพบว่าลำดับชั้นเนื้อหาวิชาคณิตศาสตร์ตามหลักสูตรมัธยมศึกษาตอนต้นได้จัดเรียงไว้เป็นลำดับ โดยผลการวิจัยได้เสนอให้เปลี่ยนแปลงลำดับชั้นของเนื้อหา 3 เรื่อง ได้แก่ วิชา ค.101 เรื่องเส้นตรงและมุม จำนวน 4 ลำดับชั้น วิชา ค.102 เรื่องการนำเสนอข้อมูล จำนวน 5 ลำดับชั้นและวิชา ค.203 เรื่องระบบจำนวนเต็ม จำนวน 6 ลำดับชั้น ส่วนผลการพัฒนาแบบฝึกหัดเสริมการเรียนรู้พบว่าแบบฝึกหัดที่พัฒนาขึ้นมีประสิทธิภาพ 82.0/82.3

### **งานวิจัยเกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์เพื่อการประเมินเชิงวินิจฉัย**

Shermis และ Wolting (1996) ได้ใช้การทดสอบแบบปรับเหมาะในการประเมินเพื่อจัดนักเรียนเข้าสู่โปรแกรมการอ่านและให้ข้อมูลสำหรับผู้สอนเพื่อการปรับปรุงการอ่าน สำหรับนักศึกษาที่กำลังศึกษาในระดับปริญญาตรี (undergraduate) กลุ่มตัวอย่างที่ใช้ ได้แก่ นักศึกษาที่เรียนวิชาพัฒนาการอ่าน (development reading course) ซึ่งเป็นวิชาที่ใช้เพื่อปรับปรุงการอ่าน จำนวน 370 คน มีการพัฒนาเครื่องมือ 2 ขั้นตอน ขั้นตอนแรกพัฒนาเครื่องมือที่ใช้การเขียนตอบบนกระดาษคำตอบ (paper-pencil) แล้วจึงพัฒนาเป็นเครื่องมือที่ใช้การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ได้คลังข้อสอบจำนวน 153 ข้อ จำแนกเป็น 4 ส่วน ส่วนที่ 1 Comprehension จำนวน 50 ข้อ ส่วนที่ 2 Word Attack จำนวน 36 ข้อ ส่วนที่ 3 Word Context จำนวน 39 ข้อ และส่วนที่ 4 Word Isolation จำนวน 28 ข้อ การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ใช้ระบบการทดสอบแบบปรับเหมาะ Hyper CAT(TM) ซึ่งใช้ได้ทั้ง Apple Macintosh และ Win Plus (TM) for Windows-based computer โดยใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นพื้นฐานสำหรับวิธีการทดสอบแบบปรับเหมาะ ผลการวิจัยพบว่า การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์สามารถประหยัดเวลาในการสอบได้ร้อยละ 30-50 และเป็นวิธีการที่มีความตรงในการจัดนักเรียนเข้าโปรแกรม (placement) สูงกว่าการทดสอบโดยใช้การเขียนตอบบนกระดาษคำตอบ

Tatsuoka และ Tatsuoka (1997) ได้ใช้ Rule space model ในการทดสอบแบบปรับเหมาะ เพื่อพัฒนากระบวนการวินิจฉัยและการสอนซ่อมเสริม เรื่องปัญหาการบวกเศษส่วน โดยใช้ คอมพิวเตอร์ กลุ่มตัวอย่างที่ใช้ ได้แก่ นักเรียนเกรด 7 และเกรด 8 ของโรงเรียนมัธยมศึกษา ในเมืองเล็ก ๆ ในรัฐ Illinois ในปี ค.ศ. 1988 จำนวน 287 คน และในปี ค.ศ. 1989 จำนวน 191 คน การดำเนินการใช้กระบวนการที่ต่อเนื่องใน 2 ปี โดยเริ่มจากการทดสอบก่อนการซ่อมเสริม (pretest) เพื่อจัดนักเรียนตามสถานะความรู้ จำนวน 33 สถานะ โดยพิจารณาจากการตอบคำถาม และพิจารณาว่าใกล้เคียงกับสถานะความรู้ใดมากที่สุด การทดสอบจะหยุดสอบเมื่อผู้สอบถูกจัดอยู่ใน สถานะความรู้ใดสถานะความรู้หนึ่งอย่างคงที่ (stabilized on some point in rule space) หลังจากนั้น จึงจัดนักเรียนเพื่อรับการซ่อมเสริม โดยใช้เครื่องมือซ่อมเสริมในระบบ PLATO (a computer - based educational teacher system developed at University of Illinois) เลือกคุณลักษณะเพื่อใช้ในการซ่อมเสริมก่อนหลังโดยพิจารณาจากค่า Mahalanobis distances ( $D^2$ ) ตามเส้นทางรูป ต้นไม้ซึ่งแสดงสถานะความรู้ที่กำหนดไว้ เมื่อนักเรียนได้รับการฝึกหัดจนผ่านทุกคุณลักษณะใน สถานะความรู้ที่ได้รับการซ่อมเสริมแล้ว นักเรียนจะได้รับการทดสอบภายหลัง (posttest) และการ ทดสอบซ้ำ (retention test) ซึ่งดำเนินการทดสอบภายหลังจากการทดสอบภายหลัง 3 เดือน สำหรับนักเรียนที่ถูกจัดอยู่ในสถานะความรู้ที่ไม่ต้องได้รับการซ่อมเสริมไม่ต้องทดสอบภายหลัง และทดสอบซ้ำ ผลการวิจัยพบว่า Rule space model สามารถใช้ในการวินิจฉัยและการจัด เครื่องมือเพื่อการซ่อมเสริมได้อย่างมีประสิทธิภาพ

McGlohen, Chang และ Wills (2003) ได้ใช้การทดสอบแบบปรับเหมาะโดยใช้ คอมพิวเตอร์เพื่อประเมินเชิงวินิจฉัยทางพุทธิปัญญา (cognitive diagnostic assessment) โดยมี วัตถุประสงค์เพื่อรวมการประเมินระดับความสามารถของบุคคล โดยเน้นการให้ข้อมูลย้อนกลับเชิง วินิจฉัยจากเวกเตอร์คุณลักษณะของแต่ละบุคคล (individual attribute vector) ที่เชื่อมโยงกับ เทคโนโลยีการทดสอบแบบใหม่ซึ่งช่วยให้ครูและนักเรียนได้รับประโยชน์จากกระบวนการทดสอบ การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์เป็นแบบ shadow testing เลือกข้อสอบโดยใช้ ยุทธวิธี Minimization of Shannon Entropy ซึ่งเป็นยุทธวิธีในการทดสอบความไม่แน่นอนของ โอกาสความน่าจะเป็น โดยต้องการความไม่แน่นอนที่มีค่าต่ำสุด (minimal uncertainty) และ Maximization of Kullback-Leibler (K-L) Information ซึ่งเป็นยุทธวิธีที่ใช้ในการเลือกข้อสอบและ เป็นดัชนีในกระบวนการเลือกข้อสอบในการประเมินเชิงวินิจฉัย โดยใช้ค่าสารสนเทศ K-L ที่มี ค่าสูงสุด ออกแบบการเลือกข้อสอบเป็น 3 กรณี กรณีแรกเลือกข้อสอบจากการประมาณค่าระดับ ความสามารถ ( $\theta$ ) เท่านั้น กรณีที่ 2 เลือกข้อสอบจากการประมาณค่าเวกเตอร์คุณลักษณะ

(α) เท่านั้นและกรณีที่สามารถเลือกข้อสอบทั้งจากการประมาณค่าระดับความสามารถและการประมาณค่าเวกเตอร์คุณลักษณะ เลือกกลุ่มตัวอย่างได้จากการสุ่มอย่างง่าย จากการประเมินระดับกว้างของรัฐ (state-mandated large-scale assessment) จำนวน 2,000 คน คลังข้อสอบประกอบด้วยข้อสอบคณิตศาสตร์ จำนวน 394 ข้อ และข้อสอบการอ่าน จำนวน 324 ข้อ ประมาณค่าพารามิเตอร์ทั้งคะแนนและแบบแผนความรอบรู้ของคุณลักษณะ (attribute mastery patterns) โดยใช้ทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และ fusion model ด้วยโปรแกรม BilogMG และ Arpeggio 1.2 ข้อมูลที่ใช้เป็นข้อมูลจากการสมมติ (simulation data) แบบการแทนค่า (replacement) จากคะแนนจริงเป็น 6,000 คน และ 3,000 คน ผลการวิจัย พบว่า การใช้การประมาณค่าระดับความสามารถให้ความถูกต้องในการประเมินมากกว่าการใช้การประมาณค่าเวกเตอร์คุณลักษณะ การใช้ทั้งสองค่าให้ความถูกต้องในการประเมินมากที่สุดและข้อสอบที่มีอัตราการแสดงข้อสอบขั้นต่ำ (minimal item exposure rates) ให้การประมาณค่าได้อย่างถูกต้อง

Noye และ Fritz (2003) ได้ใช้เทคนิคการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์เพื่อวินิจฉัยระดับความสามารถของนักเรียนใน intelligent tutoring systems (ITS) โดยใช้วิธีการคัดเลือกข้อคำถามแบบปรับเหมาะวิธีใหม่ที่เรียกว่า Bayesian diagnostic algorithm ในการประมาณค่าระดับความสามารถในปัจจุบันของนักเรียน จำแนกคำถามในคลังข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์

Wen (2003) ได้ประยุกต์ใช้ Rule space model ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ โดยมีวัตถุประสงค์เพื่อศึกษาความสามารถในการประมาณค่าความสามารถของ Rule space model เมื่อมีองค์ประกอบของแบบทดสอบที่แตกต่างกันและวิธีการเลือกข้อสอบที่แตกต่างกัน องค์ประกอบของแบบสอบ ได้แก่ ความยาวของแบบสอบ, จำนวนคุณลักษณะในแบบทดสอบ, ลำดับขั้นความสัมพันธ์ระหว่างคุณลักษณะ, ธรรมชาติของข้อสอบ (ง่ายและซับซ้อน), จำนวนข้อสอบในคลังข้อสอบ, พารามิเตอร์การเดา (guessing) หรือพารามิเตอร์ข้อที่ผิด (slipness) และพารามิเตอร์ของข้อสอบ (1 พารามิเตอร์ และ 3 พารามิเตอร์) วิธีการเลือกข้อสอบที่แตกต่างกัน 4 วิธีใช้วิธีการเลือกอย่างสุ่ม (random selection method) ในการเรียงลำดับเพื่อการคำนวณและการเปรียบเทียบคุณลักษณะที่แท้จริงของผู้สอบ ใช้ Monte Carlo simulation method ในการศึกษา

Guzmán และ Conejo (2004) ได้พัฒนาโมเดลโดยใช้การเลือกคำตอบแบบมากกว่า 2 ค่า (polytomous) สำหรับวินิจฉัยความรู้ของนักเรียนใน Intelligent Tutoring System โดยใช้การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ซึ่งเป็นแบบทดสอบแบบเนื้อหาสมดุล (content - balanced test) ประมาณค่าความรู้ของนักเรียนโดยใช้ไค์การกระจายความน่าจะเป็นของข้อสอบ

ที่นักเรียนทำ คำนวณโดย Bayes theorem คัดเลือกข้อสอบข้อต่อไปโดยใช้ค่าความแปรปรวนของการกระจายความรู้ที่คาดหวังของนักเรียนที่มีค่าน้อยที่สุด (minimizes the expectation of a posteriori student knowledge distribution variance) วิเคราะห์ค่าพารามิเตอร์ของข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า (polytomous IRT) แบบ partial credit model โดยการสมมติข้อมูล (simulation data) เพื่อเปรียบเทียบจำนวนเฉลี่ยของข้อสอบที่ใช้ค่าเฉลี่ยของความคลาดเคลื่อนในการประมาณค่าความรู้ของนักเรียนและร้อยละของความถูกต้องในการประเมินนักเรียน สำหรับข้อสอบที่มีค่าอำนาจจำแนกแตกต่างกันระหว่างการประเมินที่ใช้การเลือกคำตอบแบบ 2 ค่า (dichotomous) และการเลือกคำตอบแบบมากกว่า 2 ค่า ผลการวิจัยพบว่า การเลือกข้อสอบแบบมากกว่า 2 ค่าให้ผลดีกว่าการเลือกคำตอบแบบ 2 ค่า (dichotomous) สำหรับข้อสอบที่มีค่าการกระจายต่ำ ทั้งจำนวนข้อสอบที่ใช้ ค่าความคลาดเคลื่อนในการประมาณค่าความรู้ของนักเรียนและความถูกต้องในการประเมิน โมเดลที่พัฒนาขึ้นให้ข้อมูลการตอบของนักเรียนได้มากกว่าโมเดลการเลือกคำตอบแบบ 2 ค่า และให้ผลการประเมินนักเรียนที่ถูกต้องโดยใช้ข้อคำถามน้อยลง

Gardner และคณะ (2004) ได้พัฒนา Beck Depression Inventory (BDI) เป็นฉบับที่ใช้การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์เพื่อวินิจฉัยและคัดแยกผู้ป่วยที่มีอาการซึมเศร้า โดยใช้เกณฑ์ช่วงความเชื่อมั่น (confidence interval : CI) รอบระดับความสามารถ ( $\theta$ ) ที่สูงกว่าหรือต่ำกว่าคะแนนเกณฑ์หรือคะแนนจุดตัด มีวัตถุประสงค์เพื่อทดสอบประสิทธิภาพและการใช้ประโยชน์ในการจำแนก Major Depression Episodes (MDE) และการวัดระดับความรุนแรงของความเศร้าซึม กลุ่มตัวอย่างเป็นผู้ที่ทำ BDI และ Structured Clinical Interview for DSM-IV Axis Disorder (SCID) จำนวน 744 คน และผู้ป่วยที่ทำ Hamilton Depression Rating Scale จำนวน 285 คน วิเคราะห์ค่าพารามิเตอร์ของข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า (polytomous IRT) แบบ graded-response model โดยการสมมติข้อมูล (simulation data) ข้อสอบข้อแรกเป็นข้อที่ให้สารสนเทศสูงสุดของคะแนนความซึมเศร้าโดยเฉลี่ยของกลุ่มประชากร เลือกข้อสอบข้อต่อไปโดยพิจารณาจากข้อสอบที่ให้ค่าสารสนเทศสูงสุด ผลการวิจัยพบว่า BDI ฉบับที่ใช้การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มีพื้นที่ใต้โค้ง (Area Under the Curve : AUC) ซึ่งเป็นตัวบ่งชี้ของ SCID ในการวินิจฉัย MDE เท่ากับ 88 % เท่า ๆ กับ BDI ฉบับเต็ม แต่ใช้คำถามน้อยกว่า (5.6 กับ 21 ข้อ) นอกจากนี้ยังพบว่า คะแนนความซึมเศร้าของ BDI ฉบับที่ใช้การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มีค่าสัมประสิทธิ์สหสัมพันธ์กับคะแนนรวมของ BDI และคะแนนความซึมเศร้าของ

Hamilton เท่ากับ .92 และ .74 ตามลำดับ แสดงให้เห็นว่า BDI ฉบับที่ใช้การทดสอบแบบปรับเหมาะ โดยใช้คอมพิวเตอร์มีประสิทธิภาพมากขึ้นโดยไม่ทำให้ความถูกต้องในการประเมินน้อยลง

จากการศึกษางานวิจัยข้างต้นจะเห็นได้ว่าการนำการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ไปใช้เพื่อการประเมินเชิงวินิจฉัยมีการดำเนินการยังไม่แพร่หลาย โดยเฉพาะในประเทศไทยไม่พบเอกสารงานวิจัยเกี่ยวกับการดำเนินการดังกล่าว ดังนั้นการประยุกต์ใช้การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์เพื่อการประเมินเชิงวินิจฉัยจึงยังต้องการการดำเนินการที่หลากหลาย เพื่อให้ได้ข้อสรุปที่ชัดเจนขึ้น

### บทที่ 3

#### วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีการประเมินเชิงวินิจฉัย โดยการประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ในการประเมินสถานะความรู้และแบบการคิดที่ผิด เรื่อง การบวกลบเศษส่วน ของนักเรียนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) รวมทั้งเพื่อศึกษาคุณภาพและเปรียบเทียบคุณภาพ ตลอดจนศึกษาความคิดเห็นที่มีต่อวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นดังกล่าว มีรายละเอียดของประชากรกลุ่มตัวอย่าง เครื่องมือที่ใช้ในการวิจัย ขั้นตอนและวิธีดำเนินการวิจัย ดังนี้

#### ประชากรและกลุ่มตัวอย่าง

##### 1. ประชากร

ประชากรในการวิจัย จำแนกเป็น 3 กลุ่ม ดังนี้

นักเรียนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6)

ครูผู้สอนวิชาคณิตศาสตร์ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6)

ผู้เชี่ยวชาญ ประกอบด้วย ผู้เชี่ยวชาญด้านหลักสูตรคณิตศาสตร์ ผู้เชี่ยวชาญด้านการเรียนการสอนคณิตศาสตร์ และผู้เชี่ยวชาญด้านการวัดและประเมินผล

## 2. กลุ่มตัวอย่าง

กลุ่มตัวอย่างในการวิจัย จำแนกเป็น 10 กลุ่ม ดังนี้

กลุ่มตัวอย่างในขั้นตอนการสำรวจคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) และขั้นตอนการจัดเรียงลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ได้แก่ ครูต้นแบบ วิชาคณิตศาสตร์ ระดับประถมศึกษาของสำนักงานคณะกรรมการการศึกษาแห่งชาติ (สกศ.) ตั้งแต่ปี 2541 - 2544 จำนวน 57 คน

กลุ่มตัวอย่างในขั้นตอนการประเมินความถูกต้องของคุณลักษณะที่ใช้ในการบวกลบเศษส่วนและขั้นตอนการประเมินแผนภาพลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ได้แก่ ผู้เชี่ยวชาญด้านหลักสูตรคณิตศาสตร์หรือการเรียนการสอนคณิตศาสตร์ จำนวน 3 คน โดยมีเกณฑ์ในการคัดเลือกผู้เชี่ยวชาญ ดังนี้

มีวุฒิการศึกษา ด้านหลักสูตรคณิตศาสตร์หรือการเรียนการสอนคณิตศาสตร์

เป็นผู้ที่ปฏิบัติงานเกี่ยวกับหลักสูตรคณิตศาสตร์หรือการเรียนการสอนคณิตศาสตร์ เรื่อง การบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) อย่างน้อย 5 ปี

กลุ่มตัวอย่างในการศึกษานำร่องเพื่อจำแนกแบบการคิดของนักเรียน ได้แก่ นักเรียน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) ที่เรียนเรื่องการบวกลบเศษส่วนแล้ว จากโรงเรียนขนาดเล็ก (จำนวนนักเรียนต่ำกว่า 200 คน) โรงเรียนขนาดกลาง (จำนวนนักเรียนระหว่าง 201 – 500 คน) และโรงเรียนขนาดใหญ่ (จำนวนนักเรียนมากกว่า 500 คน) ขนาดโรงเรียนละ 70 คน รวมจำนวน 210 คน

กลุ่มตัวอย่างในการตรวจสอบความครบถ้วนและความครอบคลุมของแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน ได้แก่ ครูผู้สอนวิชาคณิตศาสตร์จากโรงเรียนแกนนำการเรียนการสอนคณิตศาสตร์ ระดับประถมศึกษา ของสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) เขตการศึกษาละ 1 โรงเรียน โรงเรียนละ 1 คน รวมจำนวน 12 คน

กลุ่มตัวอย่างในการทดลองแบบทดสอบแบบเขียนตอบบนกระดาษคำตอบ ได้แก่ นักเรียน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 ที่เรียน เรื่องการบวกลบเศษส่วนแล้ว จาก

โรงเรียนขนาดเล็ก โรงเรียนขนาดกลาง และโรงเรียนขนาดใหญ่ ขนาดละ 6 โรงเรียน รวมจำนวน 18 โรงเรียน โรงเรียนละ 100 คน รวม จำนวน 1,800 คน

กลุ่มตัวอย่างในการทดลองใช้โปรแกรมคอมพิวเตอร์ ได้แก่ นักเรียน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) จากโรงเรียนขนาดเล็ก โรงเรียนขนาดกลางและโรงเรียนขนาดใหญ่ ซึ่งมีลักษณะเหมือนกลุ่มตัวอย่างในการประเมินเชิงวิจักษ์ สถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วน ขนาดโรงเรียนละ 18 คน จำนวน 54 คน (3 คน×3 ขนาดโรงเรียน×6 โปรแกรมย่อย)

กลุ่มตัวอย่างในการระบุปัญหาที่เกิดจากการทดลองใช้โปรแกรมคอมพิวเตอร์ ได้แก่ กลุ่มตัวอย่างกลุ่มที่ 2.6 และครูผู้สอนวิชาคณิตศาสตร์ ระดับการศึกษาขั้นพื้นฐานช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) ที่รับผิดชอบดำเนินการจัดให้นักเรียนได้ทดลองใช้โปรแกรมคอมพิวเตอร์

กลุ่มตัวอย่างที่ได้รับการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ได้แก่ นักเรียน ระดับการศึกษาขั้นพื้นฐานช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน จำนวน 540 คน เลือกลุ่มตัวอย่างโดยใช้การสุ่มแบบหลายขั้นตอน (multi - stage random sampling) มีการดำเนินการ ดังนี้

ขั้นตอนที่ 1 แบ่งประชากรออกเป็นชั้นตามภาคของประเทศไทย จำนวน 6 ภาค ได้แก่ ภาคเหนือ ภาคกลาง ภาคใต้ ภาคตะวันออกเฉียงเหนือ ภาคตะวันออกและภาคตะวันตก สุ่มภาคโดยใช้การสุ่มอย่างง่าย จำนวน 2 ภาค ได้ภาคกลางและภาคตะวันออกเฉียงเหนือ

ขั้นตอนที่ 2 สุ่มจังหวัดจากแต่ละภาค โดยใช้การสุ่มอย่างง่าย ภาคละ 1 จังหวัด ได้จังหวัดประจวบคีรีขันธ์และจังหวัดชัยภูมิ

ขั้นตอนที่ 3 สุ่มสำนักงานเขตพื้นที่การศึกษาจากแต่ละจังหวัด โดยใช้การสุ่มอย่างง่าย จังหวัดละ 1 แห่ง ได้สำนักงานเขตพื้นที่การศึกษาจังหวัดประจวบคีรีขันธ์ เขต 2 และสำนักงานเขตพื้นที่การศึกษาจังหวัดชัยภูมิ เขต 1

ขั้นตอนที่ 4 แบ่งโรงเรียนในแต่ละสำนักงานเขตพื้นที่การศึกษาเป็น 3 ขนาด ได้แก่ โรงเรียนขนาดเล็ก มีจำนวนนักเรียนต่ำกว่า 200 คน โรงเรียนขนาดกลาง มีจำนวนนักเรียน 201 ถึง 500 คน และโรงเรียนขนาดใหญ่ มีจำนวนนักเรียนมากกว่า 500 คน สุ่มโรงเรียนจากแต่ละขนาดในแต่ละสำนักงานเขตพื้นที่การศึกษา โดยใช้การสุ่มอย่างง่าย ขนาดละ 3 โรงเรียน ได้จำนวน 18 โรงเรียน

ขั้นตอนที่ 5 สุ่มนักเรียน ระดับการศึกษาขั้นพื้นฐาน (ชั้นประถมศึกษาปีที่ 4– 6) จากแต่ละโรงเรียน โดยใช้การสุ่มอย่างง่าย โรงเรียนละ 30 คน ได้จำนวน 540 คน



กลุ่มตัวอย่างในการระบุปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์และการศึกษาความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น ได้แก่ กลุ่มตัวอย่างกลุ่มที่ 2.8 และครูผู้สอนวิชาคณิตศาสตร์ ระดับการศึกษาขั้นพื้นฐานช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) ที่รับผิดชอบดูแลการดำเนินการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

กลุ่มตัวอย่างที่ได้รับการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วน โดยใช้เทคนิคการคิดออกเสียง ได้แก่ นักเรียน ระดับการศึกษาขั้นพื้นฐานช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 - 6) โดยใช้ในการสุ่มอย่างง่ายจากกลุ่มตัวอย่างกลุ่มที่ 2.8 ขนาดโรงเรียนละ 18 คน รวมจำนวน 54 คน

## เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย จำแนกเป็น 6 ส่วน ดังนี้

1. เครื่องมือที่ใช้ในการกำหนดลำดับชั้นของคุณลักษณะ ได้แก่

แบบสำรวจคุณลักษณะที่ใช้ในการบวกลบเศษส่วน จำแนกเป็น 2 ตอน ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสำรวจ เป็นแบบตรวจสอบรายการและเติมคำ เพื่อตรวจสอบสภาพและข้อมูลบางประการของผู้ตอบแบบสำรวจ ตอนที่ 2 คุณลักษณะที่ใช้ในการบวกลบเศษส่วนเป็นแบบเขียนตอบเพื่อให้กลุ่มตัวอย่างจำแนกคุณลักษณะย่อยที่ใช้ในการบวกลบเศษส่วนสำหรับแต่ละคุณลักษณะหลักได้อย่างอิสระ

แบบประเมินความถูกต้องของคุณลักษณะที่ใช้ในการบวกลบเศษส่วนมีลักษณะเป็นแบบตรวจสอบรายการเพื่อให้กลุ่มตัวอย่างพิจารณาและตัดสินความถูกต้องของคุณลักษณะหลักและคุณลักษณะย่อยที่ใช้ในการบวกลบเศษส่วน พร้อมทั้งให้เหตุผลและ/หรือข้อเสนอแนะสำหรับแต่ละคุณลักษณะที่ละข้อ

แบบเปรียบเทียบลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน สร้างตามแนวคิดเทคนิคกระบวนการให้น้ำหนักรายคู่ (paired – weighting procedure) ซึ่งเป็นเทคนิค

การประเมินความต้องการจำเป็นที่อยู่ในกลุ่มเทคนิควิธีจัดเรียงลำดับ (rank หรือ setting priorities) มีลักษณะเป็นแบบเปรียบเทียบรายคู่เพื่อให้กลุ่มตัวอย่างพิจารณาและตัดสินลักษณะความสัมพันธ์เชิงลำดับชั้นของคุณลักษณะที่ละคู่

แบบประเมินความถูกต้องของแผนภาพลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน มีลักษณะเป็นแบบตรวจสอบรายการเพื่อให้กลุ่มตัวอย่างพิจารณาและตัดสินความถูกต้องของแผนภาพลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน และบันทึกเหตุผลและ/หรือข้อเสนอแนะเพื่อการปรับปรุงแก้ไข

## 2. เครื่องมือที่ใช้ในการจัดทำชุดของข้อสอบ ได้แก่

2.1 แบบทดสอบเพื่อสำรวจแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน มีลักษณะเป็นแบบทดสอบแบบแสดงวิธีทำ คุณลักษณะละ 4 ข้อ รวมจำนวน 16 ข้อ

2.2 แบบประเมินความครบถ้วนและความครอบคลุมของแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน จำแนกเป็น 2 ตอน ตอนที่ 1 ข้อมูลทั่วไปของผู้ประเมินเป็นแบบตรวจสอบรายการเพื่อสำรวจสถานภาพและข้อมูลบางประการของผู้ประเมิน ตอนที่ 2 ความครบถ้วนและความครอบคลุมของแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน เป็นแบบตรวจสอบรายการเพื่อให้ผู้ประเมินพิจารณาแบบการคิดที่ผิดสำหรับแต่ละคุณลักษณะและตัดสินความครบถ้วนและความครอบคลุมของแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนในแต่ละคุณลักษณะ พร้อมทั้งระบุแบบการคิดที่ผิดเพิ่มเติมกรณีเห็นว่าแบบการคิดที่ผิดที่กำหนดให้ไม่ครบถ้วนและครอบคลุม

3. เครื่องมือที่ใช้ในการจัดทำคลังข้อสอบ คือ แบบทดสอบสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนเป็นแบบทดสอบแบบเติมคำ จำนวน 20 ฉบับ ฉบับละ 32 ข้อ

4. เครื่องมือที่ใช้ในการทดลองใช้โปรแกรมคอมพิวเตอร์และการประเมินเชิงวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน ได้แก่

4.1 โปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะและคู่มือการใช้โปรแกรมซึ่งจำแนกเป็น 3 ส่วน ส่วนที่ 1 บทนำ ส่วนที่ 2 โปรแกรมย่อยการพิมพ์ข้อสอบและส่วนที่ 3 โปรแกรมย่อยการทดสอบ

4.2 แบบสอบถามปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ จำแนกเป็น 2 ตอน ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถามเป็นแบบตรวจสอบรายการเพื่อสำรวจสถานภาพและข้อมูลบางประการของผู้ตอบแบบสอบถาม ตอนที่ 2 ปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ เป็นแบบตรวจสอบรายการ แบบเขียนตอบและแบบมาตราส่วนประมาณค่า

เพื่อให้กลุ่มตัวอย่างตรวจสอบปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ในด้านต่าง ๆ พร้อมทั้ง ระบุสภาพปัญหาและระดับความถี่ของการเกิดปัญหา จำแนกเป็น 2 ฉบับ ฉบับครูและ ฉบับนักเรียน ทั้ง 2 ฉบับมีเนื้อหาเหมือนกัน แตกต่างกันที่รูปแบบในการตอบ

5. เครื่องมือที่ใช้ในการประเมินเชิงวิจิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน โดยใช้เทคนิคการคิดออกเสียง ได้แก่ คู่มือการประเมินเชิงวิจิจฉัยสถานะความรู้ และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนโดยใช้เทคนิคการคิดออกเสียง จำแนกเป็น 3 ส่วน ส่วนที่ 1 บทนำ ส่วนที่ 2 มโนทัศน์เบื้องต้นเกี่ยวกับเทคนิคการคิดออกเสียง และส่วนที่ 3 การดำเนินการประเมินเชิงวิจิจฉัยโดยใช้เทคนิคการคิดออกเสียง มีเครื่องมือประกอบการใช้คู่มือ ดังนี้

5.1 แบบทดสอบ เรื่อง การบวกลบเศษส่วน ใช้เป็นเครื่องมือในการวิจิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่นักเรียนใช้ในการบวกลบเศษส่วนมีจำนวน 4 ฉบับ ฉบับที่ 1 การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน ฉบับที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน ฉบับที่ 3 การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน และฉบับที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน แบบทดสอบทั้ง 4 ฉบับ มีข้อสอบ จำนวนฉบับละ 10 ข้อ คัดเลือกข้อสอบจากคลังข้อสอบในแต่ละลำดับชั้น โดยพิจารณาคัดเลือกให้มีค่าความยากของข้อสอบครอบคลุมทุกระดับความสามารถ

5.2 แบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 -6) ใช้ประกอบการพิจารณาตัดสินผลการประเมิน จำแนกเป็น 4 ส่วน ส่วนที่ 1 แบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน ส่วนที่ 2 แบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน ส่วนที่ 3 แบบการคิดที่ผิดที่ใช้ในการบวกลบจำนวนคละที่มีตัวส่วนเท่ากันและส่วนที่ 4 แบบการคิดที่ผิดที่ใช้ในการบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน

5.3 แบบบันทึกกระบวนการคิดแก้ปัญหา ในการประเมินเชิงวิจิจฉัยสถานะความรู้ และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 - 6) ใช้ในการบันทึกผลการประเมิน จำแนกเป็น 3 ตอน ตอนที่ 1 ข้อมูลทั่วไปของผู้บันทึก ตอนที่ 2 ข้อมูลทั่วไปของผู้ประเมินและผู้ได้รับการประเมิน และตอนที่ 3 สรุปกระบวนการคิดแก้ปัญหา สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน

6. เครื่องมือที่ใช้ในการศึกษาความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจิจฉัยที่พัฒนาขึ้น ได้แก่ แบบสอบถามความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจิจฉัย โดยประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ มีจำนวน 2 ฉบับ ฉบับครูและฉบับนักเรียน ทั้ง 2 ฉบับ มีลักษณะของเครื่องมือเหมือนกัน ข้อคำถามมีวัตถุประสงค์เดียวกัน แต่ต่างกันที่การใช้ภาษา แบบสอบถามจำแนกเป็น 2 ตอน ดังนี้

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม เป็นแบบตรวจสอบรายการและแบบเติมคำ เพื่อตรวจสอบสถานภาพและข้อมูลบางประการของผู้ตอบแบบสอบถาม

ตอนที่ 2 ความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจจัย โดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เป็นคำถามปลายเปิดเพื่อให้ผู้ตอบแบบสอบถามแสดงความคิดเห็นได้อย่างอิสระ มีจำนวน 5 ข้อ

ตัวอย่างเครื่องมือที่ใช้ในการวิจัยแสดงในภาคผนวก ข

### **ขั้นตอนและวิธีดำเนินการวิจัย**

การพัฒนาวิธีการประเมินเชิงวิจจัย โดยการประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ มีการดำเนินการวิจัย จำแนกเป็น 4 ระยะเวลา  
ระยะที่ 1 การพัฒนาชุดของข้อสอบ แบบทดสอบแบบเขียนตอบบนกระดาษคำตอบและคลังข้อสอบ  
ระยะที่ 2 การพัฒนาโปรแกรมคอมพิวเตอร์เพื่อใช้ในการทดสอบแบบปรับเหมาะ  
ระยะที่ 3 การประเมินเชิงวิจจัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน และระยะที่ 4 การศึกษาคุณภาพและความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจจัยที่พัฒนาขึ้น แต่ระยะนี้มีรายละเอียดขั้นตอนและวิธีการดำเนินการวิจัย ดังนี้

#### **ระยะที่ 1 การพัฒนาชุดของข้อสอบ แบบทดสอบแบบเขียนตอบบนกระดาษคำตอบและคลังข้อสอบ**

การพัฒนาชุดของข้อสอบ แบบทดสอบแบบเขียนตอบบนกระดาษคำตอบและคลังข้อสอบ มีการดำเนินการ 4 ขั้นตอน ได้แก่ การกำหนดลำดับขั้นของคุณลักษณะ การจัดทำชุดของข้อสอบ การจัดทำแบบทดสอบแบบเขียนตอบบนกระดาษคำตอบและการจัดทำคลังข้อสอบ แต่ขั้นตอนมีการดำเนินการ ดังนี้

##### **ขั้นตอนที่ 1 การกำหนดลำดับขั้นของคุณลักษณะ**

การกำหนดลำดับขั้นของคุณลักษณะเป็นการจำแนกและระบุคุณลักษณะที่ใช้ในการบวกลบเศษส่วนและนำมาจัดเรียงลำดับก่อนหลังโดยแสดงเป็นแผนภาพเพื่อใช้ในการกำหนดลักษณะเฉพาะของข้อสอบและแบบแผนการตอบข้อสอบที่คาดหวัง มีวัตถุประสงค์และวิธีดำเนินการ ดังนี้

1. ศึกษาเอกสารและงานวิจัยเกี่ยวกับการบวกลบเศษส่วนเพื่อวิเคราะห์คุณลักษณะหลักที่ใช้ในการบวกลบเศษส่วน ดำเนินการโดยวิเคราะห์เนื้อหาเกี่ยวกับการเรียนการสอนเรื่องการ

บวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) จากเอกสาร หลักสูตร หนังสือ ตำราเรียนและงานวิจัยที่เกี่ยวข้อง นำเนื้อหาที่ได้มาสังเคราะห์และสรุปเป็น คุณลักษณะหลักที่ใช้ในการบวกลบเศษส่วน

2. สํารวจคุณลักษณะย่อยที่ใช้ในการบวกลบเศษส่วน เพื่อระบุและจำแนกคุณลักษณะย่อย ที่ใช้ในการบวกลบเศษส่วนสำหรับแต่ละคุณลักษณะหลักดำเนินการโดยให้กลุ่มตัวอย่างกลุ่มที่ 2.1 ตอบแบบสำรวจคุณลักษณะที่ใช้ในการบวกลบเศษส่วน เก็บรวบรวมข้อมูลโดยส่งแบบสำรวจให้ กลุ่มตัวอย่างทางไปรษณีย์ นำแบบสำรวจที่ได้รับกลับ มาตรวจสอบความถูกต้องสมบูรณ์ของ ข้อมูลและวิเคราะห์ข้อมูลซึ่งจำแนกเป็นข้อมูลสภาพทั่วไปของผู้ตอบแบบสำรวจจากแบบสำรวจ ตอนที่ 1 วิเคราะห์ด้วยสถิติบรรยาย ได้แก่ ความถี่และร้อยละ โดยใช้โปรแกรมสำเร็จรูป SPSS/PC และข้อมูลคุณลักษณะที่ใช้ในการบวกลบเศษส่วน จากแบบสำรวจตอนที่ 2 วิเคราะห์โดยการ วิเคราะห์เนื้อหาและสรุปเป็นความเรียงเพื่อบรรยายคุณลักษณะย่อยสำหรับแต่ละคุณลักษณะหลัก

3. ประเมินความถูกต้องของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน เพื่อตรวจสอบ ความถูกต้องของคุณลักษณะหลักและคุณลักษณะย่อยที่ใช้ในการบวกลบเศษส่วน ดำเนินการโดย ให้กลุ่มตัวอย่างกลุ่มที่ 2.2 พิจารณาและตัดสินความถูกต้องของคุณลักษณะจากแบบประเมิน ความถูกต้องของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน เก็บรวบรวมข้อมูลโดยนำแบบประเมิน ไปให้กลุ่มตัวอย่างด้วยตนเอง วิเคราะห์ข้อมูลโดยตรวจสอบความถูกต้องสมบูรณ์ของข้อมูลและ คำนวณค่าสัดส่วนของความสอดคล้อง คุณลักษณะที่มีค่าสัดส่วนของความสอดคล้องต่ำกว่า 2 ใน 3 จะพิจารณาข้อเสนอแนะเพื่อปรับปรุงแก้ไขหรือตัดออก ส่วนคุณลักษณะที่มีสัดส่วนของ ความสอดคล้องมากกว่าหรือเท่ากับ 2 ใน 3 ถือว่าเป็นคุณลักษณะในการบวกลบเศษส่วนที่ถูกต้อง

3. จัดเรียงลำดับชั้นของคุณลักษณะเพื่อกำหนดลำดับก่อนและหลังของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน โดยพิจารณาจากความเป็นคุณลักษณะที่ต้องมีมาก่อน (prerequisite) ดำเนินการโดยให้กลุ่มตัวอย่าง กลุ่มที่ 2.1 เปรียบเทียบลำดับชั้นของคุณลักษณะที่ละคู่ตาม ลักษณะความสัมพันธ์ที่กำหนดในแบบแบบเปรียบเทียบลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบ เศษส่วน เก็บรวบรวมข้อมูลโดยส่งแบบเปรียบเทียบลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบ เศษส่วนให้กลุ่มตัวอย่างทางไปรษณีย์ วิเคราะห์ข้อมูลโดยตรวจสอบความถูกต้องสมบูรณ์ของ ข้อมูล คำนวณค่าความถี่และร้อยละของแต่ละคุณลักษณะในแต่ละลักษณะความสัมพันธ์ รวม คะแนนสำหรับแต่ละคุณลักษณะในแต่ละลักษณะความสัมพันธ์ของแต่ละบุคคล จากนั้นจึงนำคะแนน ของแต่ละบุคคลมารวมกันเป็นคะแนนของกลุ่ม แล้วจึงนำคะแนนของกลุ่มมาจัดเรียงลำดับชั้นของ คุณลักษณะจากคุณลักษณะต่ำสุดซึ่งเป็นคุณลักษณะที่ต้องมีมาก่อนทุกคุณลักษณะไปหา คุณลักษณะสูงสุดซึ่งเป็นคุณลักษณะที่ไม่ต้องมีมาก่อนคุณลักษณะใด ๆ โดยใช้คะแนนรวมของ

ลักษณะความสัมพันธ์กรณีเป็นคุณลักษณะที่ไม่เกี่ยวข้องสัมพันธ์กันและกรณีเป็นคุณลักษณะในลำดับชั้นเดียวกันประกอบการพิจารณา

4. จัดทำแผนภาพลำดับชั้นของคุณลักษณะเพื่อกำหนดแผนภาพซึ่งแสดงความสัมพันธ์เชิงลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ดำเนินการโดยนำคุณลักษณะที่จัดเรียงลำดับชั้นแล้วมาเขียนเป็นแผนภาพลำดับชั้นของคุณลักษณะ และให้กลุ่มตัวอย่างกลุ่มที่ 2.2 ประเมินความถูกต้องของแผนภาพ เก็บรวบรวมข้อมูลโดยนำแบบประเมินความถูกต้องของแผนภาพไปให้กลุ่มตัวอย่างด้วยตนเอง วิเคราะห์ข้อมูลโดยตรวจสอบความถูกต้องสมบูรณ์ของข้อมูล คำนวณค่าสัดส่วนของความสอดคล้อง และปรับปรุงแผนภาพลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วนตามคำแนะนำของกลุ่มตัวอย่าง จนกระทั่งมีสัดส่วนของความสอดคล้องของผลการประเมินมากกว่าหรือเท่ากับ 2 ใน 3

## ขั้นตอนที่ 2 การจัดทำชุดของข้อสอบ

การจัดทำชุดของข้อสอบเป็นการดำเนินการเพื่อให้ได้ข้อสอบและแนวทางการวินิจฉัยคำตอบ ซึ่งประกอบด้วยสถานะความรู้และแบบการคิดที่ผิดในแต่ละคุณลักษณะสำหรับแต่ละแบบแผนการตอบข้อสอบที่คาดหวัง ซึ่งจะใช้สำหรับจัดทำแบบทดสอบแบบเขียนตอบบนกระดาษคำตอบ และเป็นฐานข้อมูลของข้อสอบในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ มีวัตถุประสงค์และวิธีดำเนินการ ดังนี้

1. กำหนด binary adjacency matrix (A) เพื่อแสดงความสัมพันธ์ทางตรงระหว่างคุณลักษณะ
2. กำหนด reachability matrix (R) เพื่อแสดงความสัมพันธ์ทางตรงและความสัมพันธ์ทางอ้อมระหว่างคุณลักษณะ ซึ่งบ่งบอกถึงเงื่อนไขในโครงสร้างของลำดับชั้น เพื่อประโยชน์ในการกำหนดชุดย่อยของข้อคำถาม
3. กำหนด incidence matrix (Q) เพื่อแสดงชุดของข้อสอบและคุณลักษณะที่ต้องใช้ในการทำข้อสอบในแต่ละข้อให้ถูกต้อง
4. กำหนด reduced Q matrix (Q<sub>r</sub>) เพื่อลดจำนวนชุดของข้อสอบที่เป็นไปได้ตามเงื่อนไขโครงสร้างลำดับชั้นที่กำหนด
5. กำหนดแบบแผนการตอบข้อสอบที่คาดหวัง (expected item response pattern) และคุณลักษณะของผู้ตอบในแต่ละแบบแผนการตอบ
6. กำหนดแบบการคิดสำหรับแต่ละคุณลักษณะในแต่ละลำดับชั้น มีขั้นตอนดำเนินการ ดังนี้

6.1 ศึกษา นำร่องเพื่อจำแนกแบบการคิดที่ผิดของนักเรียน โดยให้กลุ่มตัวอย่าง กลุ่มที่ 2.3 ทำแบบทดสอบเพื่อสำรวจแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน โดยประสานงานกับ โรงเรียนเพื่อขอดำเนินการทดสอบด้วยตนเองหรือขอความร่วมมือในการดำเนินการทดสอบจากครู ประจำชั้นหรือครูที่รับผิดชอบดูแลกลุ่มตัวอย่าง วิเคราะห์ข้อมูลโดยนำผลการตอบมาตรวจให้คะแนน และวิเคราะห์จำแนกแบบการคิดที่ผิดที่นักเรียนใช้ในการบวกลบเศษส่วน

6.2 ประเมินความครบถ้วนและความครอบคลุมของแบบการคิดที่ผิดที่ใช้ในการ บวกลบเศษส่วน โดยให้กลุ่มตัวอย่าง กลุ่มที่ 2.4 ตอบประเมินความครบถ้วนและความครอบคลุม ของแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน เก็บรวบรวมข้อมูลโดยส่งแบบประเมินให้กลุ่มตัวอย่าง ทางไปรษณีย์ วิเคราะห์ข้อมูลโดยนำผลการประเมินที่ได้มาตรวจสอบความถูกต้อง สมบูรณ์ของ ข้อมูล และคำนวณค่าความถี่และร้อยละของผลการประเมิน

7. เขียนข้อคำถามสำหรับคุณลักษณะในแต่ละลำดับชั้น คุณลักษณะละ 122 ข้อ คำนวณหาคำตอบถูกและคำตอบผิดสำหรับแต่ละแบบการคิดที่ผิดในแต่ละคุณลักษณะ การ กำหนดจำนวนข้อสอบดังกล่าว คำนึงถึงขนาดที่พอเหมาะตามเกณฑ์จำนวนข้อสอบในคลังข้อสอบ ตามข้อเสนอแนะของ Weiss (1988, อ้างถึงใน รังสรรค์ มณีเล็ก, 2540) ซึ่งแนะนำให้จำนวน ข้อสอบในคลังข้อสอบ จำนวน 100 – 200 ข้อ และจากข้อค้นพบของ Urry (1970,1977 อ้างถึงใน รังสรรค์ มณีเล็ก, 2540) ที่พบว่าคลังข้อสอบควรมีข้อสอบอย่างน้อย 100 ข้อ

8. เขียนคำบรรยายเพื่อประเมินเชิงวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ใน การบวกลบเศษส่วนในแต่ละแบบแผนการตอบข้อสอบที่คาดหวัง

### ขั้นตอนที่ 3 การจัดทำแบบทดสอบแบบเขียนตอบบนกระดาษคำตอบ

การจัดทำแบบทดสอบแบบเขียนตอบบนกระดาษคำตอบเป็นการดำเนินการเพื่อให้ได้ แบบทดสอบสถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วนซึ่งจะใช้ผลการทดลอง แบบทดสอบเป็นฐานข้อมูลในโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะะ มี วัตถุประสงค์และวิธีดำเนินการ ดังนี้

1. เลือกข้อสอบแต่ละคุณลักษณะจากชุดของข้อสอบที่พัฒนาขึ้น ในขั้นตอนที่ 2 เพื่อ เป็นข้อสอบร่วม (anchor item) ในแต่ละฉบับ จำนวนคุณลักษณะละ 2 ข้อ และเพื่อเป็นข้อสอบเฉพาะ ในแต่ละฉบับ จำนวนคุณลักษณะละ 6 ข้อ

2. จัดฉบับของแบบทดสอบ ซึ่งประกอบด้วยข้อสอบในแต่ละคุณลักษณะ จำนวน 8 ข้อ จำแนกเป็นข้อสอบเฉพาะในแต่ละฉบับ จำนวน 6 ข้อ และข้อสอบร่วมในทุกฉบับ จำนวน 2 ข้อ ซึ่งจะได้

แบบทดสอบแบบเติมคำ จำนวน 20 ฉบับ ฉบับละ 32 ข้อ (ข้อสอบในแต่ละชุดของข้อสอบสำหรับแต่ละคุณลักษณะมีจำนวน 122 ข้อ)

3. นำแบบทดสอบให้ผู้เชี่ยวชาญด้านหลักสูตรการเรียนการสอนคณิตศาสตร์และด้านการวัดและประเมินผลตรวจสอบความตรงตามเนื้อหา และปรับปรุงแบบทดสอบตามคำแนะนำของผู้เชี่ยวชาญ

4. นำแบบทดสอบทั้ง 20 ฉบับ ไปทดลองใช้กับกลุ่มตัวอย่าง กลุ่มที่ 2.5 โดยจัดให้นักเรียนจากแต่ละขนาดโรงเรียนได้รับการทดสอบฉบับละ 30 คน รวมเป็นฉบับละ 90 คน

5. วิเคราะห์ค่าพารามิเตอร์ของข้อสอบและคุณภาพของแบบทดสอบ ดังนี้

5.1 วิเคราะห์ค่าพารามิเตอร์ของข้อสอบโดยใช้โมเดลการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ ได้แก่ ค่าอำนาจจำแนกของข้อสอบ (a-Parameter) ค่าความยากของข้อสอบ (b-Parameter) และค่าการเดาข้อสอบ (c-Parameter) คำนวณโดยใช้โปรแกรม ITEM ANALYSIS SYSTEM: IAS 2006 พัฒนาโดย ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี และรองศาสตราจารย์ ดร. คณิต ไช่มุกด์

5.2 วิเคราะห์ความเที่ยงของแบบทดสอบโดยวิธีตรวจสอบความสอดคล้องภายใน (Internal consistency method) ด้วยวิธีของคูเดอร์-ริชาร์ดสัน (kuder-richardson estimates) คำนวณโดยใช้โปรแกรม ITEM ANALYSIS SYSTEM: IAS 2006 พัฒนาโดย ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี และรองศาสตราจารย์ ดร. คณิต ไช่มุกด์

#### ขั้นตอนที่ 4 การจัดทำคลังข้อสอบ

การจัดทำคลังข้อสอบเป็นการดำเนินการเพื่อจัดระบบสารสนเทศของข้อสอบซึ่งจะใช้สำหรับโปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ มีวิธีดำเนินการ ดังนี้

1. แปลงค่าพารามิเตอร์ของข้อสอบแต่ละฉบับที่ได้จากการวิเคราะห์ให้อยู่บนสเกลเดียวกัน (common scale) โดยเทียบคะแนนตามแบบแผนการใช้ข้อสอบร่วม (anchor item design) ซึ่งหาค่าคงที่สำหรับการแปลงคะแนนด้วยวิธีค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐาน (mean and sigma method) มีขั้นตอนในการแปลงค่าพารามิเตอร์ (ริงสรร์ค มณีเล็ก, 2540) ดังนี้

1.1 ประมาณค่าพารามิเตอร์ของข้อสอบตามทฤษฎีตอบสนองข้อสอบโดยประมาณค่าพารามิเตอร์แยกเป็นรายฉบับ 20 ฉบับ โดยกำหนดค่าความสามารถของผู้สอบให้มีค่าเฉลี่ยเท่ากับ 0 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 1

1.2 หาค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของข้อสอบที่ใช้เป็นข้อสอบร่วมจากทุกฉบับ

1.3 ปรับค่าอำนาจจำแนกและค่าความยากง่ายของข้อสอบจากฉบับหนึ่งไปสู่อีกฉบับหนึ่ง โดยใช้ สูตร  $y = ax + \beta$



$$\begin{aligned} \text{เมื่อ} \quad \alpha &= \frac{SD_{by}}{SD_{bx}} \\ \beta &= \overline{b_y} - \frac{SD_{by}}{SD_{bx}} \overline{b_x} \\ a_y &= \frac{a_x}{\alpha} \\ b_y &= \alpha b_x + \beta \end{aligned}$$

$a_y, a_x$  แทน ค่าอำนาจจำแนกที่ได้รับการแปลงค่าและยังมิได้รับการแปลงค่า

$b_y, b_x$  แทน ค่าความยากง่ายที่ได้รับการแปลงค่าและยังมิได้รับการแปลงค่า

$\overline{b_y}, \overline{b_x}$  แทน ค่าความยากง่ายเฉลี่ยที่ได้รับการแปลงค่าและยังมิได้รับการแปลงค่า

$SD_{by}, SD_{bx}$  แทน ค่าเบี่ยงเบนมาตรฐานที่ได้รับการแปลงค่าและยังมิได้รับการแปลงค่า

1.4 หาค่าเฉลี่ยของความยากง่ายและอำนาจจำแนกของข้อสอบรวมทั้ง 2 ฉบับ หลังจากปรับสเกลแล้ว ทั้งนี้เนื่องจากเป็นข้อสอบข้อเดียวกัน จึงควรมีค่าพารามิเตอร์เดียวกัน

2. คัดเลือกข้อสอบที่มีค่าอำนาจจำแนก อยู่ระหว่าง 0.50 ถึง 2.50 ค่าความยาก อยู่ในช่วง -2.50 ถึง 2.50 ค่าการเดาไม่เกิน 0.30 โดยจะพิจารณาให้ได้ข้อสอบที่มีค่าความยากครอบคลุม ช่วงความสามารถของผู้สอบที่เป็นกลุ่มตัวอย่าง

3. สรุปสารสนเทศรายข้อเพื่อเป็นข้อมูลของข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะ โดยใช้คอมพิวเตอร์

## ระยะที่ 2 การพัฒนาโปรแกรมคอมพิวเตอร์เพื่อใช้ในการทดสอบแบบปรับเหมาะ

การพัฒนาโปรแกรมคอมพิวเตอร์เพื่อใช้ในการทดสอบแบบปรับเหมาะ มีการดำเนินการ 3 ขั้นตอน ได้แก่ การออกแบบโปรแกรมคอมพิวเตอร์ การสร้างและพัฒนาโปรแกรมคอมพิวเตอร์และการทดลองใช้โปรแกรมคอมพิวเตอร์ แต่ละขั้นตอนมีการดำเนินการ ดังนี้

### ขั้นตอนที่ 1 การออกแบบโปรแกรมคอมพิวเตอร์

การออกแบบโปรแกรมคอมพิวเตอร์เป็นการกำหนดรายละเอียดที่ใช้ในการสร้างและพัฒนาโปรแกรมคอมพิวเตอร์ ได้แก่ ลักษณะเฉพาะของโปรแกรมและข้อกำหนดในการเขียนโปรแกรม มีรายละเอียด ดังนี้

1. กำหนดลักษณะเฉพาะของโปรแกรม ให้เป็นโปรแกรมคอมพิวเตอร์ที่สามารถทำงานบนระบบปฏิบัติการวินโดวส์ได้ เนื่องจากเป็นระบบปฏิบัติการที่ใช้กันอย่างแพร่หลายสำหรับเครื่องคอมพิวเตอร์ปัจจุบัน โดยกำหนดใช้ภาษา Borland Delphi 7.0 ในการเขียนโปรแกรม

เนื่องจากเป็นภาษาที่สามารถเขียนภาพกราฟฟิกและสูตรสมการฟังก์ชันทางคณิตศาสตร์ได้ รวมทั้งเป็นภาษาที่มีโครงสร้างชัดเจน ง่ายต่อการใช้งาน ตลอดจนสามารถใช้กับเครื่องไมโครคอมพิวเตอร์ได้

2. กำหนดข้อกำหนดในการเขียนโปรแกรมสำหรับโปรแกรมการทดสอบ โดยจำแนกเป็นโปรแกรมย่อยตามตัวแปรที่ศึกษา จำนวน 6 โปรแกรม ดังแสดงในตารางที่ 9

ตารางที่ 9 รายละเอียดโปรแกรมย่อยที่ใช้ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

โปรแกรมย่อยที่	จำนวนข้อสอบที่กำหนดให้ทำในแต่ละชั้นอย่างน้อยจำนวน	เกณฑ์การเลื่อนลำดับชั้นของคุณลักษณะเพื่อเลือกข้อสอบข้อต่อไป	เกณฑ์การหยุดสอบ
1	3	ตอบข้อสอบถูกจำนวน 2 ใน 3 ข้อ หรือ ร้อยละ 67	ใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 2 ใน 3 ข้อ หรือ ร้อยละ 67
2	4	ตอบข้อสอบถูกจำนวน 3 ใน 4 ข้อ หรือ ร้อยละ 75	ใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 3 ใน 4 ข้อ หรือ ร้อยละ 75
3	5	ตอบข้อสอบถูกจำนวน 4 ใน 5 ข้อ หรือ ร้อยละ 80	ใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 4 ใน 5 ข้อ หรือ ร้อยละ 80
4	3	ตอบข้อสอบถูกจำนวน 2 ใน 3 ข้อ หรือ ร้อยละ 67 และมีค่าความคลาดเคลื่อนมาตรฐานในการ	ใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 2 ใน 3 ข้อ หรือ ร้อยละ 67 และมีค่าความคลาดเคลื่อน

		ประมาณค่าต่ำกว่า 0.3	มาตรฐานในการประมาณค่าต่ำกว่า 0.3
5	4	ตอบข้อสอบถูกจำนวน 3 ใน 4 ข้อ หรือ ร้อยละ 75 และมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3	ใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 3 ใน 4 ข้อ หรือ ร้อยละ 75 และมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3
6	5	ตอบข้อสอบถูกจำนวน 4 ใน 5 ข้อ หรือ ร้อยละ 80 และมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3	ใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 4 ใน 5 ข้อ หรือ ร้อยละ 80 และมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3

แต่ละโปรแกรมมีข้อกำหนดในการเขียนโปรแกรมร่วมกัน ดังนี้

2.1 การคัดเลือกข้อสอบข้อแรก คัดเลือกโดยกำหนดให้ข้อสอบข้อแรกเป็นข้อสอบที่อยู่ในลำดับขั้นของคุณลักษณะต่ำสุดซึ่งเป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะในขั้นที่สูงกว่าทุกคุณลักษณะ โดยมีค่าความยากของข้อสอบปานกลาง

2.2 การคัดเลือกข้อสอบจากคลังข้อสอบ ใช้วิธีคัดเลือกข้อสอบที่มีค่าฟังก์ชันสารสนเทศสูงสุด (maximum item information : I<sub>max</sub>) ในลำดับขั้นของคุณลักษณะนั้น ๆ หรือลำดับขั้นของคุณลักษณะที่สูงกว่าตามแบบแผนการตอบข้อสอบที่คาดหวัง

2.3 การประมาณค่าความสามารถของผู้สอบ ใช้วิธีของเบย์ส์ที่ปรับเปลี่ยน (Bayesian Updating) มีสูตรในการคำนวณ ดังนี้ (Owen, 1975 อ้างถึงใน รั้งสรรค์ มณีเล็ก, 2540)

ถ้าตอบถูก

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

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

ถ้าตอบผิด

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

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

เมื่อ  $\Theta(D)$  แทน ค่าออร์ดิเนต (ordinate) ของโค้งปกติที่จุด D  
 $\Phi(D)$  แทน พื้นที่ใต้โค้งปกติจากค่า D ต่ำสุด จนถึงจุด D ที่มีค่าเท่ากับ  $\frac{b - \theta_m}{\sqrt{\frac{1}{a^2} + \sigma_m^2}}$  และ  $A = c + (1 - c) \Phi(-D)$

$\theta_m$  แทน ความสามารถของผู้สอบที่ประมาณค่าได้ก่อนตอบข้อสอบข้อที่ m+1 ซึ่งตามปกติแล้ว ถ้าไม่ทราบค่าความสามารถเบื้องต้นของผู้สอบ ก็มักกำหนดให้เท่ากับ 0

$\sigma_m^2$  แทน ความแปรปรวนในการประมาณค่าความสามารถของผู้สอบก่อนตอบข้อที่ m+1 ถ้าไม่ทราบค่าความแปรปรวนดังกล่าวมาก่อน ก็มักกำหนดให้เท่ากับ 1

$\theta_{m+1}$  แทน ค่าความสามารถของผู้สอบ โดยประมาณหลังจากที่ตอบข้อที่ m+1

$\sigma_{m+1}^2$  แทน ค่าความแปรปรวนในการประมาณค่าความสามารถของผู้สอบเมื่อตอบข้อสอบข้อที่ m+1

a แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ m+1

b แทน ค่าความยากของข้อสอบข้อที่ m+1

c แทน ค่าการเดาของข้อสอบข้อที่ m+1

2.4 การรายงานผลการทดสอบ ประกอบด้วยข้อมูลสำคัญ 2 ส่วน ส่วนที่ 1 เป็นข้อมูลทั่วไปเกี่ยวกับการทดสอบและผู้ทดสอบ ได้แก่ ชื่อแบบทดสอบ ชื่อ-นามสกุลผู้สอบ ชั้นเรียน

ห้องเรียน โรงเรียน วันเดือนปีที่ทำการทดสอบ และเวลาที่เริ่มทดสอบ ส่วนที่ 2 เป็นข้อมูลเกี่ยวกับ ผลการทดสอบประกอบด้วย เงื่อนไขการทดสอบ เวลาที่ใช้ในการทดสอบ ตำแหน่งเปอร์เซ็นต์ไทล์ ความสามารถในการสอบ ความคลาดเคลื่อนมาตรฐานในการประมาณค่า และรายละเอียดการทำข้อสอบ ได้แก่ ข้อที่ ขุด/ข้อในคลังข้อสอบ คำตอบ ผลการตรวจ แบบการคิดที่ผิดที่ใช้ (ตัวอย่าง รายงานผลการทดสอบแสดงในภาคผนวก ค)

## ขั้นตอนที่ 2 การสร้างและพัฒนาโปรแกรมคอมพิวเตอร์

การสร้างและพัฒนาโปรแกรมคอมพิวเตอร์ ดำเนินการโดยเขียนผังโปรแกรมการทำงานของคอมพิวเตอร์และเขียนโปรแกรมคอมพิวเตอร์ตามผังโปรแกรมที่กำหนดไว้ ทดลองการทำงานของโปรแกรมคอมพิวเตอร์เบื้องต้น และปรับปรุงโปรแกรมคอมพิวเตอร์เบื้องต้น

## ขั้นตอนที่ 3 การทดลองใช้โปรแกรมคอมพิวเตอร์

การทดลองใช้โปรแกรมคอมพิวเตอร์มีวัตถุประสงค์เพื่อศึกษาประสิทธิภาพขั้นต้นของโปรแกรมคอมพิวเตอร์ และนำผลที่ได้มาปรับปรุงโปรแกรมคอมพิวเตอร์ให้สามารถใช้งานได้ อย่างมีประสิทธิภาพมากขึ้น มีการดำเนินการ ดังนี้

1. ทดลองใช้โปรแกรมคอมพิวเตอร์ ดำเนินการโดยผู้วิจัย เพื่อให้กลุ่มตัวอย่าง กลุ่มที่ 2.6 ทดลองใช้โปรแกรมคอมพิวเตอร์ที่ละกลุ่ม จำแนกตามโปรแกรมย่อย โดยก่อนดำเนินการทดลองใช้โปรแกรมคอมพิวเตอร์จะอธิบายและฝึกทักษะการใช้คอมพิวเตอร์เบื้องต้น ให้กับผู้สอบและทดสอบทักษะดังกล่าวจนมั่นใจว่าผู้สอบทุกคนสามารถใช้คอมพิวเตอร์ได้ ทั้งนี้ เพื่อควบคุมไม่ให้เกิดปัญหาในการใช้คอมพิวเตอร์ของผู้สอบมีผลต่อการทดสอบ วิเคราะห์ข้อมูลเชิงปริมาณ โดยการคำนวณค่าความถี่และร้อยละ ค่าเฉลี่ยเลขคณิตและส่วนเบี่ยงเบนมาตรฐาน และวิเคราะห์ข้อมูลเชิงคุณภาพโดยการวิเคราะห์เนื้อหาและสรุปเป็นความเรียง

2. สอบถามปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ ดำเนินการโดยให้กลุ่มตัวอย่างกลุ่มที่ 2.7 ตอบแบบสอบถามปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์หลังจากได้ทดลองใช้โปรแกรมเรียบร้อยแล้ว วิเคราะห์ข้อมูลเชิงปริมาณโดยการคำนวณค่าความถี่ ร้อยละ ค่าเฉลี่ยเลขคณิตและส่วนเบี่ยงเบนมาตรฐาน และวิเคราะห์ข้อมูลเชิงคุณภาพโดยการวิเคราะห์เนื้อหาและสรุปเป็นความเรียง

ระยะที่ 3 การประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน

การประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน มีการดำเนินการ 2 ขั้นตอน ได้แก่ การประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์และการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนโดยใช้เทคนิคการคิดออกเสียง แต่ละขั้นตอนมีการดำเนินการ ดังนี้

### **ขั้นตอนที่ 1 การประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์**

การประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ มีการดำเนินการ ดังนี้

1. ดำเนินการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์โดยเริ่มจากการติดตั้งโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะที่เครื่องคอมพิวเตอร์ของโรงเรียน จัดเตรียมอุปกรณ์ที่ใช้ในการทดสอบสำหรับแต่ละเครื่อง ตรวจสอบความพร้อมของเครื่องคอมพิวเตอร์และอุปกรณ์จัดนักเรียนซึ่งเป็นกลุ่มตัวอย่าง กลุ่มที่ 2.8 เข้ารับการทดสอบโปรแกรมย่อยละ 30 คน ในแต่ละขนาดโรงเรียน (6 โปรแกรมย่อย  $\times$  30 คน  $\times$  3 ขนาดโรงเรียน) ก่อนดำเนินการทดสอบจะอธิบายและฝึกทักษะการใช้คอมพิวเตอร์เบื้องต้นให้กับนักเรียนและทดสอบทักษะดังกล่าวจนมั่นใจว่านักเรียนทุกคนสามารถใช้คอมพิวเตอร์ในการดำเนินการทดสอบได้ ทั้งนี้เพื่อควบคุมไม่ให้เกิดการผิดพลาดในการใช้คอมพิวเตอร์ของนักเรียนมีผลต่อการทดสอบ นักเรียนแต่ละคนได้รับการทดสอบ จำนวน 2 ครั้ง ระยะเวลาห่างกันครั้งละ 1 สัปดาห์ วิเคราะห์ข้อมูลเชิงปริมาณโดยการคำนวณค่าความถี่และร้อยละ และวิเคราะห์ข้อมูลเชิงคุณภาพโดยใช้การวิเคราะห์เนื้อหาและสรุปเป็นความเรียง

2. สอบถามปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ ดำเนินการโดยให้กลุ่มตัวอย่าง กลุ่มที่ 2.9 ตอบแบบสอบถามปัญหาที่เกิดจากการใช้โปรแกรมหลังจากเสร็จสิ้นการประเมินครั้งที่ 2 วิเคราะห์ข้อมูลเชิงปริมาณโดยการคำนวณค่าความถี่ ร้อยละ ค่าเฉลี่ยเลขคณิตและส่วนเบี่ยงเบนมาตรฐาน และวิเคราะห์ข้อมูลเชิงคุณภาพโดยใช้การวิเคราะห์เนื้อหาและสรุปเป็นความเรียง

### **ขั้นตอนที่ 2 การประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน โดยใช้เทคนิคการคิดออกเสียง**

เทคนิคการคิดออกเสียง (think aloud technique) เป็นเทคนิคที่เหมาะสมในการศึกษากระบวนการคิดแก้ปัญหา โดยใช้วิธีให้ผู้สอบรายงานความคิด และการกระทำด้วยถ้อยคำ (verbal report) ซึ่งดำเนินการได้ทั้งในขณะที่ผู้สอบแก้ปัญหา (concurrent protocol) และหลังจากที่ผู้สอบแก้ปัญหาแล้ว (retrospective protocol) สำหรับการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน โดยใช้เทคนิคการคิดออกเสียงในครั้งนี้ เลือกดำเนินการในขณะที่ผู้สอบแก้ปัญหา โดยให้กลุ่มตัวอย่างกลุ่มที่ 2.10 ทำแบบทดสอบแบบเขียนตอบบนกระดาษคำตอบ พร้อมทั้งพูดถึงกระบวนการคิดที่ใช้ในการแก้ปัญหาแต่ละขั้นตอนสำหรับการตอบข้อสอบแต่ละข้อ ในขณะที่เดียวกันผู้ประเมินจะวิเคราะห์ถึงแบบการคิดที่ผู้สอบใช้ในการแก้ปัญหา กรณีที่ผู้สอบพูดถึงกระบวนการคิดที่ใช้ในการแก้ปัญหาไม่ชัดเจนหรือผู้ประเมินไม่มั่นใจในการจำแนกแบบการคิดที่ผู้สอบใช้ ผู้ประเมินจะดำเนินการสอบถามกลับไปกลับมา (check-recheck) หรือให้ผู้สอบทำข้อสอบข้ออื่นที่วัดคุณลักษณะในลำดับขั้นเดียวกัน จนกว่าจะมั่นใจได้ว่าผู้สอบใช้แบบการคิดนั้นจริง การประเมินกำหนดให้นักเรียน 1 คน ได้รับการประเมินจำนวน 2 ครั้ง โดยครูผู้สอนและผู้วิจัย การดำเนินการใช้คู่มือการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน โดยใช้เทคนิคการคิดออกเสียงเป็นแนวทางการดำเนินการ ศึกษาความตรงของการประเมินจากความสอดคล้องระหว่างผลการประเมินของครูผู้สอนกับผู้วิจัย (inter rater validity) วิเคราะห์ข้อมูลเชิงปริมาณโดยการคำนวณค่าความถี่และร้อยละและวิเคราะห์ข้อมูลเชิงคุณภาพโดยใช้การวิเคราะห์เนื้อหาและสรุปเป็นความเรียง

#### **ระยะที่ 4 การศึกษาคุณภาพและความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น**

การศึกษาคุณภาพและความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้นมีการดำเนินการ 3 ขั้นตอน ได้แก่ การศึกษาคุณภาพของวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น การเปรียบเทียบคุณภาพของวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้นและการศึกษาความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น แต่ละขั้นตอนมีการดำเนินการ ดังนี้

##### **ขั้นตอนที่ 1 การศึกษาคุณภาพของวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น**

คุณภาพของวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้นศึกษาจากค่าความเที่ยง ค่าความตรง จำนวนข้อสอบที่ใช้และจำนวนเวลาที่ใช้ในการทดสอบ มีรายละเอียด ดังนี้

1. ค่าความเที่ยงพิจารณาจากค่าสัดส่วนของความสอดคล้องระหว่างผลการประเมินเชิงวินิจฉัยจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง ดำเนินการวิเคราะห์ค่าสัดส่วนโดยใช้โปรแกรมสำเร็จรูป SPSS/PC
2. ค่าความตรงพิจารณาจากค่าสัดส่วนของความสอดคล้องระหว่างผลการประเมินเชิงวินิจฉัยจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (computerized adaptive testing) กับผลการประเมินเชิงวินิจฉัยจากการทดสอบโดยใช้เทคนิคการคิดออกเสียง (think aloud technique) ดำเนินการวิเคราะห์ค่าสัดส่วนโดยใช้โปรแกรมสำเร็จรูป SPSS/PC
3. จำนวนข้อสอบที่ใช้พิจารณาในภาพรวมและจำแนกตามโปรแกรมย่อยซึ่งแตกต่างกันตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้ ดำเนินการวิเคราะห์ค่าความถี่และค่าร้อยละ ค่าเฉลี่ยเลขคณิตและส่วนเบี่ยงเบนมาตรฐานของจำนวนข้อสอบที่ใช้ โดยใช้โปรแกรมสำเร็จรูป SPSS/PC
4. จำนวนเวลาที่ใช้ในการทดสอบพิจารณาในภาพรวมและจำแนกตามโปรแกรมย่อยซึ่งแตกต่างกันตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้ ดำเนินการวิเคราะห์ค่าความถี่และค่าร้อยละ ค่าเฉลี่ยเลขคณิตและส่วนเบี่ยงเบนมาตรฐานของจำนวนข้อสอบที่ใช้ โดยใช้โปรแกรมสำเร็จรูป SPSS/PC

## ขั้นตอนที่ 2 การเปรียบเทียบคุณภาพของการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น

การเปรียบเทียบคุณภาพของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น เมื่อข้อกำหนดในการตัดสินความรอบรู้ในแต่ละลำดับขั้นแตกต่างกัน มีการดำเนินการ ดังนี้

- 1) การเปรียบเทียบค่าความเที่ยงและค่าความตรงใช้การทดสอบไคสแควร์ ( $\chi^2$  - test) โดยใช้โปรแกรมสำเร็จรูป SPSS/PC
- 2) การเปรียบเทียบจำนวนข้อสอบที่ใช้และจำนวนเวลาที่ใช้การวิเคราะห์ความแปรปรวนทางเดียว (one way ANOVA) โดยใช้โปรแกรมสำเร็จรูป SPSS/PC

## ขั้นตอนที่ 3 การศึกษาความคิดเห็นที่มีต่อวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น

การศึกษาความคิดเห็นที่มีต่อวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น มีวัตถุประสงค์เพื่อให้กลุ่มตัวอย่างแสดงความคิดเห็นที่มีต่อการใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ โดยการประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะในด้านความยากง่ายในการใช้ ความเป็นไปได้ในการนำไปใช้ ประโยชน์ที่ได้รับจากการวินิจฉัยและความพึงพอใจที่มีต่อการใช้ ดำเนินการโดยให้กลุ่มตัวอย่าง กลุ่มที่ 2.9 ตอบแบบสอบถามความคิดเห็นที่มีต่อวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น หลังจากการประเมินเชิงวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบ



เศษส่วน ด้วยการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 2 เสร็จสิ้น วิเคราะห์ข้อมูล โดยตรวจสอบความถูกต้อง สมบูรณ์ของข้อมูล และวิเคราะห์ข้อมูลสภาพทั่วไปของผู้ตอบแบบสอบถาม ตอนที่ 1 ด้วยสถิติบรรยาย ได้แก่ ความถี่และร้อยละ โดยใช้โปรแกรมสำเร็จรูป SPSS/PC ส่วนข้อมูลความคิดเห็นที่มีต่อการใช้โปรแกรม ตอนที่ 2 วิเคราะห์โดยการวิเคราะห์เนื้อหา และสรุปเป็นความเรียง

## บทที่ 4

### ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีการดำเนินการวิจัย 4 ระยะ ระยะที่ 1 การพัฒนาชุดของข้อสอบ แบบทดสอบแบบเขียนตอบบนกระดาษคำตอบและคลังข้อสอบ ระยะที่ 2 การพัฒนาโปรแกรมคอมพิวเตอร์เพื่อใช้ในการทดสอบแบบปรับเหมาะ ระยะที่ 3 การประเมินเชิงวิจนิจฉัยสถานะความรู้และแบบการคิดที่ผิด

ที่ใช้ในการบวกลบเศษส่วน และระยะที่ 4 การศึกษาคุณภาพและความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น จากการดำเนินการดังกล่าวจำแนกผลการวิเคราะห์ข้อมูลได้ 4 ตอน ดังนี้

ตอนที่ 1 ผลการพัฒนาวิธีการประเมินเชิงวิจักษ์โดยประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

ตอนที่ 2 ผลการศึกษาคุณภาพของวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น

ตอนที่ 3 ผลการเปรียบเทียบคุณภาพของวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น

ตอนที่ 4 ผลการศึกษาความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น

ปรากฏผลการวิเคราะห์ข้อมูลในแต่ละตอน ดังนี้

### **ตอนที่ 1 ผลการพัฒนาวิธีการประเมินเชิงวิจักษ์โดยประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์**

องค์ประกอบสำคัญของการพัฒนาวิธีการประเมินเชิงวิจักษ์โดยประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ คือ โปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะซึ่งเป็นเครื่องมือการประเมินที่พัฒนาขึ้นจากแนวคิดโมเดลลำดับชั้นของคุณลักษณะ โปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะสำหรับการวิจัยในครั้งนี้จึงแตกต่างจากโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะทั่วไป โดยเฉพาะลักษณะของสารสนเทศในคลังข้อสอบและลำดับชั้นการเลือกข้อสอบจากคลังข้อสอบ ผลการพัฒนาวิธีการประเมินเชิงวิจักษ์โดยประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะในตอนที่ 1 นี้ จึงนำเสนอผลการดำเนินการตั้งแต่การพัฒนากรอบแนวคิดที่ใช้ในการพัฒนาวิธีการประเมิน ได้แก่ ผลการกำหนดลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน และการพัฒนาเครื่องมือการประเมิน ได้แก่ ผลการพัฒนาคลังข้อสอบและโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะ มีรายละเอียดผลการดำเนินการ ดังนี้

#### **1. ผลการกำหนดลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน**

การกำหนดลำดับชั้นของคุณลักษณะมีการดำเนินการ 3 ขั้นตอน ได้แก่ การสำรวจคุณลักษณะ การจัดเรียงลำดับชั้นของคุณลักษณะและการจัดทำแผนภาพลำดับชั้นของคุณลักษณะ มีผลการวิเคราะห์ข้อมูลจากการดำเนินการดังกล่าว ดังนี้

#### **ผลการสำรวจคุณลักษณะที่ใช้ในการบวกลบเศษส่วน**

การสำรวจคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) ดำเนินการเพื่อให้ได้คุณลักษณะหลักที่ใช้ในการบวกลบเศษส่วนด้วยการวิเคราะห์เนื้อหาจากการศึกษาเอกสารและงานวิจัยเกี่ยวกับหลักสูตรและการจัดการเรียนการสอน วิชาคณิตศาสตร์ เรื่องการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) และเพื่อให้ได้คุณลักษณะย่อยที่ใช้ในการบวกลบเศษส่วน โดยใช้แบบสำรวจ สรุปผลการสำรวจคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ได้ดังนี้

#### คุณลักษณะหลักที่ใช้ในการบวกลบเศษส่วน

จากการศึกษาเอกสารและงานวิจัยเกี่ยวกับหลักสูตรและการจัดการเรียนการสอนวิชาคณิตศาสตร์ เรื่องการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) เพื่อสำรวจเนื้อหาที่ใช้ในการบวกลบเศษส่วนพบว่า มีเนื้อหาที่สำคัญจำแนกได้ 2 ลักษณะ ลักษณะที่ 1 เป็นเนื้อหาเกี่ยวกับการบวกลบเศษส่วนชนิดเดียวกัน จำแนกเป็นการบวกเศษส่วนแท้กับเศษส่วนแท้ ทั้งที่มีตัวส่วนเท่ากันและตัวส่วนไม่เท่ากัน ระหว่างเศษส่วน 2 จำนวนและ 3 จำนวน และการบวกจำนวนคละกับจำนวนคละ ทั้งที่มีตัวส่วนเท่ากันและตัวส่วนไม่เท่ากัน ระหว่างเศษส่วน 2 จำนวนและ 3 จำนวน ลักษณะที่ 2 เป็นเนื้อหาเกี่ยวกับการบวกลบเศษส่วนต่างชนิดกัน จำแนกเป็นการบวกลบจำนวนนับกับเศษส่วนแท้ และการบวกลบเศษส่วนแท้กับจำนวนคละ ทั้งที่มีตัวส่วนเท่ากันและตัวส่วนไม่เท่ากัน

เมื่อวิเคราะห์วิธีดำเนินการเพื่อให้ได้ผลลัพธ์จากเนื้อหาดังกล่าวพบว่ามีบางเนื้อหาเป็นเนื้อหาของย่อยของบางเนื้อหา (Subset) ดังนั้นเพื่อให้การจำแนกคุณลักษณะที่ใช้ในการบวกลบเศษส่วนมีความถูกต้องและชัดเจนขึ้น ผู้วิจัยจึงจัดทำแบบทดสอบแบบแสดงวิธีทำจำนวน 1 ฉบับ โดยกำหนดให้มีข้อคำถามครอบคลุมเนื้อหาทั้งชนิดของเศษส่วน (เศษส่วนแท้ จำนวนคละและจำนวนเต็ม) ลักษณะของตัวส่วน (ตัวส่วนเท่ากันและตัวส่วนไม่เท่ากัน) และจำนวนของเศษส่วน (2 จำนวนและ 3 จำนวน) ได้ข้อสอบ จำนวน 22 ข้อ นำแบบทดสอบที่ได้ไปให้นักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 30 คน หาคำตอบโดยแสดงวิธีทำ นำผลการทำแบบทดสอบมาตรวจให้คะแนนและวิเคราะห์คุณลักษณะที่ใช้ในการบวกลบเศษส่วน โดยนำผลการวิเคราะห์เนื้อหาจากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องมาพิจารณาประกอบ สรุปคุณลักษณะหลักที่ใช้ในการบวกลบเศษส่วนได้ จำนวน 4 คุณลักษณะ ดังนี้

คุณลักษณะที่ 1 การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน

คุณลักษณะที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน

คุณลักษณะที่ 3 การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน

คุณลักษณะที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน

### 1.1.2 คุณลักษณะย่อยที่ใช้ในการบวกลบเศษส่วนสำหรับแต่ละคุณลักษณะหลัก

จากการวิเคราะห์ผลการสำรวจคุณลักษณะย่อยที่ใช้ในการบวกลบเศษส่วนจากแบบสำรวจคุณลักษณะที่ใช้ในการบวกลบเศษส่วนที่ได้รับกลับคืน จำนวน 26 ฉบับ สรุปคุณลักษณะย่อยที่ใช้ในการบวกลบเศษส่วนซึ่งมีผลการประเมินความถูกต้องของคุณลักษณะระหว่างผู้ประเมินสอดคล้องกัน จำนวน 3 ใน 3 ได้ดังแสดงในตารางที่ 10 (สภาพทั่วไปของกลุ่มตัวอย่างในการสำรวจคุณลักษณะย่อยที่ใช้ในการบวกลบเศษส่วนและกลุ่มตัวอย่างในการประเมินความถูกต้องของคุณลักษณะที่ใช้ในการบวกลบเศษส่วนแสดงในภาคผนวก ฉ) ตารางที่ 10 คุณลักษณะย่อยที่ใช้ในการบวกลบเศษส่วนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) จำแนกตามคุณลักษณะหลัก

คุณลักษณะหลัก	คุณลักษณะย่อย
1. การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน	เข้าใจความหมายของเศษส่วนแท้ที่มีตัวส่วนเท่ากัน เข้าใจวิธีบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน เข้าใจวิธีทำผลลัพธ์ให้เป็นเศษส่วนอย่างต่ำ บวกลบจำนวนนับ คูณจำนวนนับ หารจำนวนนับ
2. การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน	2.1 เข้าใจความหมายของเศษส่วนแท้ที่มีตัวส่วนเท่ากัน 2.2 เข้าใจความหมายของเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน 2.3 เข้าใจวิธีบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน 2.4 เข้าใจวิธีบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน 2.5 เข้าใจวิธีทำเศษส่วนแท้ 2 จำนวนให้มีตัวส่วนเท่ากัน 2.6 เข้าใจวิธีทำผลลัพธ์ให้เป็นเศษส่วนอย่างต่ำ 2.7 บวกลบจำนวนนับ 2.8 คูณจำนวนนับ 2.9 หารจำนวนนับ
ตารางที่ 10 คุณลักษณะย่อยที่ใช้ในการบวกลบเศษส่วนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) จำแนกตามคุณลักษณะหลัก (ต่อ)	
คุณลักษณะหลัก	คุณลักษณะย่อย
3. การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน	3.1 เข้าใจความหมายของเศษส่วนแท้ที่มีตัวส่วนเท่ากัน 3.2 เข้าใจความหมายของจำนวนคละ 3.3 เข้าใจความหมายของเศษเกิน

	3.4 เข้าใจวิธีบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน
	<b>วิธีที่ 1</b>
	1) บวกลบจำนวนนับ
	2) เข้าใจวิธีบวกลบเศษส่วนที่มีตัวส่วนเท่ากัน
	<b>วิธีที่ 2</b>
	1) เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน
	2) เข้าใจวิธีบวกลบเศษเกินที่มีตัวส่วนเท่ากัน
	3.5 เข้าใจวิธีทำผลลัพธ์ให้เป็นเศษส่วนอย่างต่ำ
	3.6 คูณจำนวนนับ
	3.7 หารจำนวนนับ
4. การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน	4.1 เข้าใจความหมายของเศษส่วนแท้ที่มีตัวส่วนเท่ากัน
	4.2 เข้าใจความหมายของเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน
	4.3 เข้าใจความหมายของจำนวนคละ
	4.4 เข้าใจความหมายของเศษเกิน
	4.5 เข้าใจวิธีบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน
	<b>วิธีที่ 1</b>
	1) บวกลบจำนวนนับ
	2) เข้าใจวิธีบวกลบเศษส่วนที่มีตัวส่วนเท่ากัน
	3) เข้าใจวิธีทำเศษส่วนแท้ 2 จำนวน ให้มีตัวส่วนเท่ากัน
	<b>วิธีที่ 2</b>
	1) เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน
	2) เข้าใจวิธีบวกลบเศษเกินที่มีตัวส่วนไม่เท่ากัน
	3) เข้าใจวิธีทำเศษเกิน 2 จำนวน ให้มีตัวส่วนเท่ากัน
	4) เข้าใจวิธีบวกลบเศษเกินที่มีตัวส่วนเท่ากัน
	5) บวกลบจำนวนนับ
	4.6 เข้าใจวิธีทำผลลัพธ์ให้เป็นเศษส่วนอย่างต่ำ
	4.7 คูณจำนวนนับ
	4.8 หารจำนวนนับ

## 1.2 ผลการจัดเรียงลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน

การจัดเรียงลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วนนี้ใช้ผลจากการเปรียบเทียบลำดับชั้นของคุณลักษณะโดยใช้กระบวนการให้น้ำหนักรายคู่ โดยการเปรียบเทียบความเป็นคุณลักษณะที่ต้องมีมาก่อน พบว่า คุณลักษณะที่ 1 การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากันเป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน คุณลักษณะที่ 3 การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน และคุณลักษณะที่ 4 การบวก

ลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน โดยมีน้ำหนักความสำคัญ ร้อยละ 80.77, 88.46 และ 88.46 ตามลำดับ เมื่อพิจารณาการเปรียบเทียบความเป็นคุณลักษณะที่ต้องมีมาก่อนของคุณลักษณะที่ 2 และคุณลักษณะที่ 3 กับคุณลักษณะที่ 4 พบว่า คุณลักษณะที่ 2 และคุณลักษณะที่ 3 เป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะที่ 4 โดยมีค่าน้ำหนักความสำคัญ ร้อยละ 92.31 และ 96.15 ตามลำดับ เมื่อพิจารณาความเป็นคุณลักษณะที่ต้องมีมาก่อนของคุณลักษณะที่ 2 กับคุณลักษณะที่ 3 พบว่า เป็นคุณลักษณะที่อยู่ในลำดับชั้นเดียวกัน เนื่องจากมีค่าน้ำหนักความสำคัญในการเปรียบเทียบในทุกกรณีใกล้เคียงกัน กล่าวคือ ในการเปรียบเทียบ กรณีคุณลักษณะที่ 2 ต้องมีมาก่อนคุณลักษณะที่ 3 มีค่าน้ำหนักความสำคัญ ร้อยละ 30.77 กรณีคุณลักษณะที่ 3 ต้องมีมาก่อนคุณลักษณะที่ 2 มีค่าน้ำหนักความสำคัญ ร้อยละ 26.92 กรณีเป็นคุณลักษณะที่ไม่มีความเกี่ยวข้องกัน มีค่าน้ำหนักความสำคัญ ร้อยละ 19.23 และเป็นคุณลักษณะที่อยู่ลำดับชั้นเดียวกัน มีค่าน้ำหนักความสำคัญ ร้อยละ 23.08 ซึ่งมีค่าสูงกว่าเมื่อเปรียบเทียบกับคุณลักษณะอื่น ๆ

จากผลการเปรียบเทียบลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน โดยใช้กระบวนการให้น้ำหนักรายคู่ดังกล่าว สรุปลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) ได้ดังนี้

ลำดับชั้นที่ 1 ได้แก่ คุณลักษณะที่ 1 การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน

ลำดับชั้นที่ 2 ได้แก่ คุณลักษณะที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน และคุณลักษณะที่ 3 การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน

ลำดับชั้นที่ 3 ได้แก่ คุณลักษณะที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน

ผลการเปรียบเทียบลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วนโดยใช้กระบวนการให้น้ำหนักรายคู่ แสดงในตารางที่ 11

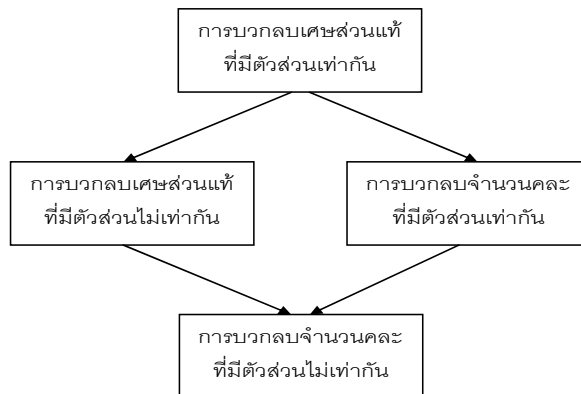
ตารางที่ 11 ผลการเปรียบเทียบลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วนโดยใช้กระบวนการให้น้ำหนักรายคู่

ผลการเปรียบเทียบ	คุณลักษณะที่ 1 กับ 2		คุณลักษณะที่ 1 กับ 3		คุณลักษณะที่ 1 กับ 4		คุณลักษณะที่ 2 กับ 3		คุณลักษณะที่ 2 กับ 4		คุณลักษณะที่ 3 กับ 4	
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ
คุณลักษณะแรก ต้องมาก่อน	21	80.77	23	88.46	23	88.46	8	30.77	24	92.31	25	96.15

คุณลักษณะหลัง												
คุณลักษณะหลัง ต้องมีมาก่อน	0	0.00	0	0.00	0	0.00	7	26.92	0	0.00	0	0.00
คุณลักษณะแรก												
คุณลักษณะ แรกและหลัง ไม่เกี่ยวข้องกัน	4	15.38	2	7.69	2	7.69	5	19.23	0	0.00	0	0.00
คุณลักษณะ แรกและหลัง อยู่ลำดับชั้น เดียวกัน	1	3.85	1	3.85	1	3.85	6	23.08	2	7.69	1	3.85
<b>รวม</b>	<b>26</b>	<b>100</b>	<b>26</b>	<b>100</b>	<b>26</b>	<b>100</b>	<b>26</b>	<b>100</b>	<b>26</b>	<b>100</b>	<b>26</b>	<b>100</b>

### 1.3 ผลการจัดทำแผนภาพลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน

จากผลการจัดเรียงลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) นำมาจัดทำแผนภาพลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ซึ่งนำไปประเมินความถูกต้องและมีผลการประเมินระหว่างกลุ่มตัวอย่างสอดคล้องกัน จำนวน 3 ใน 3 ดังแสดงในแผนภาพที่ 4



แผนภาพที่ 4 ลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6)

### 2. ผลการพัฒนาคัดข้อสอบและโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะ

การพัฒนาคัดข้อสอบและโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะ มีการดำเนินการ 2 ขั้นตอน ได้แก่ การพัฒนาคัดข้อสอบ และการพัฒนาโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะ มีผลการวิเคราะห์ข้อมูลจากการดำเนินการดังกล่าว ดังนี้

#### ผลการพัฒนาคัดข้อสอบ

การพัฒนาคลังข้อสอบดำเนินการโดยใช้แนวคิดโมเดลลำดับชั้นของคุณลักษณะ เพื่อให้ได้ข้อสอบที่มีคุณภาพสำหรับใช้ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ผลการพัฒนาจะได้ทั้งข้อสอบ ซึ่งประกอบด้วย สารสนเทศของข้อสอบ ข้อคำถาม คำตอบถูก และคำตอบผิด สำหรับแต่ละแบบการคิดที่ผิด รวมทั้งกระบวนการในการทดสอบ ได้แก่ ลำดับชั้นการเลือกข้อสอบ จากชุดของข้อสอบในคลังข้อสอบ ตลอดจนแนวทางการวินิจฉัยผลการประเมิน โดยแบ่งการนำเสนอผลการวิเคราะห์ข้อมูลจากการดำเนินการเป็น 5 ส่วน แต่ละส่วนมีรายละเอียด ดังนี้

#### ผลการกำหนดเมทริกซ์เพื่อใช้ในการพัฒนาชุดของข้อสอบ

เมทริกซ์ต่าง ๆ ที่ใช้ในการพัฒนาชุดของข้อสอบกำหนดจากความสัมพันธ์เชิงลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ซึ่งแสดงในแผนภาพลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ปรากฏผลการดำเนินการ ดังนี้

#### ก. ผลการกำหนด binary adjacency matrix (A)

การกำหนด binary adjacency matrix (A) มีวัตถุประสงค์เพื่อแสดงความสัมพันธ์ทางตรงระหว่างคุณลักษณะที่ใช้ในการบวกลบเศษส่วน จำนวน 4 คุณลักษณะ ได้แก่ การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน และการบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน มีขนาดเมทริกซ์เป็น  $4 \times 4$  ดังนี้

$$\begin{bmatrix} 0110 \\ 0001 \\ 0001 \\ 0000 \end{bmatrix}$$

จากเมทริกซ์ (A) ข้างต้นอธิบายความสัมพันธ์ทางตรงระหว่างคุณลักษณะได้ดังนี้

1) คุณลักษณะที่ 1 การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากันเป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน และคุณลักษณะที่ 3 การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน จะเห็นได้จากตำแหน่งที่ (1, 2) และตำแหน่งที่ (1, 3) แทนด้วยเลข 1



2) คุณลักษณะที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากันเป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน จะเห็นได้จากตำแหน่งที่ (2, 4) แทนด้วยเลข 1

3) คุณลักษณะที่ 3 การบวกลบจำนวนคละที่มีตัวส่วนเท่ากันเป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน จะเห็นได้จากตำแหน่งที่ (3, 4) แทนด้วยเลข 1

4) คุณลักษณะที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากันเป็นคุณลักษณะที่ไม่มีมาก่อนคุณลักษณะใดเลย จะเห็นได้จากทุกตำแหน่งแทนด้วยเลข 0

#### ข. ผลการกำหนด reachability matrix (R)

การกำหนด reachability matrix (R) มีวัตถุประสงค์เพื่อแสดงความสัมพันธ์ทางตรงและความสัมพันธ์ทางอ้อมระหว่างคุณลักษณะที่ใช้ในการบวกลบเศษส่วนทั้ง 4 คุณลักษณะ มีขนาดเมทริกซ์ เป็น  $4 \times 4$  ดังนี้

$$\begin{bmatrix} 1111 \\ 0101 \\ 0011 \\ 0000 \end{bmatrix}$$

จากเมทริกซ์ (R) ข้างต้นอธิบายความสัมพันธ์ทางตรงและความสัมพันธ์ทางอ้อมระหว่างคุณลักษณะ ได้ดังนี้

1) คุณลักษณะที่ 1 การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากันเป็นคุณลักษณะที่ต้องมีมาก่อนทุกคุณลักษณะ จะเห็นได้จากทุกตำแหน่งแทนด้วยเลข 1

2) คุณลักษณะที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากันเป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน จะเห็นได้จากตำแหน่งที่ (2, 2) และตำแหน่งที่ (2, 4) แทนด้วยเลข 1

3) คุณลักษณะที่ 3 การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน เป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน จะเห็นได้จากตำแหน่งที่ (3, 3) และตำแหน่งที่ (3, 4) แทนด้วยเลข 1

4) คุณลักษณะที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากันเป็นคุณลักษณะที่ไม่มีมาก่อนคุณลักษณะใดเลย จะเห็นได้จากทุกตำแหน่งแทนด้วยเลข 0

ค. ผลการกำหนด incidence matrix (Q)

การกำหนด *incidence matrix* (Q) มีวัตถุประสงค์เพื่อแสดงชุดของข้อสอบที่ต้องทำในแต่ละขั้นตอน ลักษณะขั้นตอนการทำข้อสอบที่เป็นไปได้มีจำนวน 15 ลักษณะ ( $2^k - 1$  เมื่อ k เป็นจำนวนคุณลักษณะ) มีขนาดเมทริกซ์ เป็น  $4 \times 15$  ดังนี้

$$\begin{bmatrix} 101010101010101 \\ 011001100110011 \\ 000111100001111 \\ 000000011111111 \end{bmatrix}$$

จากเมทริกซ์ (Q) ข้างต้นอธิบายลักษณะขั้นตอนการทำข้อสอบ ได้ดังนี้

- 1) ลักษณะที่ 1 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 เพียงชุดเดียว
- 2) ลักษณะที่ 2 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 2 เพียงชุดเดียว
- 3) ลักษณะที่ 3 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 และคุณลักษณะที่ 2
- 4) ลักษณะที่ 4 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 2 เพียงชุดเดียว
- 5) ลักษณะที่ 5 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 และคุณลักษณะที่ 3
- 6) ลักษณะที่ 6 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 2 และคุณลักษณะที่ 3
- 7) ลักษณะที่ 7 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 คุณลักษณะที่ 2 และคุณลักษณะที่ 3
- 8) ลักษณะที่ 8 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 4 เพียงชุดเดียว

- 9) ลักษณะที่ 9 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 และคุณลักษณะที่ 4
- 10) ลักษณะที่ 10 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 2 และคุณลักษณะที่ 4
- 11) ลักษณะที่ 11 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 คุณลักษณะที่ 2 และคุณลักษณะที่ 4
- 12) ลักษณะที่ 12 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 3 และคุณลักษณะที่ 4
- 13) ลักษณะที่ 13 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 คุณลักษณะที่ 3 และคุณลักษณะที่ 4
- 14) ลักษณะที่ 14 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 2 คุณลักษณะที่ 3 และคุณลักษณะที่ 4
- 15) ลักษณะที่ 15 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 คุณลักษณะที่ 2 คุณลักษณะที่ 3 และคุณลักษณะที่ 4

ง. ผลการกำหนด reduced Q matrix ( $Q_r$ )

การกำหนด *reduced Q matrix* ( $Q_r$ ) มีวัตถุประสงค์เพื่อลดจำนวนชุดของข้อสอบที่ต้องทำในแต่ละขั้นตอนตามเงื่อนไขโครงสร้างลำดับชั้นที่กำหนด *matrix*  $Q_r$  แสดงได้ดังนี้

$$\begin{bmatrix} 11111 \\ 01011 \\ 00111 \\ 00001 \end{bmatrix}$$

จากเมทริกซ์ ( $Q_r$ ) ข้างต้นจะเห็นได้ว่าลักษณะการทำข้อสอบที่เป็นไปตามเงื่อนไขลำดับชั้นที่กำหนดมีเพียง 5 ลักษณะ ดังนี้

- 1) ลักษณะที่ 1 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 เพียงชุดเดียว
- 2) ลักษณะที่ 2 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 และคุณลักษณะที่ 2
- 3) ลักษณะที่ 3 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 และคุณลักษณะที่ 3

4) ลักษณะที่ 4 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 คุณลักษณะที่ 2 และคุณลักษณะที่ 3

5) ลักษณะที่ 5 ทำข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 คุณลักษณะที่ 2 คุณลักษณะที่ 3 และคุณลักษณะที่ 4

#### ผลการกำหนดแบบแผนการตอบข้อสอบที่คาดหวัง

แบบแผนการตอบข้อสอบที่คาดหวังกำหนดได้จาก *reduced Q matrix* (Q) ซึ่งกำหนดได้ จำนวน 6 แบบแผน ดังนี้

แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 1 (0000) ผู้สอบทำข้อสอบได้ไม่ถูกต้องทั้งหมด แสดงว่าผู้สอบไม่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วน

แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 2 (1000) ผู้สอบทำข้อสอบได้ถูกต้องจำนวน 1 คุณลักษณะ คือ คุณลักษณะที่ 1 แสดงว่าผู้สอบมีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแต่ที่มีตัวส่วนเท่ากัน

แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 3 (1100) ผู้สอบทำข้อสอบได้ถูกต้องจำนวน 2 คุณลักษณะ ได้แก่ คุณลักษณะที่ 1 และคุณลักษณะที่ 2 แสดงว่าผู้สอบมีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแต่ที่มีตัวส่วนเท่ากันและการบวกลบเศษส่วนแต่ที่มีตัวส่วนไม่เท่ากัน

แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 4 (1010) ผู้สอบทำข้อสอบได้ถูกต้องจำนวน 2 คุณลักษณะ ได้แก่ คุณลักษณะที่ 1 และคุณลักษณะที่ 3 แสดงว่าผู้สอบมีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแต่ที่มีตัวส่วนเท่ากันและการบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน

แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 5 (1110) ผู้สอบทำข้อสอบได้ถูกต้องจำนวน 3 คุณลักษณะ ได้แก่ คุณลักษณะที่ 1 คุณลักษณะที่ 2 และคุณลักษณะที่ 3 แสดงว่าผู้สอบมีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแต่ที่มีตัวส่วนเท่ากัน การบวกลบเศษส่วนแต่ที่มีตัวส่วนไม่เท่ากันและการบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน

แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 6 (1111) ผู้สอบทำข้อสอบได้ถูกต้องจำนวน 4 คุณลักษณะ ได้แก่ คุณลักษณะที่ 1 คุณลักษณะที่ 2 คุณลักษณะที่ 3 และคุณลักษณะที่ 4 แสดงว่าผู้สอบมีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแต่ที่มีตัวส่วนเท่ากัน การบวกลบเศษส่วนแต่ที่มีตัวส่วนไม่เท่ากัน การบวกลบจำนวนคละที่มีตัวส่วนเท่ากันและการบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน

แบบแผนการตอบข้อสอบที่คาดหวัง จำนวนคุณลักษณะที่ทำข้อสอบได้ ถูกต้อง และสถานะความรู้ของผู้สอบแสดงในตารางที่ 12

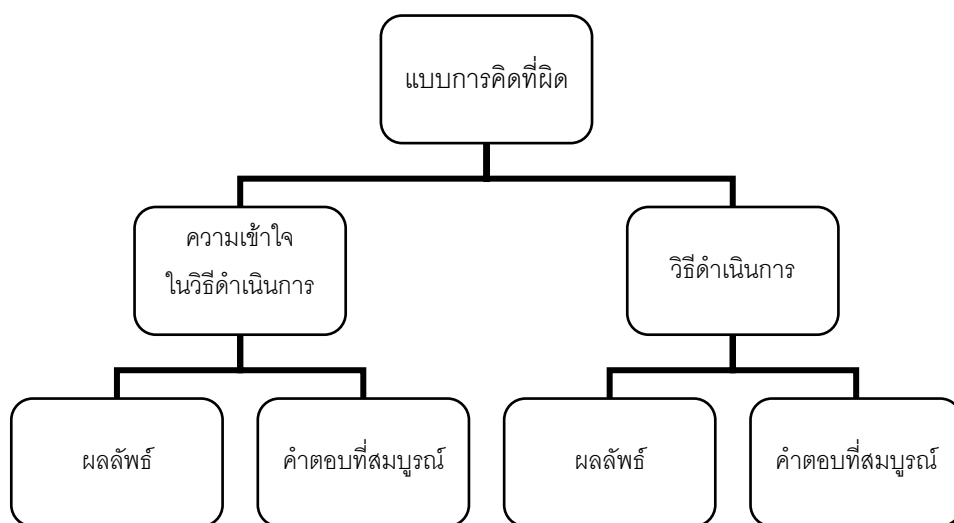
ตารางที่ 12 เมทริกซ์แสดงแบบแผนการตอบข้อสอบที่คาดหวัง จำนวนคุณลักษณะที่ทำข้อสอบได้ ถูกต้องและสถานะความรู้ของผู้สอบ

ผู้สอบ	เมทริกซ์แสดงแบบแผนการตอบข้อสอบที่คาดหวัง	จำนวนคุณลักษณะที่ทำข้อสอบได้ถูกต้อง	สถานะความรู้ของผู้สอบ
1	0000	0	ไม่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วน
2	1000	1	มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแต่ที่มีตัวส่วนเท่ากัน
3	1100	2	มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแต่ที่มีตัวส่วนเท่ากันและการบวกลบเศษส่วนแต่ที่มีตัวส่วนไม่เท่ากัน
4	1010	2	มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแต่ที่มีตัวส่วนเท่ากันและการบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน
5	1110	3	มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแต่ที่มีตัวส่วนเท่ากัน การบวกลบเศษส่วนแต่ที่มีตัวส่วนไม่เท่ากันและการบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน
6	1111	4	มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแต่ที่มีตัวส่วนเท่ากัน การบวกลบเศษส่วนแต่ที่มีตัวส่วนไม่เท่ากัน การบวกลบจำนวนคละที่มีตัวส่วนเท่ากันและการบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน

### 2.1.3 ผลการกำหนดแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน

การกำหนดแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน ดำเนินการโดยให้นักเรียนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) ที่เรียนเรื่องการบวกลบเศษส่วนแล้ว จากโรงเรียนขนาดเล็ก โรงเรียนขนาดกลางและโรงเรียนขนาดใหญ่ ขนาดโรงเรียนละ 70 คน รวมจำนวน 210 คน ทำแบบทดสอบเรื่องการบวกลบเศษส่วน จำนวน 16 ข้อ โดยให้แสดงวิธีทำ ผลการทดสอบพบว่า มีนักเรียนที่แสดงวิธีทำซึ่งสามารถนำมาวิเคราะห์แบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนได้ จำนวน 158 คน คิดเป็นร้อยละ 75.24 จำแนกเป็นนักเรียนจากโรงเรียนขนาดเล็ก จำนวน 66 คน (ร้อยละ 41.77) นักเรียนจากโรงเรียนขนาดกลาง จำนวน 42 คน (ร้อยละ 26.58) และนักเรียนจากโรงเรียนขนาดใหญ่ จำนวน 50 คน (ร้อยละ 31.65)

ผลจากการสำรวจแบบการคิดที่ผิดพบว่า สามารถจำแนกแบบการคิดที่ผิดได้เป็น 2 ส่วน ส่วนที่ 1 เป็นความเข้าใจในวิธีการดำเนินการ และส่วนที่ 2 เป็นวิธีการดำเนินการ โดยมีวัตถุประสงค์ในแต่ละส่วน จำแนกได้ 2 ประการ ประการแรกเพื่อให้ได้ผลลัพธ์และประการที่ 2 เพื่อให้ได้คำตอบที่สมบูรณ์ แบบการคิดที่ผิดทั้ง 2 ส่วนจะประกอบเป็นแบบการคิดที่ผิดในแต่ละคุณลักษณะ โดยมีส่วนที่ 1 เป็นแบบการคิดที่ผิดหลัก และส่วนที่ 2 เป็นแบบการคิดที่ผิดย่อย แบบการคิดที่ผิดแต่ละแบบอาจมีวัตถุประสงค์เพียงวัตถุประสงค์เดียวหรือมากกว่า ทั้งนี้ขึ้นอยู่กับระดับความซับซ้อนของคุณลักษณะ ตัวอย่างแบบการคิดที่ผิด สำหรับคุณลักษณะที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน เช่น ไม่เข้าใจวิธีการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน โดยใช้วิธีเศษบวกลบเศษ ตัวส่วนบวกลบตัวส่วน หรือใช้วิธีเศษส่วนตัวตั้งบวกลบกันเป็นเศษ เศษส่วนตัวบวกลบบวกลบกันเป็นตัวส่วน เป็นต้น แผนผังแนวคิดในการจำแนกแบบการคิดที่ผิดแสดงในแผนภาพที่ 5



แผนภาพที่ 5 แผนผังแนวคิดในการจำแนกแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน



รายละเอียดแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนสำหรับแต่ละคุณลักษณะและสภาพทั่วไปของกลุ่มตัวอย่างในการประเมินความครบถ้วนและความครอบคลุมของแบบการคิดที่ผิด แสดงในภาคผนวก ง และภาคผนวก ฉ ตามลำดับ

#### 2.1.4 ผลการจัดทำแบบทดสอบแบบเขียนตอบบนกระดาษคำตอบ

การจัดทำแบบทดสอบแบบเขียนตอบบนกระดาษคำตอบมีวัตถุประสงค์เพื่อให้ได้แบบทดสอบสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน จากการดำเนินการ พบว่า จัดฉบับแบบทดสอบแบบเขียนตอบบนกระดาษคำตอบได้จำนวน 20 ฉบับ ฉบับละ 32 ข้อ จำแนกเป็นข้อสอบจากชุดของข้อสอบลำดับชั้นละ 8 ข้อ ประกอบด้วยข้อสอบเฉพาะจำนวน 6 ข้อและข้อสอบร่วมจำนวน 2 ข้อ นำแบบทดสอบไปทดลองใช้กับกลุ่มตัวอย่างจากโรงเรียนขนาดเล็ก โรงเรียนขนาดกลางและโรงเรียนขนาดใหญ่ ขนาดละ 30 คน รวมจำนวน 1,800 คน โดยทดลองใช้แบบทดสอบฉบับละ 90 คน วิเคราะห์ค่าพารามิเตอร์ของข้อสอบโดยใช้โมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และวิเคราะห์ค่าความเที่ยงของแบบทดสอบโดยวิธีตรวจสอบความสอดคล้องภายใน (internal consistency method) ด้วยวิธีของคูเดอร์-ริชาร์ดสัน (Kuder-richardson estimates) คำนวณโดยใช้โปรแกรม ITEM ANALYSIS SYSTEM: IAS 2006 พัฒนาโดย ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี และรองศาสตราจารย์ ดร. คณิต ไช้มุกด์ นำเสนอผลการวิเคราะห์ข้อมูลเป็น 2 ส่วน ส่วนที่ 1 ค่าพารามิเตอร์ของข้อสอบและส่วนที่ 2 ค่าความเที่ยงของแบบทดสอบ มีรายละเอียด ดังนี้

##### ก. ค่าพารามิเตอร์ของข้อสอบ

ค่าพารามิเตอร์ของข้อสอบส่วนใหญ่อยู่ในเกณฑ์ดี พิจารณาได้จากค่าเฉลี่ยของค่าอำนาจจำแนกมีค่ามากกว่า 0.80 ทุกฉบับ ฉบับที่มีค่าเฉลี่ยของค่าอำนาจจำแนกมากที่สุด คือ ฉบับที่ 11 (1.23) ฉบับที่มีค่าเฉลี่ยของค่าอำนาจจำแนกน้อยที่สุด คือ ฉบับที่ 2 (0.82) เมื่อพิจารณาค่าการเดา พบว่า มีค่าเฉลี่ยของค่าการเดาน้อยกว่า 0.30 ทุกฉบับ ฉบับที่มีค่าเฉลี่ยของค่าการเดามากที่สุด คือ ฉบับที่ 8 (0.22) ฉบับที่มีค่าเฉลี่ยของค่าการเดาน้อยที่สุด คือ ฉบับที่ 12 (0.14) ส่วนค่าเฉลี่ยของค่าความยาก พบว่า ส่วนใหญ่มีค่าอยู่ระหว่าง 0.57 ถึง 1.44 ซึ่งอยู่ในเกณฑ์ยากง่ายปานกลาง (12 ฉบับ) นอกนั้นมีค่าเฉลี่ยของค่าความยากอยู่ระหว่าง -0.14 ถึง 0.24 ซึ่งอยู่ในเกณฑ์ง่าย (8 ฉบับ) ดังแสดงในตารางที่ 13



ข. ค่าความเที่ยงของแบบทดสอบ

แบบทดสอบแบบเขียนตอบบนกระดาษคำตอบทั้ง 20 ฉบับ มีค่าความเที่ยงสูง โดยมีค่าความเที่ยงอยู่ระหว่าง 0.82 ถึง 0.95 ฉบับที่มีค่าความเที่ยงต่ำที่สุด คือ ฉบับที่ 7 ฉบับที่มีค่าความเที่ยงสูงที่สุด คือ ฉบับที่ 11 ส่วนใหญ่มีค่าความเที่ยงมากกว่า 0.90 (15 ฉบับ) ดังแสดงในตารางที่ 13

ตารางที่ 13 ค่าพหามิตอร์ของข้อสอบและค่าความเที่ยงของแบบทดสอบ จำแนกตามฉบับ

ฉบับที่	ค่าอำนาจจำแนก		ค่าความยาก		ค่าการเดา		ค่าความเที่ยง (KR20)
	$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.	
1	1.00	0.22	-0.14	1.90	0.19	0.05	0.91
2	0.82	0.23	-0.21	1.58	0.20	0.04	0.92
3	0.94	0.44	0.82	2.10	0.20	0.04	0.88
4	1.04	0.24	0.12	1.75	0.19	0.06	0.90
5	0.97	0.28	0.61	1.94	0.19	0.05	0.92
6	1.03	0.25	0.82	1.50	0.17	0.05	0.86
7	1.05	0.33	0.64	1.75	0.20	0.06	0.82
8	0.88	0.18	0.20	1.86	0.22	0.06	0.90
9	1.11	0.65	1.04	1.20	0.17	0.05	0.89
10	1.01	0.23	0.24	1.72	0.17	0.06	0.91
11	1.23	0.33	1.44	1.48	0.19	0.05	0.95
12	1.06	0.19	0.57	1.67	0.14	0.06	0.91
13	1.04	0.23	0.74	2.17	0.16	0.05	0.90
14	1.11	0.41	0.85	1.97	0.17	0.06	0.83
15	0.99	0.19	0.00	2.09	0.20	0.06	0.92
16	1.02	0.27	-0.02	2.12	0.20	0.05	0.90
17	0.84	0.26	0.83	1.10	0.20	0.03	0.94
18	0.89	0.25	-0.17	1.51	0.17	0.06	0.91
19	1.00	0.50	1.43	1.50	0.19	0.03	0.91
20	1.09	0.23	1.11	1.78	0.17	0.05	0.90

### 2.1.5 ผลการจัดทำคลังข้อสอบ

การจัดทำคลังข้อสอบดำเนินการโดยแปลงค่าพารามิเตอร์ของข้อสอบแต่ละฉบับที่ได้จากการวิเคราะห์ให้อยู่บนสเกลเดียวกัน (common scale) โดยเทียบคะแนนตามแบบแผนการใช้ข้อสอบร่วม (anchor item design) ซึ่งหาค่าคงที่สำหรับการแปลงคะแนนด้วยวิธีค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐาน (mean and sigma method) คัดเลือกข้อสอบที่มีค่าอำนาจจำแนก อยู่ระหว่าง 0.50 ถึง 2.50 ค่าความยาก อยู่ระหว่าง -2.50 ถึง 2.50 ค่าการเดาไม่เกิน 0.30 ได้ข้อสอบจำนวน 386 ข้อ และได้คัดเลือกข้อสอบที่มีค่าความยากระหว่าง -3.00 ถึง -2.51 และ 2.49 ถึง 3.00 เพิ่มเติม ได้ข้อสอบจำนวน 52 ข้อ ทั้งนี้เพื่อให้ได้ข้อสอบที่มีค่าความยากครอบคลุมระดับความสามารถของผู้สอบ รวมข้อสอบในคลังข้อสอบทั้งหมด จำนวน 438 ข้อ นำเสนอผลการจัดทำคลังข้อสอบเป็น 5 ส่วน ส่วนที่ 1 จำนวนข้อสอบในคลัง ส่วนที่ 2 ค่าพารามิเตอร์ของข้อสอบ ส่วนที่ 3 ลักษณะการกระจายของค่าความยากของข้อสอบ ส่วนที่ 4 ค่าความน่าจะเป็นโดยเฉลี่ยในการตอบข้อสอบ และส่วนที่ 5 ค่าฟังก์ชันสารสนเทศของคลังข้อสอบ มีรายละเอียด ดังนี้

#### ก. จำนวนข้อสอบในคลัง

ข้อสอบในคลังมี จำนวน 438 ข้อ จำแนกเป็นข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน จำนวน 105 ข้อ คุณลักษณะที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน จำนวน 115 ข้อ คุณลักษณะที่ 3 การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน จำนวน 106 ข้อ และคุณลักษณะที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน จำนวน 112 ข้อ

#### ข. ค่าพารามิเตอร์ของข้อสอบในคลัง

ค่าพารามิเตอร์ของข้อสอบในคลัง โดยเฉลี่ยอยู่ในเกณฑ์ดี พิจารณาได้จากค่าเฉลี่ยของค่าอำนาจจำแนกมีค่ามากกว่า 1.00 ทุกคุณลักษณะ โดยมีค่าเฉลี่ยของค่าการเดาต่ำกว่า 0.20 ทุกคุณลักษณะเช่นกัน ส่วนค่าเฉลี่ยของค่าความยากมีค่าอยู่ระหว่าง -0.82 ถึง 1.17 ซึ่งอยู่ในเกณฑ์ง่ายถึงค่อนข้างยาก รายละเอียดดังแสดงในตารางที่ 14

ตารางที่ 14 ค่าพารามิเตอร์ของข้อสอบในคลัง จำแนกตามคุณลักษณะ

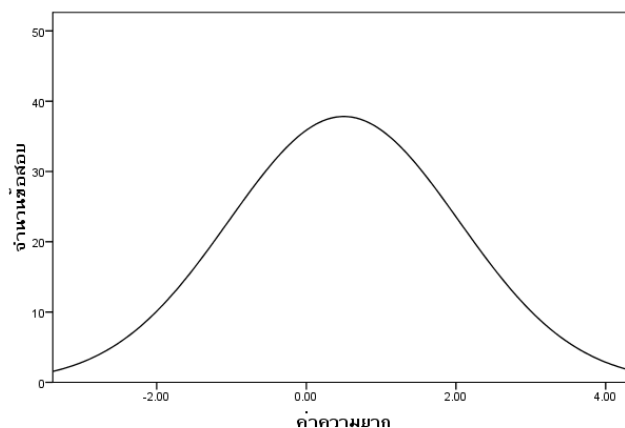
คุณลักษณะที่	ค่าอำนาจจำแนก		ค่าความยาก		ค่าการเดา	
	$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.
1	1.10	0.55	-0.82	1.62	0.19	0.05
2	1.02	0.21	0.91	1.18	0.17	0.07
3	1.03	0.25	0.65	1.46	0.17	0.07
4	1.03	0.23	1.17	1.05	0.18	0.07

ค. ลักษณะการกระจายของค่าความยากของข้อสอบในคลังข้อสอบ

ลักษณะการกระจายค่าความยากของข้อสอบในคลังข้อสอบนำเสนอในภาพรวมและจำแนกตามคุณลักษณะ ดังนี้

1) การกระจายค่าความยากของข้อสอบในคลังข้อสอบในภาพรวม

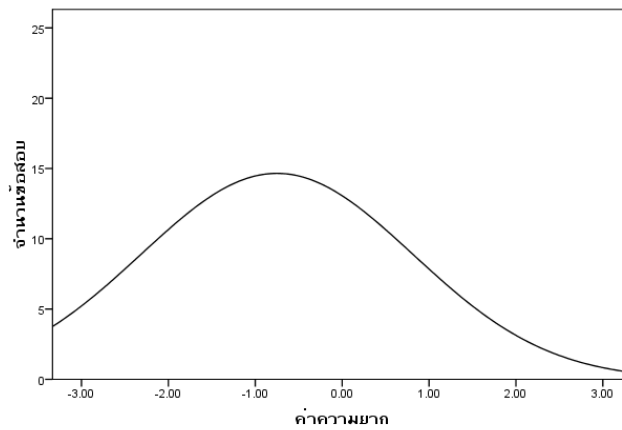
ค่าความยากของข้อสอบในคลังข้อสอบ ในภาพรวม มีค่าอยู่ระหว่าง -2.95 ถึง 2.99 มีลักษณะโค้งการแจกแจงค่อนข้างสมมาตร แสดงให้เห็นว่า ข้อสอบในคลังมีค่าความยากอยู่ในระดับปานกลาง ดังแสดงในแผนภาพที่ 6



แผนภาพที่ 6 ลักษณะการกระจายค่าความยากของข้อสอบในคลังข้อสอบ ในภาพรวม

2) การกระจายค่าความยากของข้อสอบในคลังข้อสอบ สำหรับคุณลักษณะที่ 1

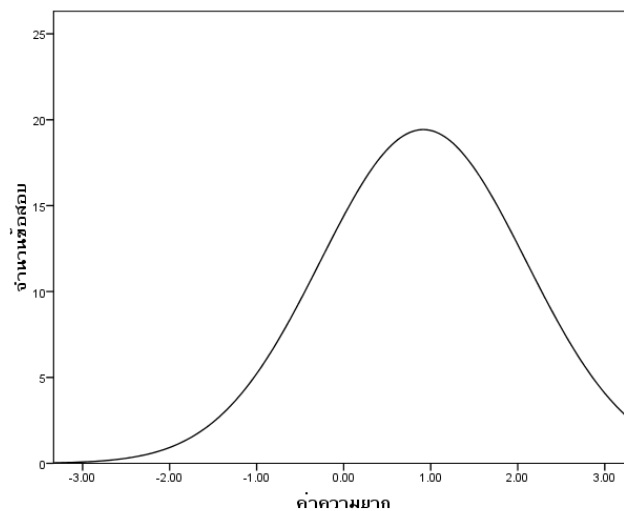
ค่าความยากของข้อสอบในคลังข้อสอบสำหรับคุณลักษณะที่ 1 มีค่าอยู่ระหว่าง -2.95 ถึง 2.90 มีลักษณะโค้งการแจกแจงเบ้ขวาเล็กน้อย แสดงให้เห็นว่า ข้อสอบสำหรับคุณลักษณะที่ 1 มีค่าความยาก อยู่ในระดับปานกลางถึงค่อนข้างง่าย ดังแสดงในแผนภาพที่ 7



แผนภาพที่ 7 ลักษณะการกระจายค่าความยากของข้อสอบในคลังข้อสอบ สำหรับคุณลักษณะที่ 1

3) การกระจายของค่าความยากของข้อสอบในคลังข้อสอบ สำหรับคุณลักษณะที่ 2

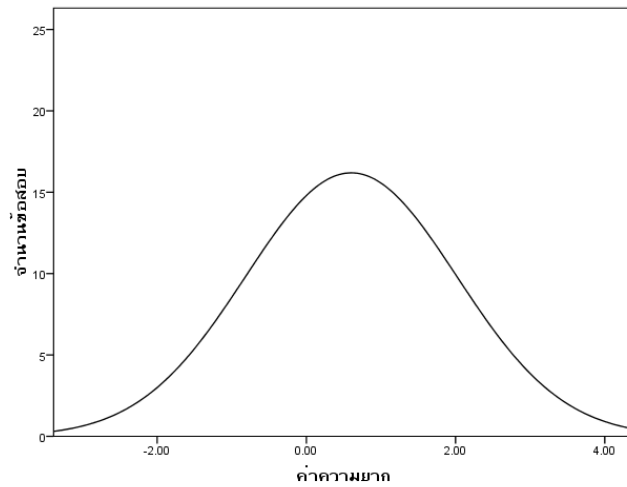
ค่าความยากของข้อสอบในคลังข้อสอบ สำหรับคุณลักษณะที่ 2 มีค่าอยู่ระหว่าง -2.70 ถึง 2.95 มีลักษณะโค้งการแจกแจงเบ้ซ้ายเล็กน้อย แสดงให้เห็นว่า ข้อสอบสำหรับคุณลักษณะที่ 2 มีค่าความยาก อยู่ในระดับปานกลางถึงค่อนข้างยาก ดังแสดงในแผนภาพที่ 8



แผนภาพที่ 8 ลักษณะการกระจายค่าความยากของข้อสอบในคลังข้อสอบ สำหรับคุณลักษณะที่ 2

4) การกระจายของค่าความยากของข้อสอบในคลังข้อสอบ สำหรับคุณลักษณะที่ 3

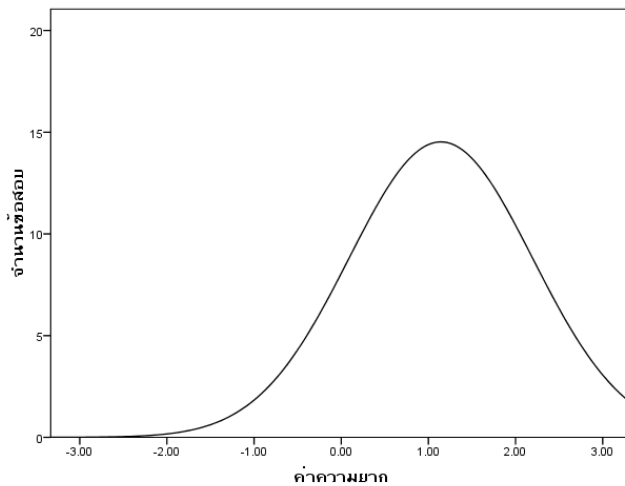
ค่าความยากของข้อสอบในคลังข้อสอบสำหรับคุณลักษณะที่ 3 มีค่าอยู่ระหว่าง -2.85 ถึง 2.99 มีลักษณะโค้งการแจกแจงค่อนข้างสมมาตร แสดงให้เห็นว่า ข้อสอบสำหรับคุณลักษณะที่ 3 มีค่าความยากอยู่ในระดับปานกลาง ดังแสดงในแผนภาพที่ 9



แผนภาพที่ 9 ลักษณะการกระจายค่าความยากของข้อสอบในคลังข้อสอบ สำหรับคุณลักษณะที่ 3

5) การกระจายของค่าความยากของข้อสอบในคลังข้อสอบ สำหรับคุณลักษณะที่ 4

ค่าความยากของข้อสอบในคลังข้อสอบ สำหรับคุณลักษณะที่ 4 มีค่าอยู่ระหว่าง -2.04 ถึง 2.99 มีลักษณะโค้งการแจกแจงเบ้ซ้ายเล็กน้อย แสดงให้เห็นว่า ข้อสอบสำหรับคุณลักษณะที่ 4 มีค่าความยากอยู่ในระดับปานกลางถึงค่อนข้างยาก ดังแสดงในแผนภาพที่ 10



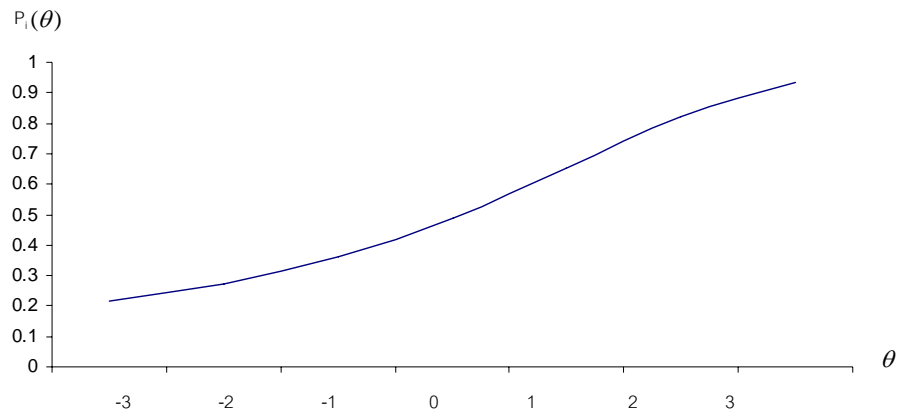
แผนภาพที่ 10 ลักษณะการกระจายค่าความยากของข้อสอบในคลังข้อสอบ สำหรับคุณลักษณะที่ 4

ง. ค่าความน่าจะเป็นโดยเฉลี่ยในการตอบข้อสอบ

1) ค่าความน่าจะเป็นโดยเฉลี่ยในการตอบข้อสอบในคลังข้อสอบในภาพรวม

ค่าความน่าจะเป็นโดยเฉลี่ยในการตอบข้อสอบในคลังข้อสอบในภาพรวมมีค่าลักษณะข้อสอบที่เขียนด้วยฟังก์ชันโลจิสติกสอดคล้องกับโมเดลการตอบสนองข้อสอบ

แบบ 3 พารามิเตอร์ แสดงให้เห็นว่าผู้สอบที่มีระดับความสามารถต่ำมีโอกาสที่จะตอบข้อสอบได้ ถูกต่องน้อยกว่าผู้สอบที่มีระดับความสามารถสูง ดังแสดงในแผนภาพที่ 11

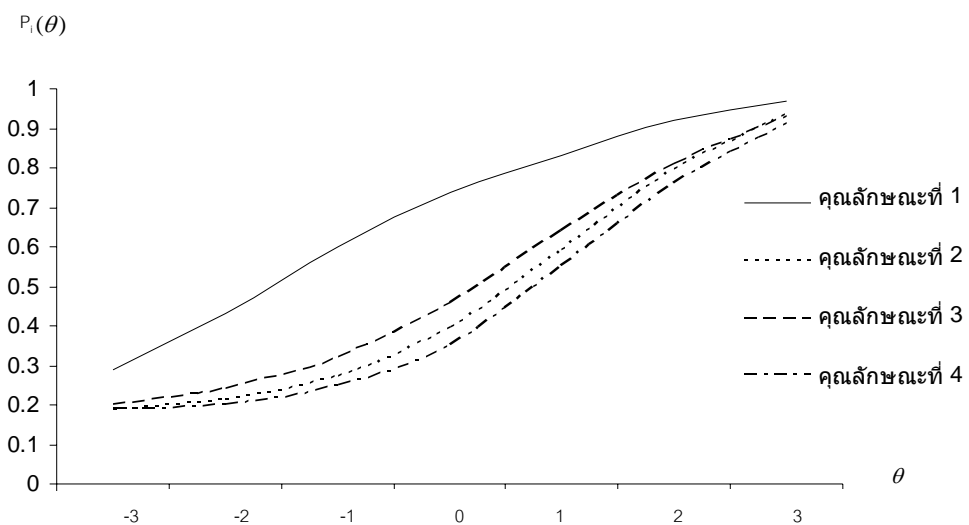


แผนภาพที่ 11 ค่าความน่าจะเป็นโดยเฉลี่ยในการตอบข้อสอบในคลังข้อสอบในภาพรวม

2) ค่าความน่าจะเป็นโดยเฉลี่ยในการตอบข้อสอบในคลังข้อสอบ

จำแนกตามคุณลักษณะ

ค่าความน่าจะเป็นโดยเฉลี่ยในการตอบข้อสอบในคลังข้อสอบ จำแนกตามคุณลักษณะมีไค้งลักษณะข้อสอบที่เขียนด้วยฟังก์ชันโลจิสติกคล้ายกับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ทุกคุณลักษณะ แสดงให้เห็นว่าผู้สอบที่มีระดับความสามารถต่ำมีโอกาสที่จะตอบข้อสอบในแต่ละคุณลักษณะได้ถูกต่องน้อยกว่าผู้สอบที่มีระดับความสามารถสูง ดังแสดงในแผนภาพที่ 12

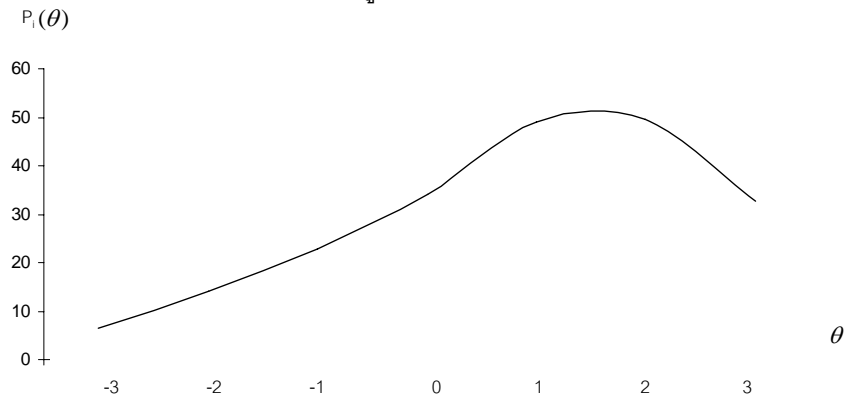


แผนภาพที่ 12 ค่าความน่าจะเป็นโดยเฉลี่ยในการตอบข้อสอบในคลังข้อสอบจำแนกตามคุณลักษณะ

๑. ค่าฟังก์ชันสารสนเทศของคลังข้อสอบ

1) ค่าฟังก์ชันสารสนเทศของคลังข้อสอบในภาพรวม

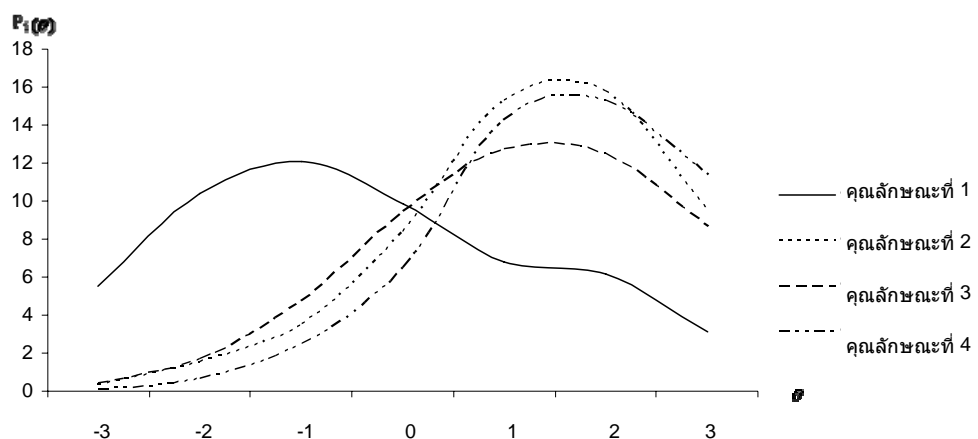
ค่าฟังก์ชันสารสนเทศของคลังข้อสอบมีลักษณะโค้งการแจกแจงเบ้ซ้ายเล็กน้อย แสดงให้เห็นว่าข้อสอบในคลังข้อสอบเหมาะกับผู้สอบที่มีระดับความสามารถปานกลางค่อนข้างไปทางระดับความสามารถสูง ดังแสดงในแผนภาพที่ 13



แผนภาพที่ 13 ค่าฟังก์ชันสารสนเทศของคลังข้อสอบ ในภาพรวม

2) ค่าฟังก์ชันสารสนเทศของคลังข้อสอบจำแนกตามคุณลักษณะ

ค่าฟังก์ชันสารสนเทศของคลังข้อสอบ ส่วนใหญ่มีลักษณะโค้งการแจกแจงเบ้ซ้ายแสดงให้เห็นว่าข้อสอบในคลังข้อสอบส่วนใหญ่เหมาะกับผู้สอบที่มีระดับความสามารถปานกลางค่อนข้างไปทางระดับความสามารถสูง ยกเว้นข้อสอบในคลังข้อสอบสำหรับคุณลักษณะที่ 1 ซึ่งมีลักษณะโค้งการแจกแจงเบ้ขวา แสดงให้เห็นว่าข้อสอบในคลังข้อสอบสำหรับคุณลักษณะที่ 1 เหมาะกับผู้สอบที่มีระดับความสามารถปานกลางค่อนข้างไปทางระดับความสามารถต่ำ ดังแสดงในแผนภาพที่ 14



แผนภาพที่ 14 ค่าฟังก์ชันสารสนเทศของคลังข้อสอบ จำแนกตามคุณลักษณะ

## 2.2 ปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ

โปรแกรมคอมพิวเตอร์ที่พัฒนาขึ้นมีการนำไปใช้ในการทดสอบแบบปรับเหมาะ 2 ขั้นตอน ขั้นตอนแรกเป็นการนำไปใช้เพื่อทดลองประสิทธิภาพการทำงานของโปรแกรมและขั้นตอนที่ 2 เป็นการนำไปใช้เพื่อประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิด การศึกษาปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ ดำเนินการโดยใช้แบบสอบถามเพื่อให้นักเรียนและครูที่รับผิดชอบดูแลการทดสอบระบุสถานะ สภาพและระดับความรุนแรงของปัญหาหลังจากการใช้โปรแกรมคอมพิวเตอร์ ปราบกฏผลการศึกษาปัญหาในแต่ละขั้นตอน ดังนี้

### 2.1.1 ปัญหาที่เกิดจากการทดลองใช้โปรแกรมคอมพิวเตอร์

ผลจากการทดลองใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ จำแนกได้ 2 ส่วน ส่วนที่ 1 เป็นผลการประเมิน ประกอบด้วยข้อมูลเบื้องต้นจากการทดสอบแบบปรับเหมาะและผลการวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน ส่วนที่ 2 เป็นผลการศึกษาปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ การนำเสนอในส่วนนี้จะนำเสนอเฉพาะผลการทดลองใช้โปรแกรมคอมพิวเตอร์ในส่วนที่ 2 และแนวทางการปรับปรุงแก้ไขโปรแกรมคอมพิวเตอร์จากปัญหาที่เกิดขึ้นเท่านั้น ส่วนผลประเมิน ในส่วนที่ 1 ได้นำเสนอไว้ในภาคผนวก ช มีรายละเอียดปัญหาที่เกิดจากการทดลองใช้โปรแกรมคอมพิวเตอร์และแนวทางการปรับปรุงแก้ไขโปรแกรม ดังนี้

#### ก. ปัญหาที่เกิดจากการทดลองใช้โปรแกรมคอมพิวเตอร์

การนำเสนอปัญหาที่เกิดจากการทดลองใช้โปรแกรมคอมพิวเตอร์ จำแนกเป็น 2 ส่วน ส่วนที่ 1 เป็นปัญหาตามความคิดเห็นของครู และส่วนที่ 2 เป็นปัญหาตามความคิดเห็นของนักเรียน มีรายละเอียด ดังนี้

##### 1) ปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ตามความคิดเห็นของครู

ครูที่ตอบแบบสอบถามปัญหาที่เกิดจากการทดลองใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะมีจำนวน 3 คน เป็นเพศหญิง 2 คน (ร้อยละ 66.7) และเพศชาย 1 คน (ร้อยละ 33.3) ทุกคนมีอายุ 46 ปี ขึ้นไป ด้านวุฒิการศึกษา พบว่ามีวุฒิการศึกษาปริญญาตรี 2 คน (ร้อยละ 66.7) และปริญญาโท 1 คน (ร้อยละ 33.3) สำเร็จการศึกษาในสาขาวิชา/วิชาเอก คณิตศาสตร์ 2 คน (ร้อยละ 66.7) และประถมศึกษา 1 คน (ร้อยละ 33.3)



ผลการวิเคราะห์สภาพและระดับปัญหาที่เกิดจากการทดลองใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะตามความคิดเห็นของครู พบว่า ครูทั้งหมดเห็นว่าไม่มีปัญหาในเกือบทุกด้าน ยกเว้นด้านการจัดรูปแบบ ด้านการพิมพ์ข้อมูลของผู้สอบและด้านการพิมพ์คำตอบ ซึ่งมีครูจำนวน 1-2 คน (ร้อยละ 33.3 – 66.7) เห็นว่ามีปัญหา เมื่อพิจารณาระดับปัญหา พบว่ามีปัญหาในระดับน้อย โดยมีค่าเฉลี่ยของปัญหาเท่ากับ 1.0 ทั้ง 3 ด้าน ส่วนด้านลักษณะของปัญหา ครูให้ความคิดเห็นว่า หน้าต่างแสดงผลมีขนาดเล็ก ช่องพิมพ์ข้อมูลของผู้สอบเว้นว่างไว้น้อยเกินไป ช่องพิมพ์คำตอบไม่ชัดเจนและนักเรียนมีปัญหาในการใช้คอมพิวเตอร์

2) ปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ตามความคิดเห็นของนักเรียน แสดงในตารางที่ 15

ตารางที่ 15 ข้อมูลทั่วไปของนักเรียนที่ตอบแบบสอบถามปัญหาที่เกิดจากการทดลองใช้โปรแกรมคอมพิวเตอร์

ข้อมูลทั่วไป	จำนวน	ร้อยละ
ขนาดโรงเรียน		
โรงเรียนขนาดเล็ก (จำนวนนักเรียนต่ำกว่า 200 คน)	18	33.3
โรงเรียนขนาดกลาง (จำนวนนักเรียน 201 ถึง 500 คน)	18	33.3
โรงเรียนขนาดใหญ่ (จำนวนนักเรียน 500 คนขึ้นไป)	18	33.3
ชั้น		
ประถมศึกษาปีที่ 4	15	27.8
ประถมศึกษาปีที่ 5	17	31.5
ประถมศึกษาปีที่ 6	22	40.7
เพศ		
หญิง	38	70.4
ชาย	16	29.6
อายุ		
10 ปี	4	7.4
11 ปี	28	51.9
12 ปี	20	37.0
13 ปี	1	1.9
14 ปี	1	1.9
รวม	54	100

จากตารางที่ 15 แสดงให้เห็นว่านักเรียนที่ตอบแบบสอบถามปัญหาที่เกิดจากการทดลองใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ มีจำนวน 54 คน ศึกษาอยู่ในโรงเรียนขนาดเล็ก โรงเรียนขนาดกลางและโรงเรียนขนาดใหญ่ ขนาดโรงเรียนละ 18 คน โดยศึกษาอยู่ในชั้นประถมศึกษาปีที่ 6 มากที่สุด (ร้อยละ 40.7) รองลงมา ได้แก่ ชั้นประถมศึกษาปีที่ 5 (ร้อยละ 31.5) และชั้นประถมศึกษาปีที่ 4 (ร้อยละ 27.8) นักเรียนส่วนใหญ่เป็นเพศหญิง (ร้อยละ 70.4) มีอายุ 11 ปี มากที่สุด (ร้อยละ 51.9) รองลงมา ได้แก่ 12 ปี (ร้อยละ 37.0) 10 ปี (ร้อยละ 7.4) 13 ปี และ 14 ปี (ร้อยละ 1.9) ตามลำดับ

ผลการวิเคราะห์สภาพและระดับปัญหาที่เกิดจากการทดลองใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ ตามความคิดเห็นของนักเรียน แสดงในตารางที่ 16

ตารางที่ 16 ปัญหาที่เกิดจากการทดลองใช้โปรแกรมคอมพิวเตอร์ตามความคิดเห็นของนักเรียน

ด้าน	สภาพปัญหา				ระดับปัญหา	
	มี		ไม่มี		$\bar{x}$	S.D.
	จำนวน	ร้อยละ	จำนวน	ร้อยละ		
1. การเข้าสู่โปรแกรม	2	3.7	52	96.3	1.50	0.70
2. การแสดงผลทางหน้าจอ						
2.1 ตัวอักษร (สี, ความชัดเจน..)	1	1.9	53	98.1	1.00	0.00
2.2 พื้นหลัง	3	5.6	51	94.4	1.33	0.57
2.3 การจัดรูปแบบ (ส่วนคำถาม, ส่วนคำตอบ, ข้อมูลอื่น ๆ .....	3	5.6	51	94.4	1.33	0.57
2.4 อื่น ๆ	2	3.7	52	96.3	1.00	0.00
3. การพิมพ์ข้อมูลของผู้สอบ	2	3.7	52	96.3	1.00	0.00
4. การคัดเลือกและนำเสนอข้อความ	4	7.4	50	92.6	1.25	0.50
5. การพิมพ์คำตอบ	11	20.4	43	79.6	1.18	0.40
6. การสั่งพิมพ์ผลการทดสอบทางเครื่องพิมพ์	3	5.6	51	94.4	1.00	0.00
7. อื่น ๆ	0	0.0	54	100	0.00	0.00

จากตารางที่ 16 แสดงให้เห็นว่านักเรียนส่วนใหญ่มีความคิดเห็นว่าการทดลองใช้คอมพิวเตอร์ในการทดสอบแบบปรับเหมาะไม่มีปัญหามากกว่ามีปัญหาในทุกด้าน โดยมีค่าร้อยละของการไม่มีปัญหา อยู่ระหว่าง 79.6 ถึง 98.1 ด้านที่ไม่มีปัญหามากที่สุด คือ ตัวอักษรที่ใช้แสดงผลทางหน้าจอ (ร้อยละ 98.1) ส่วนด้านที่มีปัญหามากที่สุด คือ การพิมพ์คำตอบ (ร้อยละ 20.4) เมื่อพิจารณาค่าเฉลี่ยของระดับปัญหาในแต่ละด้าน พบว่า อยู่ในระดับน้อยทุกด้าน โดยมีค่าเฉลี่ยของปัญหาอยู่ระหว่าง 1.00 ถึง 1.50 ดังแสดงในตารางที่ 16

ด้านลักษณะของปัญหา พบว่า ส่วนใหญ่ไม่ได้ระบุลักษณะของปัญหา ลักษณะของปัญหาที่นักเรียนระบุ ได้แก่ เข้าสู่โปรแกรมได้ยาก พิมพ์คำตอบได้ช้า พื้นหลังมีสีอ่อน ช่องพิมพ์คำตอบไม่ชัดเจน และส่งพิมพ์ผลทางเครื่องพิมพ์ได้ช้า

#### ข. การดำเนินการเพื่อปรับปรุงแก้ไขโปรแกรมคอมพิวเตอร์

ผลการศึกษาปัญหาที่เกิดจากการทดลองใช้โปรแกรมคอมพิวเตอร์ตามความคิดเห็นของครูและนักเรียน รวมทั้งจากการดำเนินการอื่น ๆ พบว่าปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์มีจำนวนและความรุนแรงของปัญหาอยู่ในระดับน้อย โดยจำแนกปัญหาที่เกิดขึ้นได้ 2 ลักษณะ ลักษณะที่ 1 เป็นปัญหาที่เกิดจากโปรแกรมคอมพิวเตอร์และเครื่องคอมพิวเตอร์ ได้แก่ หน้าต่างแสดงผลมีขนาดเล็ก ช่องพิมพ์ข้อมูลของผู้สอบเว้นว่างไว้น้อยเกินไป ช่องพิมพ์คำตอบไม่ชัดเจน เข้าสู่โปรแกรมได้ยาก พื้นหลังมีสีอ่อน และส่งพิมพ์ผลทางเครื่องพิมพ์ได้ช้า ลักษณะที่ 2 เป็นปัญหาที่เกิดจากผู้ใช้โปรแกรมคอมพิวเตอร์ ได้แก่ นักเรียนมีปัญหาในการใช้คอมพิวเตอร์ โดยเฉพาะเรื่องการพิมพ์ เพื่อให้โปรแกรมคอมพิวเตอร์สามารถใช้งานได้อย่างมีประสิทธิภาพยิ่งขึ้นจึงได้ดำเนินการปรับปรุงแก้ไขโปรแกรมคอมพิวเตอร์ ดังนี้

- 1) ด้านการเข้าสู่โปรแกรมและการส่งพิมพ์ผลทางเครื่องพิมพ์แก้ไขปรับปรุงรูปแบบคำสั่งและกำหนดลักษณะเฉพาะของคอมพิวเตอร์ที่ใช้ให้รองรับการทำงานของโปรแกรมได้มากขึ้น
- 2) ด้านการแสดงผลทางหน้าจอแก้ไขปรับปรุงสีของพื้นหลังและตัวอักษรให้มีความชัดเจนมากขึ้น
- 3) ด้านการพิมพ์ข้อมูลผู้ทดสอบแก้ไขปรับปรุงโดยเพิ่มช่องว่างในการพิมพ์ข้อมูลผู้ทดสอบให้กว้างขึ้น
- 4) ด้านการพิมพ์คำตอบแก้ไขปรับปรุงสีของช่องพิมพ์คำตอบให้ชัดเจนขึ้น

## 2.1.2 ปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ในการประเมินเชิงวิจิตร สถานะความรู้และแบบการคิดที่ผิด

การศึกษาปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ดำเนินการโดยใช้แบบสอบถามภายหลังจากนักเรียนได้รับการประเมินเชิงวิจิตรสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการรวบรวมเศษส่วน จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 2 โดยจำแนกการนำเสนอผลเป็น 2 ส่วน ส่วนที่ 1 เป็นปัญหาตามความคิดเห็นของครูและส่วนที่ 2 เป็นปัญหาตามความคิดเห็นของนักเรียน มีรายละเอียด ดังนี้

### ก. ปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ตามความคิดเห็นของครู

ครูที่ตอบแบบสอบถามปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะมีจำนวน 18 คน ส่วนใหญ่เป็นเพศหญิง (ร้อยละ 55.6) โดยมีอายุ 46 ปีขึ้นไปมากที่สุด (ร้อยละ 50.0) รองลงมา ได้แก่ 41 - 45 ปี ร้อยละ 27.8 และ 36 - 40 ปี (ร้อยละ 22.2) ด้านวุฒิการศึกษา พบว่ามีวุฒิการศึกษาปริญญาตรี 15 คน (ร้อยละ 83.3) และปริญญาโท 3 คน (ร้อยละ 16.7) สำเร็จการศึกษาในสาขาวิชา/วิชาเอก คณิตศาสตร์มากที่สุด (ร้อยละ 44.4) รองลงมา ได้แก่ ประถมศึกษา (ร้อยละ 38.9) และบริหารการศึกษา (ร้อยละ 16.7)

ผลการวิเคราะห์สภาพและระดับปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะตามความคิดเห็นของครู พบว่า ครูส่วนใหญ่เห็นว่าไม่มีปัญหามากกว่ามีปัญหาในทุกด้าน โดยมีค่าร้อยละของการไม่มีปัญหา เท่ากับ 100 เกือบทุกด้าน ยกเว้นด้านการพิมพ์ผลทางเครื่องพิมพ์ มีครูเห็นว่ามีปัญหา จำนวน 3 คน (ร้อยละ 16.7) เมื่อพิจารณาค่าเฉลี่ยของระดับปัญหา พบว่า อยู่ในระดับน้อยทุกด้าน ส่วนลักษณะของปัญหา พบว่า ครูเห็นว่าสั่งพิมพ์ผลทางเครื่องพิมพ์ได้ช้า

### ข. ปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ตามความคิดเห็นของนักเรียน

นักเรียนที่ตอบแบบสอบถามปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ มีจำนวน 540 คน มีแบบสอบถามที่สมบูรณ์สามารถนำมาวิเคราะห์ได้จำนวน 462 ฉบับ คิดเป็นร้อยละ 85.6 ผลการวิเคราะห์ข้อมูล พบว่า นักเรียนส่วนใหญ่ศึกษาอยู่ในโรงเรียนขนาดใหญ่ (ร้อยละ 35.7) รองลงมา ได้แก่ โรงเรียนขนาดกลาง (ร้อยละ 34.4) และโรงเรียนขนาดเล็ก (ร้อยละ 29.9) โดยศึกษาอยู่ในชั้นประถมศึกษาปีที่ 6 มากที่สุด (ร้อยละ 43.7) รองลงมา ได้แก่ ชั้นประถมศึกษาปีที่ 5 (ร้อยละ 42.6) และ ชั้นประถมศึกษาปีที่ 4 (ร้อยละ 13.6) นักเรียนส่วนใหญ่เป็นเพศหญิง (ร้อยละ 53.5) มีอายุ 11 ปี มากที่สุด (ร้อยละ 49.6) รองลงมา ได้แก่ 10 ปี 12 ปี 9 ปี 13 ปี คิดเป็นร้อยละ 30.5, 14.1, 3.2 และ 8 ปี 14 ปี 15 ปี คิดเป็นร้อยละ 0.2 ดังแสดงในตารางที่ 17

ตารางที่ 17 ข้อมูลทั่วไปของนักเรียนที่ตอบแบบสอบถามปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์

สภาพทั่วไป	จำนวน	ร้อยละ
ขนาดโรงเรียน		
โรงเรียนขนาดเล็ก	138	29.9
โรงเรียนขนาดกลาง	159	34.4
โรงเรียนขนาดใหญ่	165	35.7
ชั้น		
ประถมศึกษาปีที่ 4	63	13.6
ประถมศึกษาปีที่ 5	197	42.6
ประถมศึกษาปีที่ 6	202	43.7
เพศ		
หญิง	246	53.2
ชาย	216	46.8
อายุ		
8 ปี	1	0.2
9 ปี	15	3.2
10 ปี	141	30.5
11 ปี	229	49.6
12 ปี	65	14.1
13 ปี	9	1.9
14 ปี	1	0.2
15 ปี	1	0.2
รวม	462	100

ผลการวิเคราะห์สภาพและระดับปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ ตามความคิดเห็นของนักเรียน พบว่า นักเรียนส่วนใหญ่เห็นว่าไม่มีปัญหามากกว่ามีปัญหาในทุกด้าน โดยมีค่าร้อยละของการไม่มีปัญหา อยู่ระหว่าง 77.5 ถึง 92.6 ด้านที่ไม่มีปัญหามากที่สุด ยกเว้นด้าน อื่น ๆ คือ ด้านพื้นหลังของการแสดงผลทางหน้าจอ (ร้อยละ 89.2) ส่วนด้านที่มีปัญหามากที่สุด คือ การจัดรูปแบบการแสดงผลทางหน้าจอ (ร้อยละ 22.5) เมื่อพิจารณาค่าเฉลี่ยของระดับปัญหา พบว่า อยู่ในระดับน้อยทุกด้าน โดยมีค่าเฉลี่ยของปัญหาอยู่ระหว่าง 1.03 ถึง 1.19 ดังแสดงในตารางที่ 18

ด้านลักษณะของปัญหา พบว่า ส่วนใหญ่ไม่ได้ระบุลักษณะของปัญหา ลักษณะของปัญหาที่นักเรียนระบุส่วนใหญ่เป็นปัญหาเกี่ยวกับนักเรียน ได้แก่ พิมพ์ไม่เก่ง ไม่ค่อยได้ใช้คอมพิวเตอร์ เป็นต้น

ตารางที่ 18 ปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ตามความคิดเห็นของนักเรียน

ด้าน	สภาพปัญหา				ระดับปัญหา	
	มี		ไม่มี		$\bar{x}$	S.D.
	จำนวน	ร้อยละ	จำนวน	ร้อยละ		
1. การเข้าสู่โปรแกรม	97	21.0	365	79.0	1.15	0.36
2. การแสดงผลทางหน้าจอ						
2.1 ตัวอักษร (สี, ความชัดเจน..)	85	18.4	377	81.6	1.19	0.39
2.2 พื้นหลัง	50	10.8	412	89.2	1.14	0.35
2.3 การจัดรูปแบบ (ส่วนคำถาม, ส่วนคำตอบ, ข้อมูลอื่น ๆ .....	104	22.5	358	77.5	1.18	0.39
2.4 อื่น ๆ	57	12.3	405	87.7	1.19	0.40
3. การพิมพ์ข้อมูลของผู้สอบ	95	20.6	367	79.4	1.13	0.34
4. การคัดเลือกและนำเสนอข้อคำถาม	69	14.9	393	85.1	1.09	0.29
5. การพิมพ์คำตอบ	77	16.7	385	83.3	1.08	0.27
6. การส่งพิมพ์ผลการทดสอบทางเครื่องพิมพ์	77	16.7	385	83.3	1.13	0.34
7. อื่น ๆ	34	7.4	428	92.6	1.03	0.17

## ตอนที่ 2 ผลการศึกษาคุณภาพของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น

คุณภาพของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น หมายถึง การให้ผลการประเมินเชิงวินิจฉัยที่เชื่อถือได้ คุ่มค่าและสอดคล้องกับวัตถุประสงค์ของการประเมิน โดยพิจารณาจากความเที่ยงและความตรงของผลการประเมินเชิงวินิจฉัยจากวิธีการที่พัฒนาขึ้น รวมทั้งจำนวนข้อสอบและจำนวนเวลาที่ใช้ในการทดสอบ มีรายละเอียดผลการศึกษาคูณภาพแต่ละด้าน ดังนี้

### 1. ผลการศึกษาคความเที่ยงของผลการประเมินเชิงวินิจฉัยจากวิธีการที่พัฒนาขึ้น

ความเที่ยงของผลการประเมินเชิงวินิจฉัยจากวิธีการที่พัฒนาขึ้น หมายถึง คุณสมบัติของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นในด้านความคงที่ของผลการประเมินเชิงวินิจฉัย พิจารณาได้จาก

ความสอดคล้องระหว่างผลการประเมินเชิงวินิจฉัยจากการทดสอบซ้ำ (Test-retest reliability) จำแนก การศึกษาความเที่ยงออกเป็น 2 ด้าน ตามผลการประเมินเชิงวินิจฉัย ได้แก่ ความเที่ยงของผลการ ประเมินเชิงวินิจฉัยสถานะความรู้และความเที่ยงของผลการประเมินเชิงวินิจฉัยแบบการคิดที่ผิด ปราบกฎผลการศึกษาความเที่ยงในแต่ละด้าน ดังนี้

### ความเที่ยงของผลการประเมินเชิงวินิจฉัยสถานะความรู้

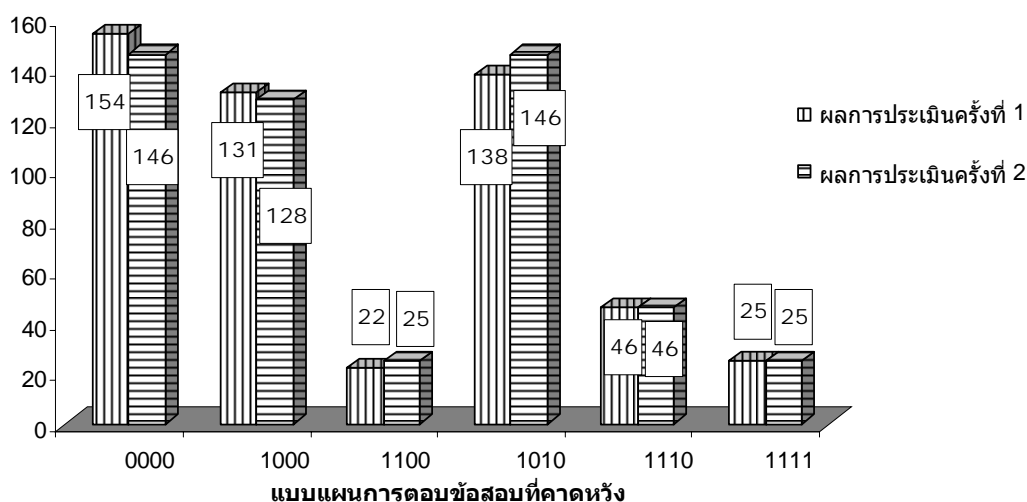
จากการศึกษาผลการประเมินเชิงวินิจฉัยสถานะความรู้ พบว่า นักเรียนมีแบบแผน การตอบข้อสอบที่คาดหวังจากการทดสอบครั้งที่ 1 และการทดสอบครั้งที่ 2 สอดคล้องกัน จำนวน 485 คน จากนักเรียนทั้งหมด จำนวน 516 คน คิดเป็นสัดส่วน เท่ากับ 0.94 จำนวนนักเรียนจำแนก ตามแบบแผนการตอบข้อสอบที่คาดหวังซึ่งวินิจฉัยได้จากการทดสอบครั้งที่ 1 และการทดสอบครั้งที่ 2 ดังแสดงในตารางที่ 19

ตารางที่ 19 จำนวนนักเรียนจำแนกตามแบบแผนการตอบข้อสอบที่คาดหวังซึ่งวินิจฉัยได้จากการ ทดสอบครั้งที่ 1 และการทดสอบครั้งที่ 2

		ผลการทดสอบครั้งที่ 2						
		0000	1000	1100	1010	1110	1111	รวม
ผลการทดสอบครั้งที่ 1	0000	139 (90.3%)	12 (7.8%)		3 (1.9%)			154 (29.8%)
	1000	6 (4.6%)	116 (88.5%)	3 (2.3%)	6 (4.6%)			131 (25.4%)
	1100			22 (100%)				22 (4.3%)
	1010	1 (0.7%)			137 (99.3%)			138 (26.7%)
	1110					46 (100%)		46 (8.9%)
	1111						25 (100%)	25 (4.8%)
	<b>รวม</b>	146 (28.3%)	128 (24.8%)	25 (4.8%)	146 (28.3%)	46 (8.9%)	25 (100%)	516 (100%)

จากผลการศึกษาค่าความเที่ยงของผลการประเมินเชิงวินิจฉัยสถานะความรู้ข้างต้นแสดงให้เห็นว่าวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นสามารถประเมินเชิงวินิจฉัยสถานะความรู้ได้อย่างมีความเที่ยง พิจารณาได้จากสัดส่วนของความสอดคล้องระหว่างผลการประเมินเชิงวินิจฉัยสถานะความรู้จากการทดสอบครั้งที่ 1 และการทดสอบครั้งที่ 2 ซึ่งแสดงถึงความคงที่ของผลการประเมินเชิงวินิจฉัยจากการทดสอบซ้ำ เมื่อจำแนกจำนวนนักเรียนในแต่ละแบบแผนการตอบข้อสอบที่คาดหวัง ปรากฏผลดังแสดงในแผนภูมิที่ 4

จำนวนนักเรียน



แผนภูมิที่ 4 สถานะความรู้ที่ใช้ในการบวกลบเศษส่วนของนักเรียน จำแนกตามแบบแผนการตอบข้อสอบที่คาดหวัง ซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 1 และครั้งที่ 2

จากแผนภูมิที่ 4 แสดงให้เห็นว่านักเรียนมีแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 1 (0000) มากที่สุด รองลงมา ได้แก่ แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 4 (1010) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 2 (1000) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 5 (1110) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 6 (1111) และแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 3 (1100) ตามลำดับ

เมื่อพิจารณาสถานะความรู้ที่ใช้ในการบวกลบเศษส่วนจากจำนวนนักเรียนในแต่ละแบบแผนการตอบข้อสอบที่คาดหวัง พบว่า สามารถจำแนกนักเรียนตามสถานะความรู้ได้ 3 กลุ่ม ดังนี้

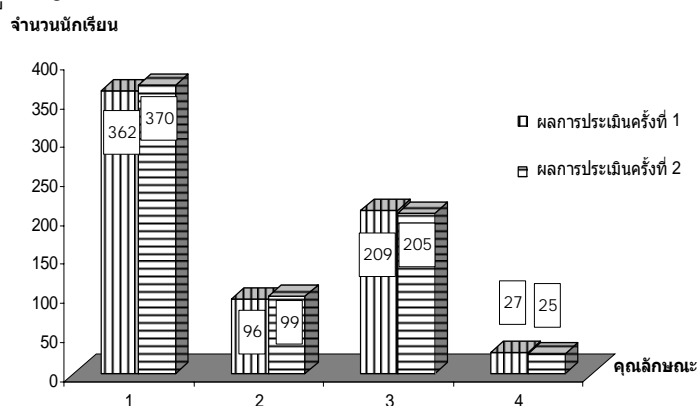
กลุ่มที่ 1 เป็นกลุ่มที่ไม่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนทุกเรื่อง ทั้งการบวกลบเศษส่วนที่มีตัวส่วนเท่ากันและไม่เท่ากัน รวมทั้งจำนวนคละที่มีตัวส่วนเท่ากันและไม่เท่ากัน จำนวน 154 คน คิดเป็นร้อยละ 29.9 และจำนวน 146 คน คิดเป็นร้อยละ 28.3 จากการประเมินครั้งที่ 1 และครั้งที่ 2 ตามลำดับ



กลุ่มที่ 2 เป็นกลุ่มที่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนเพียงบางส่วน จำนวน 337 คน คิดเป็นร้อยละ 65.3 และจำนวน 345 คน คิดเป็นร้อยละ 66.9 จากการประเมินครั้งที่ 1 และครั้งที่ 2 ตามลำดับ เมื่อพิจารณาสถานะความรู้ที่มี พบว่า นักเรียนส่วนใหญ่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนที่มีตัวส่วนเท่ากันมากกว่าการบวกลบเศษส่วนที่มีตัวส่วนไม่เท่ากัน ทั้งที่เป็นเศษส่วนแท้และจำนวนคละ จะเห็นได้จากมีจำนวนนักเรียนที่มีแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 2 และแบบแผนที่ 4 ซึ่งเป็นเรื่องการบวกลบเศษส่วนที่มีตัวส่วนเท่ากันมากกว่าแบบแผนการตอบข้อสอบที่คาดหวัง แบบแผนที่ 3 แบบแผนที่ 5 และแบบแผนที่ 6 ซึ่งเป็นเรื่องการบวกลบเศษส่วนที่มีตัวส่วนไม่เท่ากัน นอกจากนี้ยังมีจำนวนนักเรียนที่มีแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 4 และแบบแผนที่ 5 ซึ่งเป็นเรื่องการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากันและไม่เท่ากัน รวมทั้งจำนวนคละที่มีตัวส่วนเท่ากันมากกว่าแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 3 ซึ่งเป็นเรื่องการบวกเศษส่วนแท้ที่มีตัวส่วนเท่ากันและไม่เท่ากัน

กลุ่มที่ 3 เป็นกลุ่มที่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนทุกเรื่อง จำนวน 27 คน คิดเป็นร้อยละ 4.8 และจำนวน 25 คน คิดเป็นร้อยละ 4.3 จากการประเมินครั้งที่ 1 และการประเมินครั้งที่ 2 ตามลำดับ

เมื่อจำแนกจำนวนนักเรียนตามสถานะความรู้ในแต่ละคุณลักษณะ ปรากฏผลดังแสดงในแผนภูมิที่ 5



แผนภูมิที่ 5 สถานะความรู้ที่ใช้ในการบวกลบเศษส่วนของนักเรียน จำแนกตามคุณลักษณะ ซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 1 และครั้งที่ 2

จากแผนภูมิที่ 5 แสดงให้เห็นว่านักเรียนส่วนใหญ่มีความรู้ ทักษะและกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน (คุณลักษณะที่ 1) มากที่สุด รองลงมา คือ มีความรู้ ทักษะและกระบวนการที่ใช้ในการบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน (คุณลักษณะที่ 3) มีความรู้ ทักษะและกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่

เท่ากัน (คุณลักษณะที่ 2) และมีความรู้ ทักษะและกระบวนการที่ใช้ในการบวกจำนวนคละที่มีตัวส่วนไม่เท่ากัน (คุณลักษณะที่ 4) ตามลำดับ

เมื่อจำแนกจำนวนนักเรียนที่มีผลการประเมินสถานะความรู้สอดคล้องกันจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง จำแนกตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้ พบว่า นักเรียนที่ใช้เกณฑ์ในการตัดสินความรอบรู้ เกณฑ์ที่ 1 คือ ทำข้อสอบในแต่ละลำดับชั้นอย่างน้อย 3 ข้อและตอบข้อสอบได้ถูกต้องหรือใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 2 ใน 3 ส่วนใหญ่มีแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 1 (0000) มากที่สุด รองลงมา ได้แก่ แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 4 (1010) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 2 (1000) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 5 (1110) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 6 (1111) และแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 3 (1100)

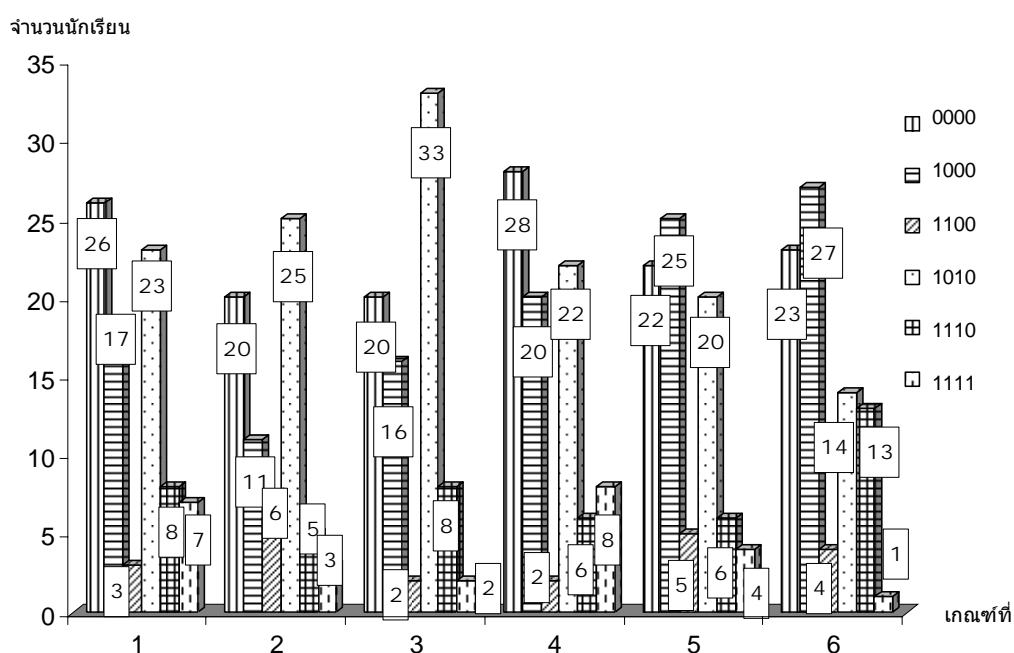
นักเรียนที่ใช้เกณฑ์ในการตัดสินความรอบรู้เกณฑ์ที่ 2 คือ ทำข้อสอบในแต่ละลำดับชั้นอย่างน้อย 4 ข้อและตอบข้อสอบได้ถูกต้องหรือใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 3 ใน 4 ส่วนใหญ่มีแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 4 (1010) มากที่สุด รองลงมา ได้แก่ แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 1 (0000) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 2 (1000) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 3 (1100) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 5 (1110) และแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 6 (1111)

นักเรียนที่ใช้เกณฑ์ในการตัดสินความรอบรู้เกณฑ์ที่ 3 คือ ทำข้อสอบในแต่ละลำดับชั้นอย่างน้อย 5 ข้อและตอบข้อสอบได้ถูกต้องหรือใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 4 ใน 5 ส่วนใหญ่มีแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 4 (1010) มากที่สุด รองลงมา ได้แก่ แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 1 (0000) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 2 (1000) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 5 (1110) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 3 (1100) และแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 6 (1111)

นักเรียนที่ใช้เกณฑ์ในการตัดสินความรอบรู้เกณฑ์ที่ 4 คือ ทำข้อสอบในแต่ละลำดับชั้นอย่างน้อย 3 ข้อและตอบข้อสอบได้ถูกต้องหรือใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 2 ใน 3 และมีความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3 ส่วนใหญ่มีแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 1 (0000) มากที่สุด รองลงมา ได้แก่ แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 4 (1010) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 2 (1000) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 6 (1111) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 5 (1110) และแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 3 (1100)

นักเรียนที่ใช้เกณฑ์ในการตัดสินความรอบรู้เกณฑ์ที่ 5 คือ ทำข้อสอบในแต่ละลำดับชั้นอย่างน้อย 4 ข้อและตอบข้อสอบได้ถูกต้องหรือใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 3 ใน 4 และมีความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3 มีแบบแผนการตอบข้อสอบเหมือนกับแบบแผนการตอบข้อสอบของนักเรียนที่ใช้เกณฑ์ในการตัดสินความรอบรู้เกณฑ์ที่ 6 คือ ทำข้อสอบในแต่ละลำดับชั้นอย่างน้อย 5 ข้อและตอบข้อสอบได้ถูกต้องหรือใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 4 ใน 5 และมีความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3 ซึ่งพบว่าส่วนใหญ่มีแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 2 (1000) มากที่สุด รองลงมา ได้แก่ แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 1 (0000) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 4 (1010) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 5 (1110) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 3 (1100) และแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 6 (1111)

ดังแสดงในแผนภูมิที่ 6



แผนภูมิที่ 6 สถานะความรู้ที่ใช้ในการบวกลบเศษส่วนของนักเรียน จำแนกตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้

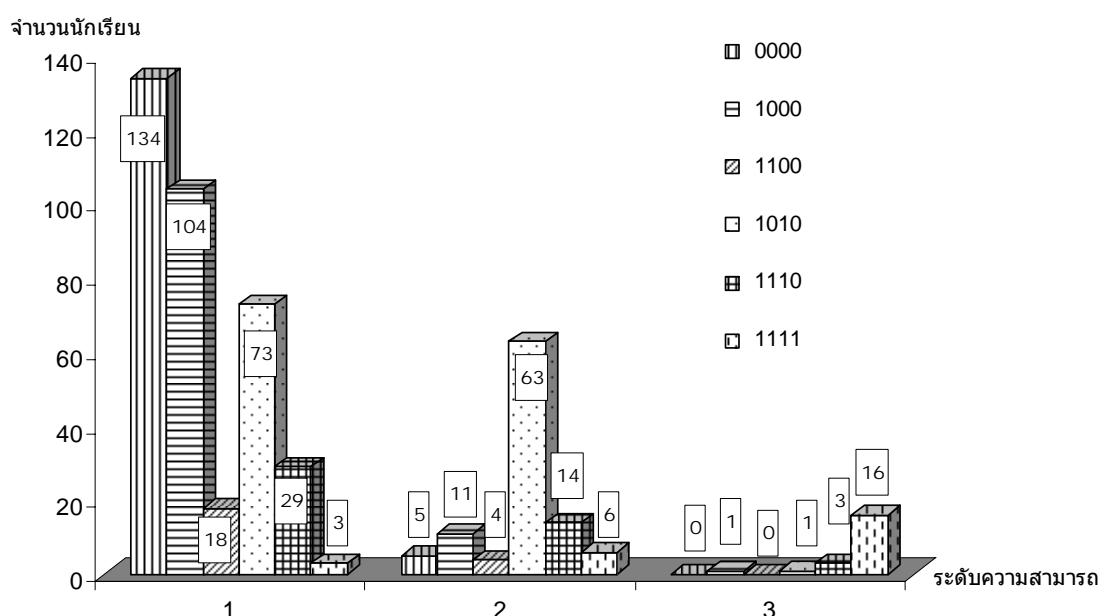
เมื่อจำแนกจำนวนนักเรียนที่มีผลการประเมินสถานะความรู้สอดคล้องกันจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง จำแนกตามระดับความสามารถของนักเรียนใน 3 ระดับ ได้แก่ ระดับต่ำ ( $\theta \leq -0.60$ ) ระดับปานกลาง ( $-0.59 \leq \theta \leq -0.59$ ) และระดับสูง ( $\theta \geq 0.60$ ) พบว่า นักเรียนที่มีความสามารถระดับต่ำ มีแบบแผนการตอบข้อสอบแบบแผนที่ 1 (0000)

มากที่สุด รองลงมา ได้แก่ แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 2 (1000) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 4 (1010) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 5 (1110) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 3 (1100) และแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 6 (1111)

นักเรียนที่มีความสามารถระดับปานกลาง มีแบบแผนการตอบข้อสอบแบบแผนที่ 4 (1010) มากที่สุด รองลงมา ได้แก่ แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 5 (1110) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 2 (1000) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 6 (1111) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 1 (0000) และแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 3 (1100)

นักเรียนที่มีความสามารถระดับสูง มีแบบแผนการตอบข้อสอบแบบแผนที่ 6 (1111) มากที่สุด รองลงมา ได้แก่ แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 5 (1110) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 2 (1000) และแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 4 (1010) ส่วนแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 1 (0000) และแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 3 (1100) ไม่มีนักเรียนที่มีความสามารถระดับสูงมีแบบแผนดังกล่าว

ดังแสดงในแผนภูมิที่ 7

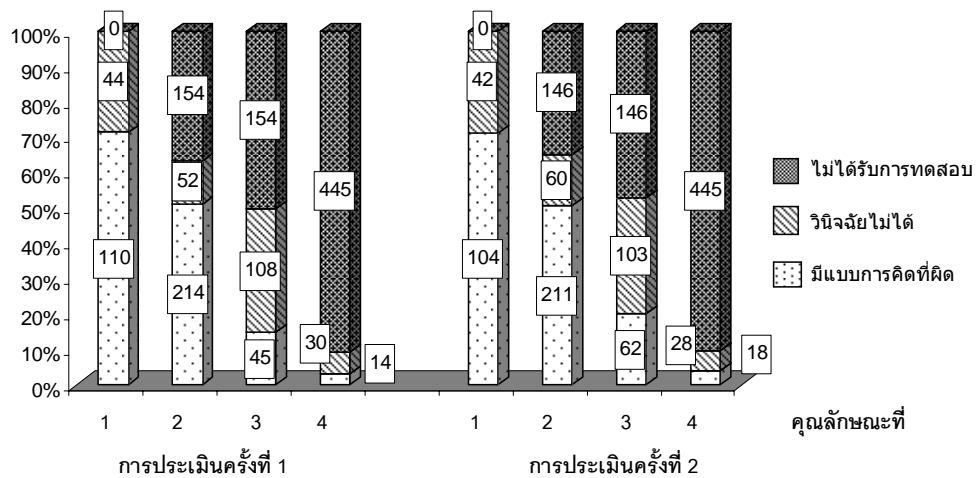


แผนภูมิที่ 7 สถานะความรู้ที่ใช้ในการบวกลบเศษส่วนของนักเรียน จำแนกตามระดับความสามารถของนักเรียน

### ความเที่ยงของผลการประเมินเชิงวินิจฉัยแบบการคิดที่ผิด

การประเมินเชิงวินิจฉัยแบบการคิดที่ผิดพิจารณาจากนักเรียนกลุ่มที่ไม่มีความรู้ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนในแต่ละคุณลักษณะซึ่งจำแนกได้เป็น 3 กลุ่ม กลุ่มที่ 1 เป็นกลุ่มที่มีแบบการคิดที่ผิด พิจารณาจากมีผลการตอบข้อสอบที่ไม่ถูกต้อง โดยแสดงให้เห็นถึงการใช้แบบการคิดที่ผิดในการบวกลบเศษส่วนแบบใดแบบหนึ่งตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้ กลุ่มที่ 2 กลุ่มที่มีผลการประเมินวินิจฉัยไม่ได้ พิจารณาจากมีผลการตอบข้อสอบที่ไม่ถูกต้อง ซึ่งไม่ได้แสดงให้เห็นถึงการใช้แบบการคิดที่ผิดในการบวกลบเศษส่วนแบบใดแบบหนึ่งตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้ โดยส่วนใหญ่พบว่าเป็นผลการตอบข้อสอบที่ไม่ถูกต้องเนื่องจากการดูเครื่องหมายผิดหรือบวกลบจำนวนนับผิด และกลุ่มที่ 3 กลุ่มที่ไม่ได้รับการทดสอบในคุณลักษณะนั้น ซึ่งเป็นไปตามเงื่อนไขความสัมพันธ์เชิงลำดับขั้นของคุณลักษณะตามแผนภาพลำดับขั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ผลการประเมินเชิงวินิจฉัยของนักเรียนกลุ่มที่ไม่มีสถานะความรู้ที่ใช้ในการบวกลบเศษส่วน จำแนกตามคุณลักษณะ ดังแสดงในแผนภูมิที่ 8

คำร้อยละของผลการวินิจฉัย



แผนภูมิที่ 8 ผลการประเมินเชิงวินิจฉัยของนักเรียนกลุ่มที่ไม่มีสถานะความรู้ที่ใช้ในการบวกลบเศษส่วน จำแนกตามคุณลักษณะ ซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 1 และครั้งที่ 2

จากแผนภูมิที่ 8 สรุปจำนวนนักเรียนตามผลการวินิจฉัยในแต่ละคุณลักษณะได้ดังนี้  
คุณลักษณะที่ 1 การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน มีนักเรียนที่ไม่มี  
ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนสำหรับคุณลักษณะนี้ จากการประเมิน  
ครั้งที่ 1 และการประเมินครั้งที่ 2 จำนวน 154 คน และ 146 คน ตามลำดับ ผลการประเมินทั้ง 2 ครั้ง

สรุปได้ว่าเป็นนักเรียนที่มีแบบการคิดที่ผิดมากกว่ามีผลการประเมินที่วินิจัยไม่ได้ และไม่มีนักเรียนที่ไม่ได้รับการทดสอบในคุณลักษณะนี้เลย เนื่องจากเป็นคุณลักษณะที่ต้องมีมาก่อนทุกคุณลักษณะและอยู่ในลำดับขั้นแรกของแผนภาพลำดับขั้นของคุณลักษณะ

คุณลักษณะที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน มีนักเรียนที่ไม่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนสำหรับคุณลักษณะนี้ จากการประเมินครั้งที่ 1 และการประเมินครั้งที่ 2 จำนวน 420 คน และ 417 คน ตามลำดับ ผลการประเมินทั้ง 2 ครั้งสรุปได้ว่าเป็นนักเรียนที่มีแบบการคิดที่ผิดมากที่สุด รองลงมา คือ เป็นนักเรียนที่ไม่ได้รับการทดสอบในคุณลักษณะนี้และมีผลการประเมินที่วินิจัยไม่ได้ ตามลำดับ

คุณลักษณะที่ 3 การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน มีนักเรียนที่ไม่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนสำหรับคุณลักษณะนี้ จากการประเมินครั้งที่ 1 และการประเมินครั้งที่ 2 จำนวน 307 คน และ 311 คน ตามลำดับ ผลการประเมินทั้ง 2 ครั้งสรุปได้ว่าเป็นนักเรียนที่ไม่ได้ทดสอบในคุณลักษณะนี้มากที่สุด รองลงมา คือ มีผลการประเมินที่วินิจัยไม่ได้และมีแบบการคิดที่ผิด ตามลำดับ

คุณลักษณะที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน มีนักเรียนที่ไม่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนสำหรับคุณลักษณะนี้ จากการประเมินครั้งที่ 1 และการประเมินครั้งที่ 2 จำนวน 489 คน และ 491 คน ตามลำดับ ผลการประเมินทั้ง 2 ครั้งสรุปได้ว่าเป็นนักเรียนที่ไม่ได้รับการทดสอบในคุณลักษณะมากที่สุด รองลงมา คือ มีผลการประเมินที่วินิจัยไม่ได้และเป็นนักเรียนที่มีแบบการคิดที่ผิด ตามลำดับ

ผลการศึกษาความเที่ยงของผลการประเมินเชิงวินิจัยแบบการคิดที่ผิด ดังแสดงในตารางที่ 20

ตารางที่ 20 จำนวนนักเรียนจำแนกตามแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนซึ่งวินิจฉัยได้จากการทดสอบครั้งที่ 1 และการทดสอบครั้งที่ 2

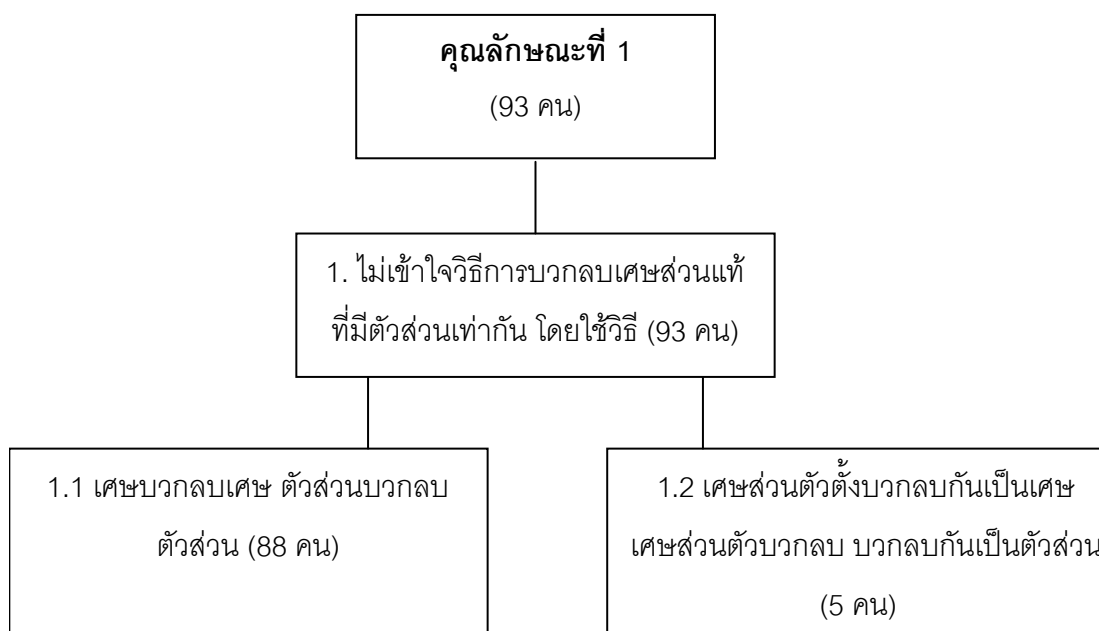
		จำนวนนักเรียนที่ใช้แบบการคิดที่ผิด จากการทดสอบครั้งที่ 2																				
		ไม่มีแบบ การคิดที่ ผิด	วินิจฉัย ไม่ได้	ไม่ได้ ทดสอบ	1.1.1	1.1.2	2.1.1	2.1.2	2.1.3	2.2.1	2.2.2	3.1.2	3.2.1	3.2.2	3.3.2	3.4.4	4.2.5	4.3.3	4.3.7	4.4.2	รวม	
จำนวนนักเรียนที่ใช้แบบการคิดที่ผิด จากการทดสอบครั้งที่ 1	ไม่มีแบบ การคิดที่ ผิด				4 (80.0%)	1 (20.0%)															5 (1.3%)	
	วินิจฉัย ไม่ได้									1 (7.1%)					6 (42.9%)		5 (35.7%)	2 (14.3%)			14 (3.5%)	
	ไม่ได้ ทดสอบ									6 (33.4%)		4 (2.2%)	4 (22.2%)		4 (22.2%)						18 (4.6%)	
	1.1.1				88 (94.6%)	5 (5.4%)															93 (23.5%)	
	1.1.2					5 (100%)															5 (1.3%)	
	2.1.1						25 (100%)														25 (6.3%)	
	2.1.2							10 (100%)													10 (2.5%)	
	2.1.3							10 (83.3%)	2 (16.7%)												12 (3.0%)	
	2.2.1									156 (100%)											156 (39.5%)	
	2.2.2										2 (100%)										2 (0.5%)	
	3.1.2											1 (100%)									1 (0.3%)	
	3.2.1												17 (100%)								17 (4.3%)	
	3.2.2													24 (100%)							24 (6.1%)	
	3.3.2														1 (100%)					1 (0.3%)		
	3.4.4	1 (50.0%)															1 (50.0%)				2 (0.5%)	
	4.2.5																	1 (100%)			1 (0.3%)	
	4.3.3																		2 (100%)		2 (0.5%)	
	4.3.7																			6 (100%)	6 (1.5%)	
	4.4.2																				2 (100%)	2 (0.5%)
	<b>รวม</b>	1 (0.3%)			92 (23.3%)	11 (2.8%)	35 (8.9%)	12 (3.0)		163 (42.3%)	2 (0.5%)	5 (1.3%)	21 (5.3%)	24 (6.1%)	11 (2.8%)	1 (0.3%)	6 (1.5%)	4 (1.0%)	6 (1.5%)	2 (0.5%)	395 (100%)	





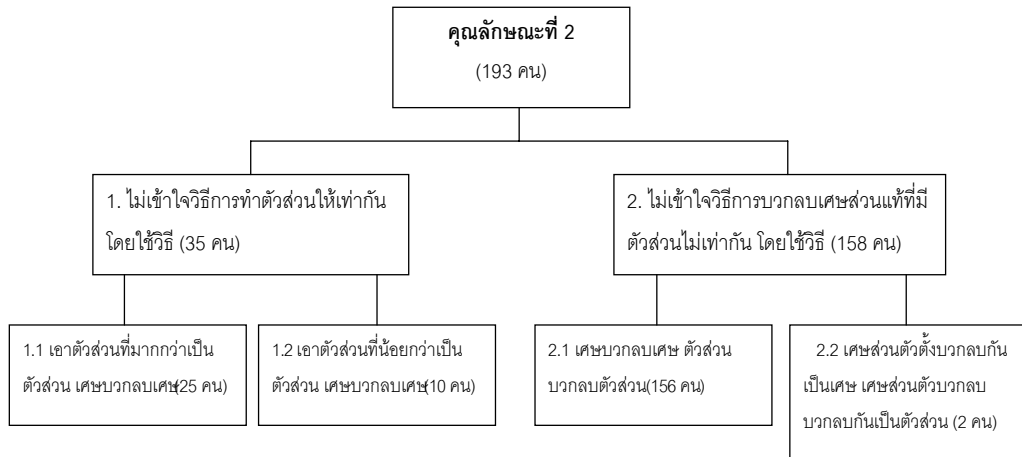
จากตารางที่ 20 แสดงให้เห็นว่านักเรียนใช้แบบการคิดที่ผิดในการบวกเศษส่วน จากการทดสอบครั้งที่ 1 และ การทดสอบครั้งที่ 2 สอดคล้องกัน จำนวน 341 คน จาก นักเรียนที่มีผลการวินิจฉัยว่ามีแบบการคิดที่ผิดทั้งหมด จำนวน 395 คน คิดเป็นสัดส่วน เท่ากับ 0.86 แสดงให้เห็นว่าวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นสามารถประเมินเชิงวินิจฉัยแบบการคิดที่ผิดได้อย่างมีความเที่ยง สรุปลักษณะแบบการคิดที่ผิดที่ใช้ในการบวกเศษส่วนสำหรับแต่ละคุณลักษณะ ได้ดังนี้

คุณลักษณะที่ 1 การบวกเศษส่วนแท้ที่มีตัวส่วนเท่ากัน พบว่า มีนักเรียนที่ใช้แบบการคิดที่ผิด จำนวน 93 คน เป็นนักเรียนที่ไม่เข้าใจวิธีการบวกเศษส่วนแท้ที่มีตัวส่วนเท่ากันทั้งหมด โดยส่วนใหญ่ใช้วิธีเศษบวกเศษ ตัวส่วนบวกตัวส่วนมากกว่าใช้วิธีเศษส่วนตัวตั้งบวกกันเป็นเศษ เศษส่วนตัวบวกบวกกันเป็นตัวส่วน ดังแสดงในแผนภาพที่ 15



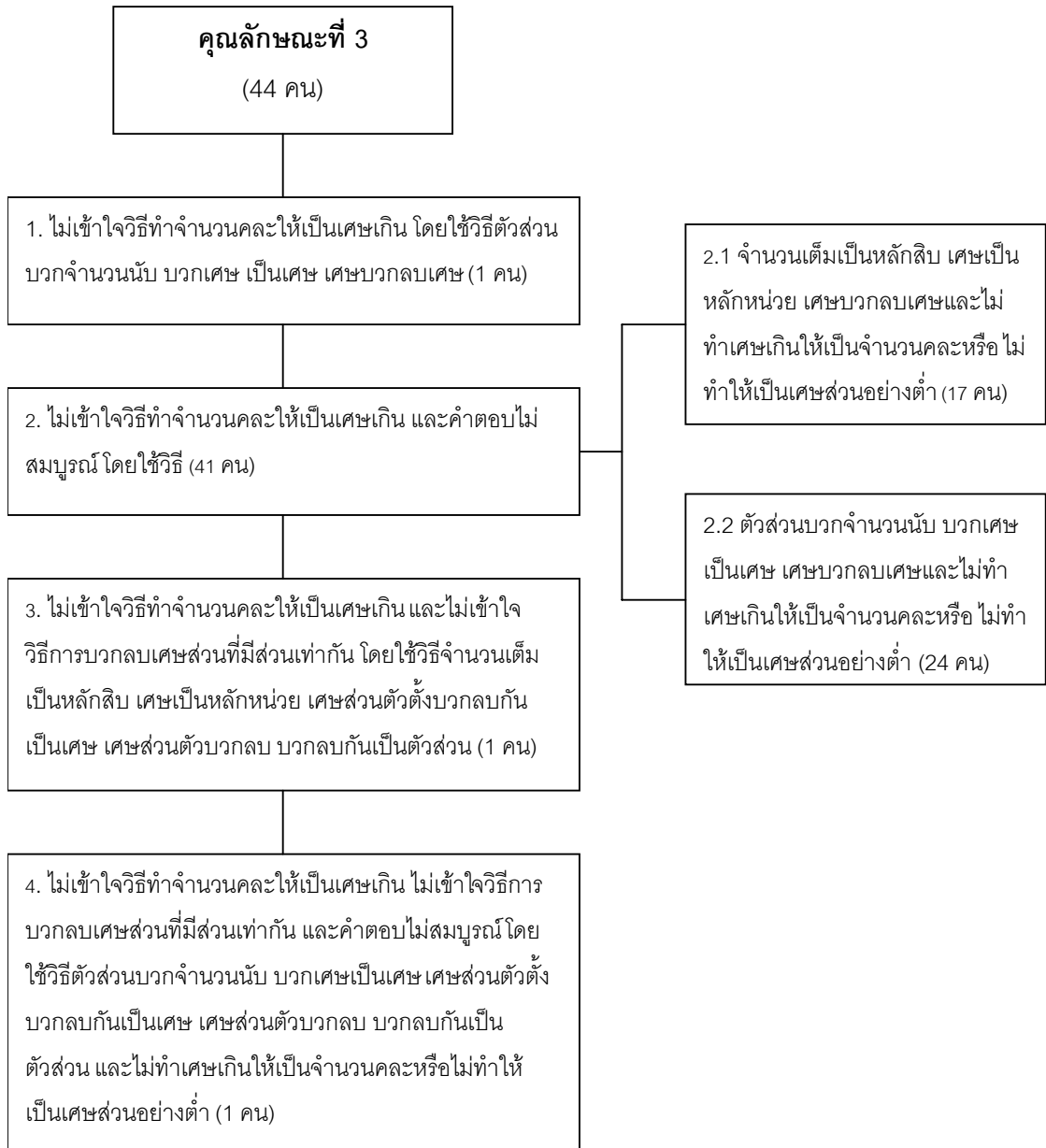
แผนภาพที่ 15 จำนวนนักเรียนจำแนกตามแบบการคิดที่ผิดที่ใช้ในการบวกเศษส่วน สำหรับคุณลักษณะที่ 1

คุณลักษณะที่ 2 การบวกเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน พบว่า มีนักเรียนที่ใช้แบบการคิดที่ผิด จำนวน 193 คน ส่วนใหญ่เป็นนักเรียนที่ไม่เข้าใจวิธีการบวกเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน โดยใช้วิธีเศษบวกเศษ ตัวส่วนบวกตัวส่วนมากกว่าใช้วิธีเศษส่วนตัวตั้งบวกกันเป็นเศษ เศษส่วนตัวบวกบวกกันเป็นตัวส่วน รองลงมา คือ นักเรียนไม่เข้าใจวิธีการทำตัวส่วนให้เท่ากัน โดยใช้วิธีเอาตัวส่วนที่มากกว่าเป็นตัวส่วน เศษบวกเศษมากกว่าใช้วิธีเอาตัวส่วนที่น้อยกว่าเป็นส่วน เศษบวกเศษ ดังแสดงในแผนภาพที่ 16



แผนภาพที่ 16 จำนวนนักเรียนจำแนกตามแบบการคิดที่ผิดที่ใช้ในการบวกเศษส่วน สำหรับ  
คุณลักษณะที่ 2

คุณลักษณะที่ 3 การบวกจำนวนคละที่มีตัวส่วนเท่ากัน พบว่า มีนักเรียนที่มีแบบการคิดที่ผิด จำนวน 44 คน ส่วนใหญ่เป็นนักเรียนที่ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกินและคำตอบไม่สมบูรณ์ โดยใช้วิธี ตัวส่วนบวกจำนวนนับบวกเศษเป็นเศษ เศษบวกเศษ และไม่ทำเศษเกินให้เป็นจำนวนคละหรือไม่ทำให้เป็นเศษส่วนอย่างต่ำมากกว่าใช้วิธีจำนวนเต็มเป็นหลักสิบเศษเป็นหลักหน่วย เศษบวกเศษและไม่ทำเศษเกินให้เป็นจำนวนคละหรือไม่ทำให้เป็นเศษส่วนอย่างต่ำ รองลงมา ได้แก่ ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน โดยใช้วิธีตัวส่วนบวกจำนวนนับบวกเศษเป็นเศษ เศษบวกเศษ ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกินและไม่เข้าใจวิธีบวกเศษส่วนที่มีตัวส่วนเท่ากัน โดยใช้วิธีจำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เศษส่วนตัวตั้งบวกกลับกันเป็นเศษ เศษส่วนตัวบวกกลับบวกกลับกันเป็นตัวส่วน และไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน ไม่เข้าใจวิธีการบวกเศษส่วนที่มีส่วนเท่ากันและคำตอบไม่สมบูรณ์ โดยใช้วิธีตัวส่วนบวกจำนวนนับบวกเศษเป็นเศษ เศษส่วนตัวตั้งบวกกลับกันเป็นเศษ เศษส่วนตัวบวกกลับบวกกลับกันเป็นตัวส่วนและไม่ทำเศษเกินให้เป็นจำนวนคละหรือไม่ทำให้เป็นเศษส่วนอย่างต่ำ ดังแสดงในแผนภาพที่ 17



แผนภาพที่ 17 จำนวนนักเรียนจำแนกตามแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน สำหรับ  
คุณลักษณะที่ 3

คุณลักษณะที่ 4 การบวกลบจำนวนคณะที่มีตัวส่วนไม่เท่ากัน พบว่า มีนักเรียนที่มี  
แบบการคิดที่ผิด จำนวน 11 คน ส่วนใหญ่เป็นนักเรียนที่ไม่เข้าใจวิธีทำจำนวนคณะให้เป็นเศษเกิน  
ไม่เข้าใจวิธีการทำตัวส่วนให้เท่ากันและคำตอบไม่สมบูรณ์ โดยใช้วิธีตัวส่วนบวกจำนวนนับบวก  
เศษเป็นเศษ เอาตัวส่วนคูณกันเป็นตัวส่วน เศษบวกลบเศษ หรือหา ค.ร.น เป็นตัวส่วน เศษบวกลบเศษ  
และไม่ทำเศษเกินให้เป็นจำนวนคณะหรือไม่ทำให้เป็นเศษส่วนอย่างต่ำมากกว่าใช้วิธีจำนวนเต็ม



เมื่อพิจารณาแบบการคิดที่ผิดจำแนกตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้ พบว่า  
คุณลักษณะที่ 1 การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน นักเรียนส่วนใหญ่ไม่เข้าใจวิธีการบวกลบ  
เศษส่วนแท้ที่มีตัวส่วนเท่ากัน โดยนักเรียนที่ใช้เกณฑ์ในการตัดสินความรอบรู้เกณฑ์ที่ 1 เกณฑ์ที่ 4  
และเกณฑ์ที่ 6 ใช้วิธีเศษบวกลบเศษตัวส่วนบวกลบตัวส่วนมากกว่าใช้วิธีเศษส่วนตัวตั้งบวกลบกัน  
เป็นเศษ เศษส่วนตัวบวกลบบวกลบกันเป็นตัวส่วน ส่วนนักเรียนที่ใช้เกณฑ์ในการตัดสินความรอบรู้  
เกณฑ์ที่ 2 เกณฑ์ที่ 3 และเกณฑ์ที่ 5 พบว่าทั้งหมดใช้วิธีเศษบวกลบเศษตัวส่วนบวกลบตัวส่วน

คุณลักษณะที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากันพบว่านักเรียนที่เกณฑ์  
ในการตัดสินความรอบรู้ทั้ง 6 เกณฑ์ ส่วนใหญ่ไม่เข้าใจวิธีการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน  
มากกว่าไม่เข้าใจวิธีการทำตัวส่วนให้เท่ากัน โดยใช้วิธีเศษบวกลบเศษตัวส่วนบวกลบตัวส่วนมากที่สุด

คุณลักษณะที่ 3 การบวกลบจำนวนคละที่มีตัวส่วนเท่ากันพบว่านักเรียนที่ใช้  
เกณฑ์ในการตัดสินความรอบรู้เกณฑ์ที่ 1 เกณฑ์ที่ 2 เกณฑ์ที่ 3 เกณฑ์ที่ 4 และเกณฑ์ที่ 5 ส่วนใหญ่  
ไม่เข้าใจวิธีการทำจำนวนคละให้เป็นเศษเกินและคำตอบไม่สมบูรณ์ ส่วนนักเรียนที่ใช้เกณฑ์ใน  
การตัดสินความรอบรู้เกณฑ์ที่ 6 มีแบบการคิดที่ผิดจำนวน 1 คน เป็นนักเรียนที่ไม่เข้าใจวิธีการทำ  
จำนวนคละให้เป็นเศษเกินและไม่เข้าใจวิธีการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน โดยใช้วิธี  
จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เศษส่วนตัวตั้งบวกลบกันเป็นเศษ เศษส่วนตัวบวกลบ  
บวกลบกันเป็นตัวส่วน

คุณลักษณะที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากันพบว่านักเรียนที่ใช้  
เกณฑ์ในการตัดสินความรอบรู้ทั้ง 6 เกณฑ์ ส่วนใหญ่ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน ไม่  
เข้าใจวิธีการทำตัวส่วนให้เท่ากันและคำตอบไม่สมบูรณ์

ดังแสดงในตารางที่ 21

ตารางที่ 21 จำนวนนักเรียนที่มีแบบการคิดที่ผิด จำแนกตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้

โปรแกรม ที่	จำนวนนักเรียนที่มีแบบการคิดที่ผิด														
	คุณลักษณะที่ 1		คุณลักษณะที่ 2				คุณลักษณะที่ 3				คุณลักษณะที่ 4				
	1.1.1	1.1.2	2.1.1	2.1.2	2.2.1	2.2.2	3.1.2	3.2.1	3.2.2	3.3.2	3.4.4	4.2.5	4.3.3	4.3.7	4.4.2
1	15	2	8	1	25		1	5	9			1		1	2
2	14		1		32			3	8				1	2	
3	15		6	4	33			3	6					1	
4	16	2		3	24	1		5					1		
5	11		6	2	26			1	1		1			1	
6	17	1	4		16	1				1				1	
<b>รวม</b>	<b>93</b>				<b>193</b>			<b>44</b>					<b>11</b>		

เมื่อพิจารณาแบบการคิดที่ผิดจำแนกตามระดับความสามารถ พบว่านักเรียนที่มีความสามารถระดับต่ำและระดับปานกลางมีแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนทุกคุณลักษณะ โดยส่วนใหญ่ไม่เข้าใจวิธีการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากันโดยใช้วิธีเศษบวกลบเศษ ตัวส่วนบวกลบตัวส่วน ส่วนนักเรียนที่มีระดับความสามารถสูงมีแบบการคิดที่ผิดจำนวน 1 คน เป็นนักเรียนที่ไม่เข้าใจวิธีการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากันโดยใช้วิธีเศษบวกลบเศษ ตัวส่วนบวกลบตัวส่วนเช่นกัน

ดังแสดงในตารางที่ 22

ตารางที่ 22 จำนวนนักเรียนที่มีแบบการคิดที่ผิด จำแนกตามระดับความสามารถของนักเรียน

ระดับ ความสามารถ	จำนวนนักเรียนที่มีแบบการคิดที่ผิด														
	คุณลักษณะที่ 1		คุณลักษณะที่ 2				คุณลักษณะที่ 3				คุณลักษณะที่ 4				
	1.1.1	1.1.2	2.1.1	2.1.2	2.2.1	2.2.2	3.1.2	3.2.1	3.2.2	3.3.2	4.2.5	4.3.3	4.3.7	4.4.2	
<b>ต่ำ</b>	83	5	21	6	106	1	1	16	23	1	1	1	3	2	
<b>ปานกลาง</b>	5		4	4	49	1		1	2			1	3		
<b>สูง</b>					1										
<b>รวม</b>	<b>93</b>				<b>193</b>			<b>44</b>				<b>11</b>			

## 2. ผลการศึกษาความตรงของผลการประเมินเชิงวินิจฉัยจากวิธีการที่พัฒนาขึ้น

ความตรงของผลการประเมินเชิงวินิจฉัยจากวิธีการที่พัฒนาขึ้น หมายถึง คุณสมบัติของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นในด้านความสามารถในการประเมินเชิงวินิจฉัยได้ตรงตามสภาพของผู้สอบ (Concurrent validity) พิจารณาจากความสอดคล้องระหว่างผลการประเมินเชิงวินิจฉัยจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง กับผลการประเมินเชิงวินิจฉัยโดยใช้เทคนิคการคิดออกเสียง จำแนกการศึกษาความตรงออกเป็น 2 ด้านตามผลการประเมินเชิงวินิจฉัย ได้แก่ ความตรงของผลการประเมินเชิงวินิจฉัยสถานะความรู้และความตรงของผลการประเมินเชิงวินิจฉัยแบบการคิดที่ผิด ปรากฏผลการศึกษาความตรงในแต่ละด้าน ดังนี้

### 2.1 ความตรงของผลการประเมินเชิงวินิจฉัยสถานะความรู้

การศึกษาผลการประเมินเชิงวินิจฉัยสถานะความรู้ พบว่า นักเรียนมีแบบแผนการตอบข้อสอบที่คาดหวังจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 1 และครั้งที่ 2 สอดคล้องกับแบบแผนการตอบข้อสอบที่คาดหวังจากการใช้เทคนิคการคิดออกเสียง จำนวน 45 คน และ 44 คน จากนักเรียนทั้งหมด จำนวน 54 คน คิดเป็นสัดส่วน เท่ากับ 0.83 และ 0.81 ตามลำดับ ดังแสดงในตารางที่ 23 และตารางที่ 24

ตารางที่ 23 จำนวนนักเรียนจำแนกตามแบบแผนการตอบข้อสอบที่คาดหวังซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 1 และจากการใช้เทคนิคการคิดออกเสียง

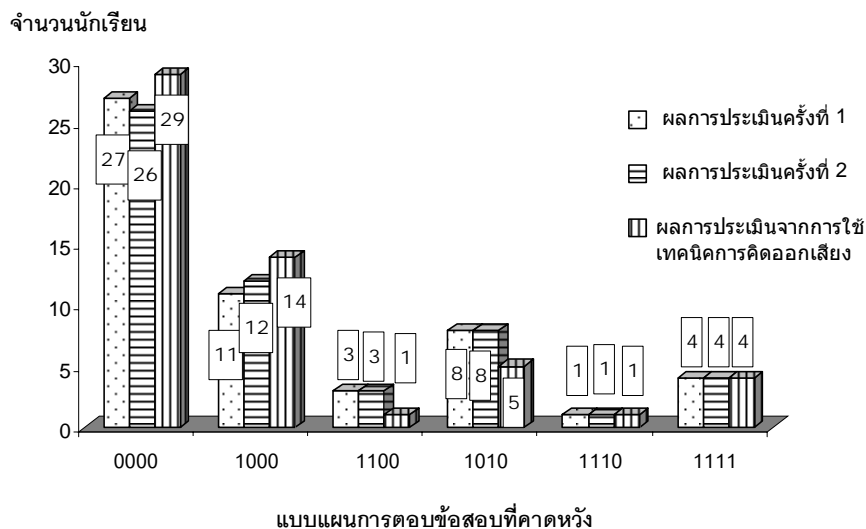
		ผลการประเมินเชิงวินิจฉัยโดยใช้เทคนิคการคิดออกเสียง						
		0000	1000	1100	1010	1110	1111	รวม
ผลการประเมินเชิงวินิจฉัยจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 1	0000	26 (89.7%)	3 (10.3%)					29 (53.7%)
	1000	1 (7.1%)	8 (57.1%)	2 (14.3%)	3 (21.4%)			14 (25.9%)
	1100			1 (100%)				1 (1.9%)
	1010				5 (100%)			5 (9.2%)
	1110					1 (100%)		1 (1.9%)
	1111						4 (100%)	4 (7.4%)
	<b>รวม</b>	<b>27</b> (50.0%)	<b>11</b> (20.4%)	<b>3</b> (5.5%)	<b>8</b> (14.8%)	<b>1</b> (1.9%)	<b>4</b> (7.4%)	<b>54</b> (100%)

ตารางที่ 24 จำนวนนักเรียนจำแนกตามแบบแผนการตอบข้อสอบที่คาดหวังซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 2 และจากการใช้เทคนิคการคิดออกเสียง

ผลการประเมินเชิงวินิจฉัยโดยใช้เทคนิคการคิดออกเสียง

		0000	1000	1100	1010	1110	1111	รวม
ผลการประเมินเชิงวินิจฉัยจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 2	0000	25 (86.2%)	4 (13.8%)					29 (53.7%)
	1000	1 (7.1%)	8 (57.2%)	2 (14.3%)	3 (21.4%)			14 (25.9%)
	1100			1 (100%)				1 (1.9%)
	1010				5 (100%)			5 (9.2%)
	1110					1 (100%)		1 (1.9%)
	1111						4 (100%)	4 (7.4%)
	รวม	26 (48.1%)	12 (22.2%)	3 (5.6%)	8 (14.8%)	1 (1.9%)	4 (7.4%)	54 (100%)

เมื่อจำแนกจำนวนนักเรียนในแต่ละแบบแผนการตอบข้อสอบที่คาดหวังปรากฏผลดังแสดงในแผนภูมิที่ 9

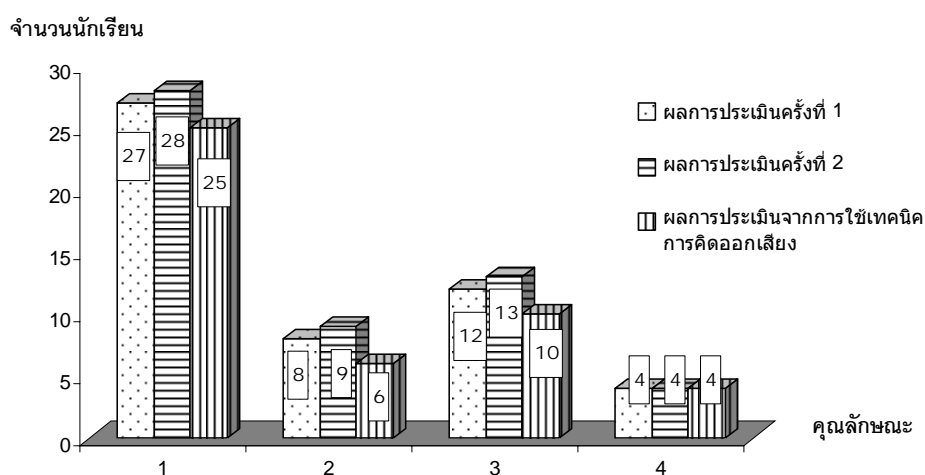


แผนภูมิที่ 9 สถานะความรู้ที่ใช้ในการบวกลบเศษส่วนของนักเรียน จำแนกตามแบบแผนการตอบข้อสอบที่คาดหวัง ซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง และจากการใช้เทคนิคการคิดออกเสียง



จากแผนภูมิที่ 9 แสดงให้เห็นว่านักเรียนมีแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 1 (0000) มากที่สุด รองลงมา ได้แก่ แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 2 (1000) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 4 (1010) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 6 (1111) แบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 3 (1100) และแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 5 (1110) แสดงให้เห็นว่านักเรียนส่วนใหญ่เป็นนักเรียนกลุ่มที่ไม่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนทุกเรื่อง จำนวน 27 คน คิดเป็นร้อยละ 50.0 จำนวน 26 คน คิดเป็นร้อยละ 48.1 และจำนวน 29 คน คิดเป็นร้อยละ 53.7 จากการประเมินครั้งที่ 1 การประเมินครั้งที่ 2 และการประเมินโดยใช้เทคนิคการคิดออกเสียงตามลำดับ รองลงมาได้แก่ เป็นนักเรียนที่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนเพียงบางส่วน จำนวน 23 คน คิดเป็นร้อยละ 52.6 จำนวน 24 คน คิดเป็นร้อยละ 44.4 และจำนวน 21 คน คิดเป็นร้อยละ 38.9 จากการประเมินครั้งที่ 1 การประเมินครั้งที่ 2 และการประเมินโดยใช้เทคนิคการคิดออกเสียง ตามลำดับโดยเป็นนักเรียนกลุ่มที่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนที่มีตัวส่วนเท่ากัน ทั้งเศษส่วนแท้และจำนวนคละมากกว่าการบวกลบเศษส่วนที่มีตัวส่วนไม่เท่ากัน ส่วนนักเรียนกลุ่มที่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนทุกเรื่องเป็นกลุ่มนักเรียนที่มีจำนวนน้อยที่สุด คือ จำนวน 4 คน คิดเป็นร้อยละ 7.4 เท่ากันทั้งจากการประเมินครั้งที่ 1 การประเมินครั้งที่ 2 และการประเมินโดยใช้เทคนิคการคิดออกเสียง

เมื่อจำแนกจำนวนนักเรียนตามสถานะความรู้ในแต่ละคุณลักษณะ ปรากฏผลดังแสดงในแผนภูมิที่ 10

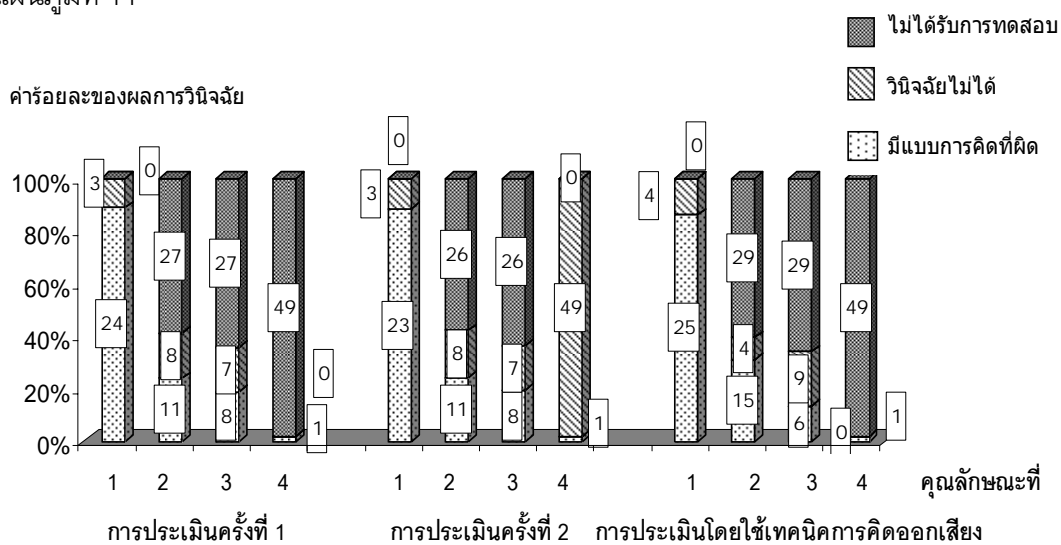


แผนภูมิที่ 10 สถานะความรู้ที่ใช้ในการบวกลบเศษส่วนของนักเรียน จำแนกตามคุณลักษณะซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้งและจากการใช้เทคนิคการคิดออกเสียง

จากแผนภูมิที่ 10 แสดงให้เห็นว่านักเรียนส่วนใหญ่มีความรู้ ทักษะหรือคุณลักษณะที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน รองลงมา ได้แก่ มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากันและมีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน ดังแสดงในแผนภูมิที่ 10

### 2.2 การศึกษาความตรงของผลการประเมินเชิงวินิจฉัยัยแบบการคิดที่ผิด

การศึกษาความตรงของผลการประเมินเชิงวินิจฉัยัยแบบการคิดที่ผิดพิจารณาจากนักเรียนกลุ่มที่มีผลการวินิจฉัยัยว่าไม่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนในแต่ละคุณลักษณะซึ่งมีผลการวินิจฉัยัยว่ามีแบบการคิดที่ผิด เนื่องจากสามารถจำแนกแบบการคิดที่ผิดตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้ได้ จำนวนนักเรียนกลุ่มที่ไม่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนจำแนกตามผลการวินิจฉัยัยในแต่ละคุณลักษณะ ดังแสดงในแผนภูมิที่ 11



แผนภูมิที่ 11 ผลการประเมินเชิงวินิจฉัยัยของนักเรียนกลุ่มที่ไม่มีสถานะความรู้ที่ใช้ในการบวกลบเศษส่วน จำแนกตามคุณลักษณะซึ่งวินิจฉัยัยได้จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง และจากการใช้เทคนิคการคิดออกเสียง

จากแผนภูมิที่ 11 แสดงให้เห็นว่ามีนักเรียนที่มีผลการวินิจฉัยัยว่ามีแบบการคิดที่ผิดจากการประเมินเชิงวินิจฉัยัยโดยใช้การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 1 ครั้งที่ 2 และการประเมินโดยใช้เทคนิคการคิดออกเสียง จำนวน 44 คน 43 คน และ 47 คน ตามลำดับ

เมื่อพิจารณาความสอดคล้องระหว่างผลการประเมิน พบว่า นักเรียนมีแบบแผนการตอบข้อสอบที่คาดหวังจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 1 และครั้งที่ 2 สอดคล้องกับแบบแผนการตอบข้อสอบที่คาดหวังจากการใช้เทคนิคการคิดออกเสียงจำนวน 38 คน และ 37 คน จากนักเรียนทั้งหมด จำนวน 44 คน คิดเป็นสัดส่วน เท่ากับ 0.86 และ 0.84 ตามลำดับ ดังแสดงในตารางที่ 25 และตารางที่ 26 แสดงให้เห็นว่าวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นสามารถประเมินเชิงวินิจฉัยแบบการคิดที่ผิดได้อย่างมีความตรง สรุปลักษณะแบบการคิดที่ผิดสำหรับแต่ละคุณลักษณะได้ดังนี้

คุณลักษณะที่ 1 การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน พบว่า มีนักเรียนที่ใช้แบบการคิดที่ผิด จำนวน 22 คน เป็นนักเรียนที่ไม่เข้าใจวิธีการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากันทั้งหมด โดยส่วนใหญ่ใช้วิธีเศษบวกลบเศษ ตัวส่วนบวกลบตัวส่วน (21 คน) มากกว่าใช้วิธีเศษส่วนตัวตั้งบวกลบกันเป็นเศษ เศษส่วนตัวบวกลบบวกลบกันเป็นตัวส่วน (1 คน)

คุณลักษณะที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน พบว่า มีนักเรียนที่ใช้แบบการคิดที่ผิด จำนวน 11 คน เป็นนักเรียนที่ไม่เข้าใจวิธีการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน โดยใช้วิธีเศษบวกลบเศษ ตัวส่วนบวกลบตัวส่วน (10 คน) มากกว่าเป็นนักเรียนที่ไม่เข้าใจวิธีการทำตัวส่วนให้เท่ากัน โดยใช้วิธีเอาตัวส่วนที่มากกว่าเป็นตัวส่วน เศษบวกลบเศษ (1 คน)

คุณลักษณะที่ 3 การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน พบว่า มีนักเรียนที่ใช้แบบการคิดที่ผิด จำนวน 4 คน โดยทั้งหมดเป็นนักเรียนที่ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน และคำตอบไม่สมบูรณ์ โดยใช้วิธีตัวส่วนบวกจำนวนนับบวกเศษเป็นเศษ เศษบวกลบเศษและไม่ทำเศษเกินให้เป็นจำนวนคละหรือไม่ทำให้เป็นเศษส่วนอย่างต่ำ

คุณลักษณะที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน พบว่า มีนักเรียนที่ใช้แบบการคิดที่ผิด จำนวน 1 คน เป็นนักเรียนที่ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน ไม่เข้าใจวิธีการทำตัวส่วนให้เท่ากันและคำตอบไม่สมบูรณ์ โดยใช้วิธีตัวส่วนบวกจำนวนนับบวกเศษเป็นเศษ เอาตัวส่วนคูณกันเป็นตัวส่วนเศษบวกลบเศษหรือหา ค.ร.น.เป็นตัวส่วน เศษบวกลบเศษ ไม่ทำเศษเกินให้เป็นจำนวนคละหรือไม่ทำให้เป็นเศษส่วนอย่างต่ำ

ตารางที่ 25 จำนวนนักเรียนจำแนกตามแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 1 และจากการใช้เทคนิคการคิดออกเสียง

		จำนวนนักเรียนที่ใช้แบบการคิดที่ผิดจากการประเมินเชิงวินิจฉัยโดยใช้เทคนิคการคิดออกเสียง										
		1.1.1	1.1.2	2.1.1	2.2.1	3.1.2	3.2.1	3.2.2	3.3.2	3.4.4	4.3.7	รวม
จำนวนนักเรียนที่ใช้แบบการคิดที่ผิดจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ครั้งที่ 1	ไม่มีแบบการคิดที่ผิด		1									1
			(100%)									(2.3%)
	ไม่ได้ทดสอบ					2	1					3
						(66.7%)	(33.3%)					(6.8%)
	1.1.1	21	1									22
		(95.5%)	(4.5%)									(50.0%)
	1.1.2		1									1
			(100%)									(2.3%)
	2.1.1			1								1
				(100%)								(2.3%)
	2.2.1				10							10
					(100%)							(22.7%)
3.2.2								4			4	
								(100%)			(9.1%)	
3.4.4									1		1	
									(100%)		(2.3%)	
4.3.7										1	1	
										(100%)	(2.3%)	
รวม	21	3	1	10	2	1	4	1		1	44	
	(47.7%)	(6.8%)	(2.3%)	(22.7%)	(4.5%)	(2.3%)	(9.1%)	(2.3%)		(2.3%)	(100%)	

ตารางที่ 26 จำนวนนักเรียนจำแนกตามแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนซึ่งวินิจฉัยได้จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 2 และจากการใช้เทคนิคการคิดออกเสียง

		จำนวนนักเรียนที่ใช้แบบการคิดที่ผิดจากการประเมินเชิงวินิจฉัยโดยใช้เทคนิคการคิดออกเสียง										
		1.1.1	1.1.2	2.1.1	2.2.1	3.1.2	3.2.1	3.2.2	3.3.2	3.4.4	4.3.7	รวม
จำนวนนักเรียนที่ใช้แบบการคิดที่ผิดจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ครั้งที่ 2	ไม่มีแบบการคิดที่ผิด		1									1
			(100%)									(2.3%)
	วินิจฉัยไม่ได้		1									1
			(100%)									(2.3%)
	ไม่ได้ทดสอบ					2	1					3
						(66.7%)	(33.3%)					(6.8%)
	1.1.1	21	1									22
		(95.5%)	(4.5%)									(50.0%)
	2.1.1			1								1
				(100%)								(2.3%)
	2.2.1				10							10
					(100%)							(22.7%)
3.2.2								4			4	
								(100%)			(9.1%)	
3.4.4									1		1	
									(100%)		(2.3%)	
4.3.7										1	1	
										(100%)	(2.3%)	
รวม	21	3	1	10	2	1	4	1		1	44	
	(47.7%)	(6.8%)	(2.3%)	(22.7%)	(4.5%)	(2.3%)	(9.1%)	(2.3%)		(2.3%)	(100%)	

### 3. ผลการศึกษาจำนวนเวลาที่ใช้ในการทดสอบ

ผลการศึกษาจำนวนเวลาที่ใช้ในการทดสอบ ดังแสดงในตารางที่ 27

ตารางที่ 27 จำนวนเวลาที่ใช้ในการทดสอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์  
จำแนกตามครั้งที่ทดสอบ

เกณฑ์ที่	จำนวน	การทดสอบครั้งที่ 1				การทดสอบครั้งที่ 2				รวม	
		$\bar{x}$	S.D.	ต่ำสุด	สูงสุด	$\bar{x}$	S.D.	ต่ำสุด	สูงสุด	$\bar{x}$	S.D.
1	90	14.5	12.3	1	58	14.1	12.0	1	53	14.3	12.1
2	73	17.3	11.3	1	54	17.0	11.1	1	49	17.2	11.2
3	83	20.5	13.9	1	54	20.1	13.7	1	49	20.3	13.8
4	90	19.1	13.8	1	50	18.5	13.4	1	47	18.8	13.6
5	90	15.8	11.1	1	43	15.4	10.8	1	41	15.6	10.9
6	90	16.4	8.4	1	38	15.7	8.0	1	34	16.1	8.1
รวม	516	17.2	12.1	1	58	16.8	11.8	1	53	17.0	11.9

จากการศึกษาจำนวนเวลาที่ใช้ในการทดสอบจำนวน 2 ครั้ง พบว่า นักเรียนใช้เวลาในการทดสอบเฉลี่ย เท่ากับ 17.0 นาที โดยใช้เวลาเฉลี่ยในการทดสอบครั้งที่ 1 และการทดสอบครั้งที่ 2 เท่ากับ 17.2 นาที และ 16.8 นาที ตามลำดับ สำหรับเวลาต่ำสุดที่ใช้ในการทดสอบ พบว่า การทดสอบทั้ง 2 ใช้เวลาต่ำสุด เท่ากับ 1 นาที ส่วนเวลาสูงสุดที่ใช้ในการทดสอบ พบว่า ครั้งที่ 1 ใช้เวลาในการทดสอบสูงสุด 58 นาที และครั้งที่ 2 ใช้เวลาในการทดสอบสูงสุด 53 นาที เมื่อพิจารณาค่าเฉลี่ยของเวลาที่ใช้ในการทดสอบจำแนกตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้ พบว่า เกณฑ์ที่ใช้ในการตัดสินความรอบรู้ที่ใช้เวลาเฉลี่ยในการทดสอบทั้ง 2 ครั้ง ต่ำสุด คือ เกณฑ์ที่ 3 รองลงมา ได้แก่ เกณฑ์ที่ 4 เกณฑ์ที่ 2 เกณฑ์ที่ 6 เกณฑ์ที่ 5 และเกณฑ์ที่ 1 ตามลำดับ ดังแสดงในตารางที่ 27

### 4. ผลการศึกษาจำนวนข้อสอบที่ใช้ในการทดสอบ

จากการศึกษาจำนวนข้อสอบที่ใช้ในการทดสอบจำนวน 2 ครั้ง พบว่า นักเรียนใช้ข้อสอบในการทดสอบเฉลี่ย จำนวน 40.7 ข้อ โดยใช้ข้อสอบเฉลี่ยในการทดสอบครั้งที่ 1 และการทดสอบครั้งที่ 2 จำนวน 41.0 ข้อ และ 40.3 ข้อ ตามลำดับ สำหรับจำนวนข้อสอบต่ำสุดที่ใช้ในการทดสอบ พบว่าการทดสอบทั้ง 2 ใช้ข้อสอบต่ำสุด จำนวน 3 ข้อ ส่วนจำนวนข้อสอบสูงสุดที่ใช้ในการทดสอบ พบว่าครั้งที่ 1 ใช้ข้อสอบสูงสุด จำนวน 160 ข้อ และครั้งที่ 2 ใช้ข้อสอบสูงสุดจำนวน 154 ข้อ เมื่อพิจารณาค่าเฉลี่ย

ของจำนวนข้อสอบที่ใช้ในการทดสอบจำแนกตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้ พบว่า เกณฑ์ในการตัดสินความรอบรู้ที่ใช้จำนวนข้อสอบเฉลี่ยในการทดสอบทั้ง 2 ครั้ง สูงสุด คือ เกณฑ์ที่ 5 รองลงมา ได้แก่ เกณฑ์ที่ 4 เกณฑ์ที่ 6 เกณฑ์ที่ 2 เกณฑ์ที่ 3 และเกณฑ์ที่ 1 ตามลำดับ ดังแสดงในตารางที่ 28 ตารางที่ 28 จำนวนข้อสอบที่ใช้ในการทดสอบจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำแนกตามครั้งที่ทดสอบ

เกณฑ์ที่	จำนวน	การทดสอบครั้งที่ 1				การทดสอบครั้งที่ 2				รวม	
		$\bar{x}$	S.D.	ต่ำสุด	สูงสุด	$\bar{x}$	S.D.	ต่ำสุด	สูงสุด	$\bar{x}$	S.D.
1	90	33.6	36.4	3	121	35.6	36.0	3	132	34.6	34.3
2	73	40.2	31.0	4	113	38.2	27.2	4	115	39.2	29.0
3	83	40.0	31.4	5	160	38.4	29.0	5	154	39.2	30.1
4	90	44.7	28.0	18	125	44.2	27.0	15	109	44.5	27.5
5	90	46.8	31.9	12	121	44.4	27.8	12	110	45.6	30.0
6	90	40.7	21.9	18	119	40.6	22.0	15	119	40.6	22.0
รวม	516	41.0	30.5	3	160	40.3	28.5	3	154	40.7	29.1

### ตอนที่ 3 ผลการเปรียบเทียบคุณภาพของวิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้น

คุณภาพของวิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้น พิจารณาจากความเที่ยง ความตรง จำนวนเวลาและจำนวนข้อสอบที่ใช้ในการทดสอบ การเปรียบเทียบคุณภาพของวิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้นทั้ง 4 ด้าน ดำเนินการวิเคราะห์ข้อมูลระหว่างโปรแกรมย่อยที่ใช้ในการทดสอบแบบปรับเหมาะ จำนวน 6 โปรแกรม ซึ่งมีเกณฑ์ที่ใช้ในการตัดสินความรอบรู้ในแต่ละลำดับชั้นที่แตกต่างกัน ปราบกฎผลการเปรียบเทียบคุณภาพแต่ละด้าน ดังนี้

#### 1. ผลการเปรียบเทียบความเที่ยง

จากการเปรียบเทียบความเที่ยงของผลการประเมินเชิงวินิจจัยสถานะความรู้และแบบการคิดที่ผิดระหว่างเกณฑ์ที่ใช้ในการตัดสินความรอบรู้จำนวน 6 เกณฑ์ พบว่าทั้ง 6 เกณฑ์มีความเที่ยงของผลการประเมินเชิงวินิจจัยสถานะความรู้และแบบการคิดที่ผิดไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยมีค่าไคสแควร์ เท่ากับ 5.473 ที่ชั้นแห่งความเป็นอิสระ เท่ากับ 5

สำหรับผลการประเมินเชิงวิจัญัยสถานะความรู้ และมีค่าไคสแควร์ เท่ากับ 4.593 ที่ชั้นแห่งความ เป็นอิสระ เท่ากับ 5 สำหรับผลการประเมินเชิงวิจัญัยแบบการคิดที่ผิด ดังแสดงในตารางที่ 29 ตารางที่ 29 ผลการเปรียบเทียบความเที่ยงของผลการประเมินเชิงวิจัญัยจากการทดสอบแบบปรับเหมาะ โดยใช้คอมพิวเตอร์จำนวน 2 ครั้ง จำแนกตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้

เกณฑ์ที่	ค่าสัดส่วนความสอดคล้องของผลการประเมินเชิงวิจัญัยจากการ ทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์จำนวน 2 ครั้ง	
	สถานะความรู้	แบบการคิดที่ผิด
1	0.93	0.91
2	0.96	0.92
3	0.97	0.91
4	0.95	0.95
5	0.91	0.88
6	0.91	0.88
$\chi^2 = 5.473$ df = 5 p = .361		$\chi^2 = 4.593$ df = 5 p = .467

## 2. ผลการเปรียบเทียบความตรง

จากการเปรียบเทียบความตรงของผลการประเมินเชิงวิจัญัยสถานะความรู้และแบบ การคิดที่ผิดระหว่างเกณฑ์ที่ใช้ในการตัดสินความรอบรู้จำนวน 6 เกณฑ์ พบว่า ทั้ง 6 เกณฑ์มีความ ตรงของผลการประเมินเชิงวิจัญัยสถานะความรู้และแบบการคิดที่ผิดไม่แตกต่างกันอย่างมี นัยสำคัญทางสถิติที่ระดับ .05 โดยมีค่าไคสแควร์ เท่ากับ 6.100 ที่ชั้นแห่งความเป็นอิสระ เท่ากับ 5 ทั้งผลการประเมินเชิงวิจัญัยสถานะความรู้และผลการประเมินเชิงวิจัญัยแบบการคิดที่ผิด ดังแสดงในตารางที่ 30

ตารางที่ 30 ผลการเปรียบเทียบความตรงของผลการประเมินเชิงวิจัญัยจากการทดสอบแบบปรับเหมาะ โดยใช้คอมพิวเตอร์และจากการใช้เทคนิคการคิดออกเสียง จำแนกตามเกณฑ์ที่ใช้ในการ ตัดสินความรอบรู้

เกณฑ์ที่	ค่าสัดส่วนความสอดคล้องของผลการประเมินเชิงวิจัญัยจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์และจากการใช้เทคนิคการคิดออกเสียง	
	สถานะความรู้	แบบการคิดที่ผิด
1	0.78	0.78
2	0.78	0.78
3	0.89	0.89
4	0.78	0.78
5	0.89	0.89
6	0.89	0.89
$\chi^2 = 6.100$ df = 5 p = .412		$\chi^2 = 6.100$ df = 5 p = .412

### 3. ผลการเปรียบเทียบจำนวนข้อสอบ

จากการเปรียบเทียบจำนวนข้อสอบที่ใช้ระหว่างเกณฑ์ที่ใช้ในการตัดสินความรอบรู้จำนวน 6 เกณฑ์ พบว่า นักเรียนที่ใช้เกณฑ์ในการตัดสินความรอบรู้ทั้ง 6 เกณฑ์มีจำนวนข้อสอบที่ใช้ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ดังแสดงในตารางที่ 31

ตารางที่ 31 ผลการเปรียบเทียบจำนวนข้อสอบที่ใช้ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

แหล่งความแปรปรวน	SS	df	MS	F
ระหว่างกลุ่ม	7194.5	5	1438.9	1.714
ภายในกลุ่ม	428204.2	510	839.6	
รวม	435398.6	515		

### 4. ผลการเปรียบเทียบเวลาที่ใช้ในการทดสอบ

จากการเปรียบเทียบเวลาที่ใช้ในการทดสอบระหว่างเกณฑ์ที่ใช้ในการตัดสินความรอบรู้จำนวน 6 เกณฑ์ พบว่า นักเรียนที่ใช้เกณฑ์ในการตัดสินความรอบรู้ทั้ง 6 เกณฑ์ใช้เวลาใน



การทดสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เมื่อศึกษาเปรียบเทียบรายคู่ พบว่า เกณฑ์การตัดสินความรอบรู้ที่นักเรียนใช้เวลาแตกต่างกัน คือเกณฑ์การตัดสินความรอบรู้เกณฑ์ที่ 1 และเกณฑ์การตัดสินความรอบรู้เกณฑ์ที่ 3 ดังแสดงในตารางที่ 32

ตารางที่ 32 ผลการเปรียบเทียบเวลาที่ใช้ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

แหล่งความแปรปรวน	SS	df	MS	F
ระหว่างกลุ่ม	2129.7	5	425.9	3.070*
ภายในกลุ่ม	70760.0	510	138.7	
รวม	72889.7	515		

\* มีนัยสำคัญทางสถิติที่ระดับ .05

#### ตอนที่ 4 ผลการศึกษาความคิดเห็นที่มีต่อวิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้น

การศึกษาค้นคว้าความคิดเห็นที่มีต่อวิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้นดำเนินการโดยใช้แบบสอบถามภายหลังจากนักเรียนได้รับการประเมินเชิงวินิจจัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ครั้งที่ 2 โดยศึกษาค้นคว้าความคิดเห็นทั้งจากนักเรียนและครูที่รับผิดชอบดำเนินการจัดให้นักเรียนได้รับการทดสอบปรากฏผลการวิเคราะห์ข้อมูล ดังนี้

##### 1. ความคิดเห็นของครูที่มีต่อวิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้น

ครูที่ตอบแบบสอบถามความคิดเห็นที่มีต่อวิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้นมีจำนวนและข้อมูลทั่วไปเหมือนกับครูที่ตอบแบบสอบถามปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ เนื่องจากเป็นกลุ่มตัวอย่างกลุ่มเดียวกัน ผลการศึกษาค้นคว้าความคิดเห็นของครูที่มีต่อวิธีการประเมินเชิงวินิจจัยในแต่ละด้าน สรุปได้ดังนี้

##### 1.1 ด้านความยากง่ายในการปฏิบัติ

ครูส่วนใหญ่ (ร้อยละ 88.9) เห็นว่าวิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้นสามารถนำไปใช้ได้ง่าย เนื่องจากไม่ต้องออกข้อสอบบ่อย ๆ การประเมินสะดวกรวดเร็ว รู้ผลได้ทันที โปรแกรมติดตั้งง่าย ไม่มีปัญหาขณะใช้งาน แต่นักเรียนบางคนยังไม่มีทักษะการใช้คอมพิวเตอร์ ทำให้บางขั้นตอนดำเนินการได้ช้า เช่น การกรอกข้อมูลผู้สอบ เป็นต้น ส่วนครูที่เห็นว่าวิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้นสามารถนำไปใช้ได้ยาก (ร้อยละ 11.1) ให้เหตุผลว่า

นักเรียนส่วนใหญ่ไม่มีทักษะการใช้คอมพิวเตอร์ต้องดูแลนักเรียนเป็นรายบุคคลทำให้เกิดความยุ่งยากและเสียเวลา นอกจากนี้ครูส่วนใหญ่ยังให้ข้อเสนอแนะว่าในรายงานผลการทดสอบควรมีการสรุปจำนวนข้อที่นักเรียนตอบถูกและตอบผิด รวมทั้งบอกรายละเอียดว่าแบบการคิดที่ผิดแต่ละข้อหมายถึงอะไร จะได้ไม่เสียเวลานับจำนวนข้อและดูเอกสารอื่นประกอบ

## 1.2 ด้านความเป็นไปได้ในการนำไปใช้

ครูส่วนใหญ่ (ร้อยละ 83.3) เห็นว่ามีความเป็นไปได้มากในการนำวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นไปใช้ในโรงเรียน เนื่องจากเป็นวิธีการที่ง่าย สะดวก รวดเร็ว และโรงเรียนส่วนใหญ่มีคอมพิวเตอร์ที่สามารถรองรับโปรแกรมได้ โดยให้ข้อเสนอแนะว่าควรมีการเตรียมความพร้อมเรื่องทักษะการใช้คอมพิวเตอร์ของนักเรียนก่อนจะนำไปใช้ ส่วนครูที่เห็นว่ามีความเป็นไปได้น้อยในการนำวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นไปใช้ในโรงเรียน (ร้อยละ 16.7) ให้เหตุผลว่าโรงเรียนยังไม่มีความพร้อมเรื่องการพัฒนาคลังข้อสอบ เนื่องจากบางโรงเรียนมีครูประจำวิชาเพียง 1 หรือ 2 คน ประกอบกับครูมีภาระงานมากจึงมีความเป็นไปได้น้อยที่จะเตรียมการประเมินในลักษณะนี้ โดยให้ข้อเสนอแนะว่าควรจะมีการดำเนินการจัดทำโปรแกรมคอมพิวเตอร์ที่ใช้ในการประเมินสำหรับแต่ละวิชาจากหน่วยงานกลางเพื่อให้โรงเรียนนำไปใช้ได้โดยไม่ต้องพัฒนาเอง

## 1.3 ด้านประโยชน์ที่ได้รับจากผลการประเมินเชิงวินิจฉัย

ครูทุกคนเห็นว่าผู้ที่เกี่ยวข้องได้รับประโยชน์จากผลการประเมินเชิงวินิจฉัยมาก ดังนี้

1) โรงเรียนมีข้อมูลพื้นฐานเกี่ยวกับระดับความสามารถและจุดบกพร่องของโรงเรียนซึ่งเป็นประโยชน์ต่อการกำหนดนโยบายและการกำหนดแผนงานในการพัฒนานักเรียน อีกทั้งยังเป็นข้อมูลเชิงประจักษ์ที่สามารถแสดงต่อผู้ปกครองและหน่วยงานที่เกี่ยวข้องได้ ครูได้ใช้ผลการประเมินเพื่อการปรับปรุงหรือวางแผนการจัดการเรียนรู้และการพัฒนานักเรียน รวมทั้งยังสามารถใช้ผลการประเมินประกอบการพิจารณาจัดทำผลงานทางวิชาการได้ด้วย

2) นักเรียนได้ทราบระดับความสามารถและจุดที่ควรแก้ไขปรับปรุงของตนเอง ซึ่งจะเป็นประโยชน์ต่อการพัฒนาตนเอง

3) ผู้ปกครองได้ทราบระดับความสามารถและจุดที่ควรแก้ไขปรับปรุงของนักเรียน ซึ่งจะเป็นประโยชน์ต่อการร่วมมือกันระหว่างครูและผู้ปกครองในการพัฒนานักเรียน

#### 1.4 ด้านความพึงพอใจที่มีต่อการใช้

ครูส่วนใหญ่ชอบและจะเลือกใช้วิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้น (ร้อยละ 88.9 และ 83.3 ตามลำดับ) เนื่องจากเป็นวิธีการที่ง่าย สะดวก รวดเร็ว ให้ผลการประเมินที่ชัดเจน เป็นประโยชน์ ส่วนครูที่ไม่ชอบและจะไม่เลือกใช้วิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้น (ร้อยละ 11.1 และ 16.7 ตามลำดับ) ให้เหตุผลว่าตนเองไม่มีทักษะในการใช้คอมพิวเตอร์ กลัวว่าจะมีปัญหาในการนำไปใช้ โดยเฉพาะการให้คำแนะนำกับนักเรียนหรือการแก้ไขปัญหาที่เกิดขึ้นระหว่างการใช้

จำนวนครูที่แสดงความคิดเห็นในแต่ละด้าน แสดงในตารางที่ 33

ตารางที่ 33 ความคิดเห็นของครูที่มีต่อวิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้น

ความคิดเห็นด้าน	จำนวน	ร้อยละ
1. ความยากง่ายในการใช้วิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้น		
1.1 ง่าย	16	88.9
1.2 ยาก	2	11.1
2. ความเป็นไปได้ในการนำวิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้นไปใช้		
2.1 เป็นไปได้	15	83.3
2.2 เป็นไปไม่ได้	3	16.7
3. การได้รับประโยชน์จากผลการประเมินโดยการทดสอบ แบบปรับเหมาะโดยใช้คอมพิวเตอร์		
3.1 ได้รับมาก	18	100
3.2 ได้รับน้อย	0	0.0
4. ความชอบในการใช้วิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้น		
4.1 ชอบ	16	88.9
4.2 ไม่ชอบ	2	11.1
5. การเลือกใช้วิธีการประเมินเชิงวินิจจัยที่พัฒนาขึ้น		
5.1 เลือกใช้	15	83.3
5.2 ไม่เลือกใช้	3	16.7
รวม	18	100

## 2. ความคิดเห็นของนักเรียนที่มีต่อวิธีการประเมินเชิงวิจิตรที่พัฒนาขึ้น

นักเรียนที่ตอบแบบสอบถามความคิดเห็นที่มีต่อการประเมินเชิงวิจิตรที่พัฒนาขึ้น มีข้อมูลทั่วไปดังแสดงในตารางที่ 34

ตารางที่ 34 ข้อมูลทั่วไปของนักเรียนที่ตอบแบบสอบถามความคิดเห็นที่มีต่อการประเมินเชิงวิจิตรที่พัฒนาขึ้น

สภาพทั่วไป	จำนวน	ร้อยละ
ขนาดโรงเรียน		
โรงเรียนขนาดเล็ก (จำนวนนักเรียนต่ำกว่า 200 คน)	138	29.7
โรงเรียนขนาดกลาง (จำนวนนักเรียน 201 ถึง 500 คน)	161	34.6
โรงเรียนขนาดใหญ่ (จำนวนนักเรียน 500 คนขึ้นไป)	166	35.7
ชั้น		
ประถมศึกษาปีที่ 4	63	13.5
ประถมศึกษาปีที่ 5	204	43.9
ประถมศึกษาปีที่ 6	198	42.6
เพศ		
หญิง	246	52.9
ชาย	219	47.1
อายุ		
8 ปี	1	0.2
9 ปี	15	3.2
10 ปี	144	31.0
11 ปี	226	48.6
12 ปี	70	15.1
13 ปี	7	1.5
14 ปี	1	0.2
15 ปี	1	0.2
รวม	465	100

จากตารางที่ 34 แสดงให้เห็นว่านักเรียนที่ตอบแบบสอบถามความคิดเห็นที่มีต่อการประเมินเชิงวิจจัยที่พัฒนาขึ้นมีจำนวน 540 คน มีแบบสอบถามที่สมบูรณ์สามารถนำมาวิเคราะห์ได้จำนวน 465 ฉบับ คิดเป็นร้อยละ 86.1 ผลการวิเคราะห์ข้อมูล พบว่า นักเรียนส่วนใหญ่ศึกษาอยู่ในโรงเรียนขนาดใหญ่ (ร้อยละ 35.7) รองลงมา ได้แก่ โรงเรียนขนาดกลาง (ร้อยละ 34.6) และโรงเรียนขนาดเล็ก (ร้อยละ 29.7) โดยศึกษาอยู่ในชั้นประถมศึกษาปีที่ 5 มากที่สุด (ร้อยละ 43.9) รองลงมา ได้แก่ ชั้นประถมศึกษาปีที่ 6 (ร้อยละ 42.6) และชั้นประถมศึกษาปีที่ 4 (ร้อยละ 13.5) เมื่อพิจารณาด้านเพศของนักเรียน พบว่า นักเรียนส่วนใหญ่เป็นเพศหญิง (ร้อยละ 53.5) มีอายุ 11 ปี มากที่สุด (ร้อยละ 48.6) รองลงมา ได้แก่ 10 ปี 12 ปี 9 ปี 13 ปี คิดเป็นร้อยละ 31.0, 15.1, 3.2 และ 8 ปี 14 ปี 15 ปี คิดเป็นร้อยละ 0.2 ดังแสดงในตารางที่ 34

## 2.1 ความคิดเห็นของนักเรียนที่มีต่อวิธีประเมินเชิงวิจจัยที่พัฒนาขึ้น

ผลการศึกษาความคิดเห็นของนักเรียนที่มีต่อวิธีการประเมินเชิงวิจจัยที่พัฒนาขึ้นในแต่ละด้าน สรุปได้ดังนี้

### 2.2.1 ด้านความยากง่ายในการใช้

นักเรียนส่วนใหญ่ (ร้อยละ 81.1) เห็นว่าการทำแบบทดสอบโดยใช้คอมพิวเตอร์ปฏิบัติได้ง่าย เนื่องจาก เคยเรียนหรือเคยฝึกมาแล้ว มีทักษะในการใช้คอมพิวเตอร์ เข้าใจวิธีการใช้ ขั้นตอนน้อยเข้าใจโปรแกรมได้เลย พิมพ์คำตอบง่ายกว่าเขียนในกระดาษ แค่พิมพ์ตัวเลขลงในช่อง ไม่ต้องเขียน สะดวก รวดเร็ว ได้คิด ได้ความรู้ใหม่ ลบได้ ไม่จำกัดเวลา สนุก ส่วนนักเรียนที่เห็นว่าการทำแบบทดสอบโดยใช้คอมพิวเตอร์ปฏิบัติได้ยาก (ร้อยละ 18.9) ให้เหตุผลว่า ไม่มีทักษะในการใช้คอมพิวเตอร์ ไม่เข้าใจวิธีการใช้ พิมพ์ไม่เก่ง พิมพ์ผิด ไม่มีสะดวกในการทดเลข ไม่คุ้นเคย เป็นต้น

### 2.2.2 ด้านความเป็นไปได้ในการนำไปใช้

นักเรียนส่วนใหญ่ (ร้อยละ 84.5) เห็นว่าการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มีความเป็นไปได้สูงในการนำไปใช้ในโรงเรียน เนื่องจาก สะดวกรวดเร็ว ง่าย เคยเรียนหรือเคยฝึกมาแล้ว นักเรียนเข้าใจและทำได้ โรงเรียนมีคอมพิวเตอร์เพียงพอ ครูอาจชอบที่จะประเมินด้วยวิธีนี้ ไม่ต้องออกข้อสอบบ่อย ๆ ไม่เสียเวลา อยากให้ครูใช้ เป็นวิธีการที่ทันสมัย เป็นประโยชน์กับนักเรียน ได้ฝึกการพิมพ์และฝึกการใช้คอมพิวเตอร์ หน้าจอมีภาพชัดเจน มีที่เก็บข้อมูล ครูเก่งครูจะรู้รึว่านักเรียนเข้าใจหรือไม่ ครูมีสิทธิเลือก ได้พัฒนานักเรียน ไม่เสียเวลาตรวจ ส่วนนักเรียนที่เห็นว่าการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มีความเป็นไปได้ต่ำในการนำไปใช้ในโรงเรียน

(ร้อยละ 5.5) ให้เหตุผลว่า ทำไม่เป็น ครูชอบให้ฝึกคิดและเขียน ครูไม่ได้ออกข้อสอบเอง ปฏิบัติได้ยาก นักเรียนเกิดความสับสน เปลืองไฟฟ้า โรงเรียนไม่มีโปรแกรมทดสอบแบบนี้ เป็นต้น

### 2.2.3 ด้านประโยชน์ที่ได้รับจากผลการประเมินเชิงวิจักษ์

นักเรียนส่วนใหญ่ (ร้อยละ 90.8) เห็นว่าได้รับประโยชน์จากผลการประเมินเชิงวิจักษ์มาก โดยได้รับประโยชน์ในด้านได้ความรู้ ได้ฝึกการพิมพ์และฝึกการใช้คอมพิวเตอร์ ได้ฝึกการคิดคำนวณ ได้พัฒนาสมอง ได้รับประสบการณ์ใหม่ ๆ ได้ฝึกสมาธิ นำไปใช้ประโยชน์ในวิชาคณิตศาสตร์ได้ ได้รู้ระดับความสามารถของตน ได้รู้จักวิธีการทดสอบแบบใหม่ ส่วนนักเรียนที่เห็นว่าได้รับประโยชน์จากผลการประเมินเชิงวิจักษ์น้อย (ร้อยละ 9.2) ไม่ได้แสดงผลประกอบ

### 2.2.4 ด้านความพึงพอใจในการใช้

นักเรียนส่วนใหญ่ (ร้อยละ 88.2) ชอบทำแบบทดสอบแบบปรับเหมาะ โดยใช้คอมพิวเตอร์ เนื่องจาก สนุก ง่าย สะดวกรวดเร็ว ทำได้ มีความรู้เพิ่มเติมในการใช้คอมพิวเตอร์และอุปกรณ์ได้ถูกต้อง ได้ความรู้ ชอบคอมพิวเตอร์ ได้รับประโยชน์มาก มีความรู้ในการทำแบบทดสอบแบบใหม่ ได้ฝึกสมาธิ ได้ฝึกการคิดคำนวณ ช่วยให้เก่งคณิตศาสตร์ ไม่ต้องเขียน ไม่ต้องลอกโจทย์ ลบได้ไม่สกปรก เขียนหนังสือไม่สวย ชอบพิมพ์มากกว่าเขียน รู้ว่าตนเองมีความรู้ระดับใด รู้ผลสอบทันที ได้พัฒนาความสามารถของตน คิดและพิมพ์คำตอบได้เร็ว ส่วนนักเรียนที่ไม่ชอบทำแบบทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (ร้อยละ 11.8) ให้เหตุผลว่า ปฏิบัติได้ยาก พิมพ์ผิด ไม่เข้าใจ ไม่มีทักษะในการใช้คอมพิวเตอร์ ยุ่งยากในการทดเลข ยุ่งยากในการพิมพ์ สับสนในการใช้

เมื่อให้นักเรียนเลือกทำแบบทดสอบระหว่างคอมพิวเตอร์กับการเขียนตอบบนกระดาษคำตอบ พบว่า นักเรียนส่วนใหญ่ (ร้อยละ 85.8) เลือกทำแบบทดสอบโดยใช้คอมพิวเตอร์ เนื่องจาก สะดวก รวดเร็ว ง่าย สนุก ได้ใช้คอมพิวเตอร์ ได้ฝึกพิมพ์ ไม่ต้องเขียน ไม่ต้องลอกโจทย์ ง่ายกว่าการเขียน ได้เรียนรู้สิ่งใหม่ ๆ ชอบคอมพิวเตอร์ ไม่สกปรก ลบได้ ได้ฝึกสมาธิ สีสันสดใส อยู่ในห้องแอร์ วัตถุประสงค์ที่ได้ คำนวณผลได้รวดเร็ว ประหยัดกระดาษ ช่วยรักษาสิ่งแวดล้อม ส่วนนักเรียนที่เลือกทำแบบทดสอบโดยใช้การเขียนตอบบนกระดาษคำตอบ (ร้อยละ 14.2) ให้เหตุผลว่า ไม่มีทักษะในการใช้คอมพิวเตอร์ ไม่มีใครช่วยแนะนำ ได้ฝึกและพัฒนาการเขียน เขียนสะดวกและง่ายกว่าพิมพ์ ทดเลขได้ง่าย ได้กา ก. ข. ค. ง . ประหยัดไฟฟ้า

จำนวนนักเรียนที่แสดงความคิดเห็นในแต่ละด้าน แสดงในตารางที่ 35

ตารางที่ 35 ความคิดเห็นของนักเรียนที่มีต่อวิธีการประเมินเชิงวิจจัยที่พัฒนาขึ้น

ความคิดเห็นด้าน	จำนวน	ร้อยละ
1. ความยากง่ายในการทำแบบทดสอบโดยใช้คอมพิวเตอร์		
1.1 ง่าย	377	81.1
1.2 ยาก	88	18.9
2. ความเป็นไปได้ในการนำการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ไปใช้เพื่อประเมินนักเรียน		
2.1 เป็นไปได้	393	84.5
2.2 เป็นไปไม่ได้	72	5.5
3. การได้รับประโยชน์จากผลการประเมินโดยการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์		
3.1 ได้รับมาก	422	90.8
3.2 ได้รับน้อย	43	9.2
4. ความชอบในการทำแบบทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์		
4.1 ชอบ	410	88.2
4.2 ไม่ชอบ	55	11.8
5. การเลือกทำแบบทดสอบระหว่างการเขียนคำตอบบนกระดาษและการทำโดยใช้คอมพิวเตอร์		
5.1 คอมพิวเตอร์	399	85.8
5.2 กระดาษ	66	14.2
รวม	465	100

## บทที่ 5

### สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้เป็นการพัฒนาวิธีการประเมินเชิงวินิจฉัยโดยประยุกต์ใช้แนวคิดโมเดลลำดับชั้นของคุณลักษณะซึ่งเป็นโมเดลการวินิจฉัยทางพุทธิปัญญาในขั้นตอนการพัฒนาเครื่องมือการประเมินและประยุกต์ใช้การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ซึ่งเป็นการใช้ทฤษฎีการทดสอบแนวใหม่เป็นวิธีการในขั้นตอนการประเมิน มีวัตถุประสงค์หลักของการวิจัย 4 ประการ ประการแรก เพื่อพัฒนาวิธีการประเมินเชิงวินิจฉัย โดยประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ในการประเมินสถานะความรู้และแบบการคิดที่ผิด เรื่องบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) โดยมีวัตถุประสงค์เฉพาะ เพื่อกำหนดลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน เพื่อพัฒนาชุดของข้อสอบและโปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ และเพื่อวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนของนักเรียนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) วัตถุประสงค์ประการที่ 2 เพื่อศึกษาคุณภาพของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น โดยมีวัตถุประสงค์เฉพาะ เพื่อศึกษาความเที่ยงและความตรงของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น รวมทั้งจำนวนเวลาและจำนวนข้อสอบที่ใช้ในการทดสอบ วัตถุประสงค์ประการที่ 3 เพื่อเปรียบเทียบคุณภาพของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น เมื่อข้อกำหนดในการตัดสินความรอบรู้ในแต่ละลำดับชั้นแตกต่างกัน วัตถุประสงค์ข้อสุดท้ายเพื่อศึกษาความคิดเห็นที่มีต่อวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นในด้านความยากง่ายในการใช้ความเป็นไปได้ในการนำไปใช้ ประโยชน์ที่ได้รับจากการวินิจฉัยและความพึงพอใจที่มีต่อการใช้

การพัฒนาวิธีการประเมินเชิงวินิจฉัยจำแนกการดำเนินการเป็น 4 ระยะ ระยะแรกเป็นการพัฒนาชุดของข้อสอบ แบบทดสอบแบบเขียนตอบบนกระดาษคำตอบและคลังข้อสอบ จำแนกเป็น 4 ขั้นตอน ขั้นตอนแรกเป็นการกำหนดลำดับชั้นของคุณลักษณะดำเนินการโดยการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง รวมทั้งสำรวจคุณลักษณะที่ใช้ในการบวกลบเศษส่วนจากครุต้นแบบ วิชาคณิตศาสตร์ ระดับประถมศึกษาของสำนักงานคณะกรรมการการศึกษาแห่งชาติ (สกศ.) เพื่อจำแนกคุณลักษณะที่ใช้ในการบวกลบเศษส่วน นำคุณลักษณะที่ใช้ในการบวกลบเศษส่วนที่จำแนกได้ให้ผู้เชี่ยวชาญด้านการจัดการเรียนการสอนคณิตศาสตร์ ด้านหลักสูตรและการสอนคณิตศาสตร์ และด้านการวัดและประเมินผล ตรวจสอบความถูกต้องของคุณลักษณะ นำคุณลักษณะที่ผ่านการ



ตรวจสอบความถูกต้องให้ครูซึ่งเป็นกลุ่มเดียวกับครูที่ตอบแบบสำรวจคุณลักษณะจัดเรียงลำดับ  
ชั้นของคุณลักษณะโดยใช้กระบวนการให้น้ำหนักรายคู่ (Paired – weighting procedure) นำ  
ลำดับชั้นของคุณลักษณะที่ได้มาจัดทำเป็นแผนภาพลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบ  
เศษส่วน นำแผนภาพลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วนให้ผู้เชี่ยวชาญกลุ่มเดียวกับ  
ผู้เชี่ยวชาญในการตรวจสอบความถูกต้องของคุณลักษณะตรวจสอบความถูกต้องของแผนภาพ

ขั้นตอนที่ 2 เป็นการจัดทำชุดของข้อสอบ ดำเนินการโดย กำหนด binary adjacency matrix  
(A) เพื่อแสดงความสัมพันธ์ทางตรงระหว่างคุณลักษณะที่ใช้ในการบวกลบเศษส่วน กำหนด  
reachability matrix (R) เพื่อแสดงความสัมพันธ์ทางตรงและความสัมพันธ์ทางอ้อมระหว่าง  
คุณลักษณะที่ใช้ในการบวกลบเศษส่วน กำหนด incidence matrix (Q) เพื่อแสดงชุดของข้อสอบที่  
ต้องทำในแต่ละขั้นตอน และกำหนด reduced Q matrix (Q<sub>r</sub>) เพื่อลดจำนวนชุดของข้อสอบที่ต้อง  
ทำในแต่ละขั้นตอนตามเงื่อนไขโครงสร้างลำดับชั้นที่กำหนด หลังจากนั้นจึงกำหนดแบบแผนการ  
ตอบข้อสอบที่คาดหวังจาก reduced Q matrix (Q<sub>r</sub>) ขั้นตอนต่อไปดำเนินการกำหนดแบบการคิดที่  
ผิดที่ใช้ในการบวกลบเศษส่วน โดยให้นักเรียนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษา  
ปีที่ 4-6) ที่เรียนเรื่องการบวกลบเศษส่วนแล้ว จากโรงเรียนขนาดเล็ก (จำนวนนักเรียนต่ำกว่า 200  
คน) โรงเรียนขนาดกลาง (จำนวนนักเรียนระหว่าง 201 – 500 คน) และโรงเรียนขนาดใหญ่  
(จำนวนนักเรียนมากกว่า 500 คน) จำนวน 210 คน ทำแบบทดสอบแบบแสดงวิธีทำ เรื่องการบว  
กลบเศษส่วน จำนวน 16 ข้อ นำผลการสอบที่ได้มาวิเคราะห์แบบการคิดที่ผิดที่ใช้ในการบวกลบ  
เศษส่วนที่เป็นไปได้ทั้งหมดและนำไปให้ครูผู้สอนวิชาคณิตศาสตร์จากโรงเรียนแกนนำการเรียน  
การสอนคณิตศาสตร์ ระดับประถมศึกษาของสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี  
(สสวท.) จำนวน 12 คน ประเมินความครบถ้วนและความครอบคลุมของแบบการคิด  
ที่ผิดที่ใช้ในการบวกลบเศษส่วน หลังจากนั้นจึงจัดทำชุดของข้อสอบสำหรับแต่ละคุณลักษณะ  
โดยกำหนดให้แต่ละคุณลักษณะมีข้อสอบจำนวน 122 ข้อ รวมจำนวน 488 ข้อ

ขั้นตอนที่ 3 เป็นการจัดทำแบบทดสอบแบบเขียนตอบบนกระดาษคำตอบ ดำเนินการโดย  
นำข้อสอบในแต่ละลำดับชั้นมาจัดทำแบบทดสอบแบบเขียนตอบบนกระดาษคำตอบ ได้แบบทดสอบ  
จำนวน 20 ฉบับ ฉบับละ 32 ข้อ จำแนกเป็นข้อสอบจากชุดของข้อสอบลำดับชั้นละ 8 ข้อ ประกอบด้วย  
ข้อสอบเฉพาะ จำนวน 6 ข้อและข้อสอบร่วม จำนวน 2 ข้อ นำแบบทดสอบไปทดลองใช้กับนักเรียน  
ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) จากโรงเรียนขนาดเล็ก โรงเรียนขนาด  
กลางและโรงเรียนขนาดใหญ่ จำนวน 1,800 คน นำผลการสอบที่ได้มาวิเคราะห์ค่าพารามิเตอร์ของ  
ข้อสอบโดยใช้โมเดลการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ และวิเคราะห์ค่าความเที่ยงของ

แบบทดสอบโดยวิธีตรวจสอบความสอดคล้องภายใน (internal consistency method) ด้วยวิธีของ คูเดอร์-ริชาร์ดสัน (Kuder-Richardson estimates) โดยใช้โปรแกรม ITEM ANALYSIS SYSTEM: IAS 2006 พัฒนาโดย ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี และรองศาสตราจารย์ ดร. คณิต ไข่มุกด์

ขั้นตอนสุดท้าย เป็นการจัดทำคลังข้อสอบ ดำเนินการโดยการแปลงค่าพารามิเตอร์ของ ข้อสอบแต่ละฉบับที่ได้จากการวิเคราะห์ให้อยู่บนสเกลเดียวกัน (common scale) โดยเทียบ คะแนนตามแบบแผนการใช้ข้อสอบร่วม (anchor item design) ซึ่งหาค่าคงที่สำหรับการแปลง คะแนนด้วยวิธีค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐาน (mean and sigma method) คัดเลือกข้อสอบที่มีค่าอำนาจจำแนก อยู่ระหว่าง 0.50 ถึง 2.50 ค่าความยาก อยู่ระหว่าง -2.50 ถึง 2.50 ค่าการเดาไม่ เกิน 0.30 ได้ข้อสอบจำนวน 386 ข้อ และคัดเลือกข้อสอบที่มีค่าความยากระหว่าง -3.00 ถึง -2.51 และ 2.49 ถึง 3.00 เพิ่มเติมเพื่อให้ข้อสอบมีค่าความยากครอบคลุมระดับความสามารถของผู้สอบ ได้ข้อสอบจำนวน 52 ข้อ รวมข้อสอบในคลังทั้งหมด จำนวน 438 ข้อ

ระยะที่ 2 เป็นการพัฒนาโปรแกรมคอมพิวเตอร์เพื่อใช้ในการประเมินเชิงวินิจฉัย จำแนกการดำเนินการเป็น 3 ขั้นตอน ขั้นตอนแรกเป็นการออกแบบโปรแกรมคอมพิวเตอร์ ดำเนินการโดยกำหนดลักษณะเฉพาะของโปรแกรมให้สามารถทำงานบนระบบปฏิบัติการวินโดวส์ และใช้กับเครื่องไมโครคอมพิวเตอร์ได้ โดยกำหนดใช้ภาษา Borland Delphi 7.0 ในการเขียน โปรแกรม เนื่องจากเป็นภาษาที่สามารถเขียนภาพกราฟิกและเขียนสูตรสมการฟังก์ชันทาง คณิตศาสตร์ได้ รวมทั้งเป็นภาษาที่มีโครงสร้างชัดเจน ง่ายต่อการใช้งาน หลังจากนั้นจึงกำหนด ข้อกำหนดในการเขียนโปรแกรม โดยจำแนกเป็นโปรแกรมย่อย 6 โปรแกรม แต่ละโปรแกรมกำหนดใช้ วิธีการคัดเลือกข้อสอบข้อแรก วิธีการคัดเลือกข้อสอบจากคลัง วิธีการประมาณค่าความสามารถ ของผู้สอบและการรายงานผลการทดสอบเหมือนกันแต่ต่างกันที่จำนวนข้อสอบที่กำหนดให้ทำใน แต่ละลำดับขั้น เกณฑ์การเลื่อนลำดับขั้นของคุณลักษณะเพื่อเลือกข้อสอบข้อต่อไปและเกณฑ์การ หยุดสอบ ขั้นตอนที่ 2 เป็นการสร้างและพัฒนาโปรแกรมคอมพิวเตอร์ดำเนินการโดยเขียน ผังโปรแกรมการทำงานของคอมพิวเตอร์ในการดำเนินการสอบและเขียนโปรแกรมคอมพิวเตอร์ ตามผังโปรแกรมที่กำหนดไว้ ทดลองการทำงานของโปรแกรมคอมพิวเตอร์เบื้องต้นและปรับปรุง โปรแกรมคอมพิวเตอร์เบื้องต้น ขั้นตอนสุดท้ายเป็นการทดลองใช้โปรแกรมคอมพิวเตอร์ ดำเนินการโดยนำโปรแกรมคอมพิวเตอร์ที่พัฒนาขึ้นไปให้นักเรียนชั้นประถมศึกษาปีที่ 4 – 6 จาก โรงเรียนขนาดเล็ก โรงเรียนขนาดกลางและโรงเรียนขนาดใหญ่ จำนวน 54 คน ทดลองใช้โปรแกรม คอมพิวเตอร์ โปรแกรมย่อยละ 9 คน โดยให้นักเรียนและครูตอบแบบสอบถามปัญหาที่เกิดจาก การใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะหลังการทดลองใช้โปรแกรม นำผลที่ได้

จากการทดลองใช้โปรแกรมคอมพิวเตอร์มาวิเคราะห์และสรุปเพื่อปรับปรุงแก้ไขโปรแกรมคอมพิวเตอร์ให้สามารถใช้งานได้อย่างมีประสิทธิภาพยิ่งขึ้น

ระยะที่ 3 เป็นการประเมินสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน จำแนกการดำเนินการเป็น 2 ขั้นตอน ขั้นตอนแรกเป็นการประเมินสถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ดำเนินการ โดยนำโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะไปทดสอบนักเรียน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) จากโรงเรียนขนาดเล็ก โรงเรียนขนาดกลางและโรงเรียนขนาดใหญ่ จำนวน 540 คน โดยดำเนินการทดสอบ จำนวน 2 ครั้ง ห่างกัน 1 สัปดาห์ และให้ครูและนักเรียนตอบแบบสอบถามปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะหลังการทดสอบครั้งที่ 2 ขั้นตอนที่ 2 เป็นการดำเนินการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วนโดยใช้เทคนิคการคิดออกเสียงดำเนินการโดยผู้สอนตัวอย่างนักเรียนที่ได้รับการประเมินเชิงวิจักษ์ด้วยการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์จากโรงเรียนขนาดเล็ก โรงเรียนขนาดกลางและโรงเรียนขนาดใหญ่ ขนาดโรงเรียนละ 18 คน จำแนกตามโปรแกรมย่อย โปรแกรมย่อยละ 9 คน เพื่อให้ครูผู้สอนวิชาคณิตศาสตร์และผู้วิจัยดำเนินการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนโดยใช้เทคนิคการคิดออกเสียง ตามคู่มือการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนโดยใช้เทคนิคการคิดออกเสียงที่กำหนดให้

ระยะสุดท้ายเป็นการศึกษาคุณภาพและความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น จำแนกการดำเนินการเป็น 3 ขั้นตอน ขั้นตอนแรกเป็นการศึกษาคุณภาพของวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น ดำเนินการโดยนำผลการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้งมาวิเคราะห์ค่าความเที่ยงโดยการคำนวณค่าสัดส่วนของความสอดคล้อง ส่วนค่าความตรงดำเนินการโดยนำผลการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์และการใช้เทคนิคการคิดออกเสียงมาคำนวณค่าสัดส่วนของความสอดคล้อง สำหรับจำนวนข้อสอบและจำนวนเวลาที่ใช้ในการทดสอบวิเคราะห์จากผลการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ทั้ง 2 ครั้ง ดำเนินการโดยคำนวณค่าความถี่และร้อยละ ค่าเฉลี่ยเลขคณิตและส่วนเบี่ยงเบนมาตรฐาน ขั้นตอนที่ 2 เป็นการเปรียบเทียบคุณภาพของวิธีการประเมินเชิงวิจักษ์ ดำเนินการเปรียบเทียบค่าความเที่ยงและ

ค่าความตรง เมื่อเกณฑ์ที่ใช้ในการตัดสินความรอบรู้แตกต่างกัน โดยทำการทดสอบไคสแควร์เพื่อทดสอบความแตกต่างของค่าสัดส่วนของความสอดคล้อง และดำเนินการเปรียบเทียบจำนวนข้อสอบและจำนวนเวลาที่ใช้ในการทดสอบเมื่อเกณฑ์ที่ใช้ในการตัดสินความรอบรู้แตกต่างกัน โดยใช้การวิเคราะห์ความแปรปรวนทางเดียวเพื่อทดสอบความแตกต่างของค่าเฉลี่ย ขั้นตอนสุดท้ายเป็นการศึกษาความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจิตรที่พัฒนาขึ้นดำเนินการโดยให้นักเรียนและครูที่เป็นกลุ่มตัวอย่างในขั้นตอนการประเมินเชิงวิจิตรสถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ตอบแบบสอบถามความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจิตรที่พัฒนาขึ้นหลังการประเมินครั้งที่ 2

การวิเคราะห์ข้อมูล ดำเนินการหลังจากเก็บรวบรวมข้อมูลแต่ละขั้นตอนเรียบร้อยแล้ว จำแนกการดำเนินการเป็น 3 ขั้นตอน ขั้นตอนแรกเป็นการตรวจสอบความถูกต้องสมบูรณ์ของข้อมูล ขั้นตอนที่ 2 เป็นการจัดเตรียมข้อมูลเพื่อการวิเคราะห์และขั้นตอนสุดท้ายเป็นการดำเนินการวิเคราะห์ข้อมูลซึ่งจำแนกเป็นข้อมูลเชิงปริมาณและข้อมูลเชิงคุณภาพ ข้อมูลเชิงปริมาณวิเคราะห์โดยใช้โปรแกรมสำเร็จรูป SPSS/PC ประกอบด้วยการวิเคราะห์สถิติบรรยาย ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ยเลขคณิตและส่วนเบี่ยงเบนมาตรฐานและการวิเคราะห์สถิติอนุมาน ได้แก่ การทดสอบไคสแควร์และการวิเคราะห์ความแปรปรวนทางเดียว ส่วนข้อมูลเชิงคุณภาพใช้การวิเคราะห์เนื้อหาและสรุปเป็นความเรียง นำเสนอข้อมูลโดยใช้ตาราง แผนภาพ และแผนภูมิประกอบคำบรรยาย

## สรุปผลการวิจัย

### 1. ผลการพัฒนาวิธีการประเมินเชิงวิจิตรโดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

วิธีการประเมินเชิงวิจิตรโดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มีเครื่องมือการประเมินที่สำคัญ คือ โปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ซึ่งประกอบด้วยองค์ประกอบหลัก จำนวน 2 องค์ประกอบ ได้แก่ คลังข้อสอบและวิธีการทดสอบ คลังข้อสอบพัฒนาโดยใช้แนวคิดโมเดลลำดับขั้นของคุณลักษณะ ประกอบด้วยชุดของข้อสอบแบบเติมคำจำนวน 4 ชุด ได้แก่ คุณลักษณะที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน มีจำนวน 105 ข้อ คุณลักษณะที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน มีจำนวน 115 ข้อ คุณลักษณะที่ใช้ในการบวกลบจำนวนคละที่มีตัวส่วนเท่ากันมีจำนวน 106 ข้อ และคุณลักษณะที่ใช้ในการบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากันมีจำนวน 112 ข้อ รวมจำนวน 438 ข้อ ชุดของข้อสอบมีความสัมพันธ์เชิงลำดับขั้น 3 ลำดับ

ลำดับขั้นที่ 1 ได้แก่ คุณลักษณะที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน ลำดับขั้นที่ 2 ได้แก่ คุณลักษณะที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากันและคุณลักษณะที่ใช้ในการบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน ลำดับขั้นที่ 3 ได้แก่ คุณลักษณะที่ใช้ในการบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน ข้อสอบแต่ละข้อจะประกอบด้วยคำตอบถูกและคำตอบผิดสำหรับแต่ละแบบการคิดที่ผิด โดยมีแนวคิดในการจำแนกแบบการคิดที่ผิดเป็น 2 ส่วน ได้แก่ ส่วนที่เป็นความเข้าใจในวิธีดำเนินการและส่วนที่เป็นวิธีดำเนินการ แต่ละส่วนจำแนกวัตถุประสงค์ได้ 2 ประการ ประการแรก เพื่อให้ได้ผลลัพธ์และประการที่ 2 เพื่อให้ได้คำตอบที่สมบูรณ์ แบบการคิดที่ผิดสำหรับแต่ละคุณลักษณะในแต่ละส่วนอาจประกอบด้วยวัตถุประสงค์เพียงข้อเดียวหรือมากกว่า ทั้งนี้ขึ้นอยู่กับระดับความซับซ้อนของคุณลักษณะ

ด้านวิธีการทดสอบใช้การทำข้อสอบผ่านโปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ซึ่งจำแนกเป็นโปรแกรมการพิมพ์ข้อสอบและโปรแกรมการทดสอบ โปรแกรมการพิมพ์ข้อสอบมีไว้สำหรับเพิ่ม ลดหรือเปลี่ยนแปลงแก้ไขข้อสอบในคลังข้อสอบ โดยกำหนดสารสนเทศเบื้องต้นของข้อสอบไว้ 4 ส่วน ได้แก่ ข้อคำถาม คำตอบถูก คำตอบผิดสำหรับแต่ละแบบการคิดที่ผิด และค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยาก ค่าอำนาจจำแนกและค่าการเดา ส่วนโปรแกรมการทดสอบจำแนกเป็นโปรแกรมย่อยจำนวน 6 โปรแกรม แต่ละโปรแกรมใช้วิธีการคัดเลือกข้อสอบข้อแรก วิธีการคัดเลือกข้อสอบจากคลังข้อสอบ วิธีการประมาณค่าความสามารถของผู้สอบและการรายงานผลการทดสอบเหมือนกัน แต่ต่างกันที่จำนวนข้อสอบที่กำหนดให้ทำในแต่ละลำดับขั้น เกณฑ์การเลื่อนลำดับขั้นของคุณลักษณะเพื่อเลือกข้อสอบข้อต่อไปและเกณฑ์การหยุดสอบ ดังนี้

โปรแกรมย่อยที่ 1 กำหนดข้อสอบขั้นต่ำให้ทำจำนวน 3 ข้อ ตอบถูก จำนวน 2 ใน 3 ข้อ คัดเลือกข้อสอบจากชุดของข้อสอบลำดับขั้นต่อไป ตอบผิดโดยใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 2 ใน 3 ข้อ หยุดสอบ

โปรแกรมย่อยที่ 2 กำหนดข้อสอบขั้นต่ำให้ทำจำนวน 4 ข้อ ตอบถูก จำนวน 3 ใน 4 ข้อ คัดเลือกข้อสอบจากชุดของข้อสอบลำดับขั้นต่อไป ตอบผิดโดยใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 3 ใน 4 ข้อ หยุดสอบ

โปรแกรมย่อยที่ 3 กำหนดข้อสอบขั้นต่ำให้ทำจำนวน 5 ข้อ ตอบถูก จำนวน 4 ใน 5 ข้อ คัดเลือกข้อสอบจากชุดของข้อสอบลำดับขั้นต่อไป ตอบผิดโดยใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 4 ใน 5 ข้อ หยุดสอบ

โปรแกรมย่อยที่ 4 กำหนดข้อสอบขั้นต่ำให้ทำจำนวน 3 ข้อ ตอบถูก จำนวน 2 ใน 3 ข้อ คัดเลือกข้อสอบจากชุดของข้อสอบลำดับขั้นต่อไป ตอบผิดโดยใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 2 ใน 3 ข้อ และมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3 หยุดสอบ

โปรแกรมย่อยที่ 5 กำหนดข้อสอบขั้นต่ำให้ทำจำนวน 4 ข้อ ตอบถูก จำนวน 3 ใน 4 ข้อ คัดเลือกข้อสอบจากชุดของข้อสอบลำดับขั้นต่อไป ตอบผิดโดยใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 3 ใน 4 ข้อและมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3 หยุดสอบ

โปรแกรมย่อยที่ 6 กำหนดข้อสอบขั้นต่ำให้ทำจำนวน 5 ข้อ ตอบถูก จำนวน 4 ใน 5 ข้อ คัดเลือกข้อสอบจากชุดของข้อสอบลำดับขั้นต่อไป ตอบผิดโดยใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 4 ใน 5 ข้อและมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3 หยุดสอบ

โปรแกรมย่อยทั้ง 6 โปรแกรม กำหนดลำดับขั้นในการเลือกข้อสอบจากคลังข้อสอบ เหมือนกัน โดยกำหนดเป็น 4 กรณี ดังนี้

กรณีที่ 1 เลือกข้อสอบจากคุณลักษณะที่ 1 ผ่านคุณลักษณะที่ 1 เลือกข้อสอบจากคุณลักษณะที่ 2 ผ่านคุณลักษณะที่ 2 เลือกข้อสอบจากคุณลักษณะที่ 3 ผ่านคุณลักษณะที่ 3 เลือกข้อสอบจากคุณลักษณะที่ 4 หยุดสอบ

กรณีที่ 2 เลือกข้อสอบจากคุณลักษณะที่ 1 ผ่านคุณลักษณะที่ 1 เลือกข้อสอบจากคุณลักษณะที่ 2 ไม่ผ่านคุณลักษณะที่ 2 เลือกข้อสอบจากคุณลักษณะที่ 3 ผ่านคุณลักษณะที่ 3 หยุดสอบ

กรณีที่ 3 เลือกข้อสอบจากคุณลักษณะที่ 1 ผ่านคุณลักษณะที่ 1 เลือกข้อสอบจากคุณลักษณะที่ 2 ผ่านคุณลักษณะที่ 2 เลือกข้อสอบจากคุณลักษณะที่ 3 ไม่ผ่านคุณลักษณะที่ 3 หยุดสอบ

กรณีที่ 4 เลือกข้อสอบจากคุณลักษณะที่ 1 ไม่ผ่านคุณลักษณะที่ 1 หยุดสอบ

ด้านผลการประเมินเชิงวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนของนักเรียน พบว่า ส่วนใหญ่เป็นกลุ่มนักเรียนที่มีความรู้ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนเพียงบางส่วน โดยมีความรู้ ทักษะหรือกระบวนการในการบวกลบเศษส่วนที่มีตัวส่วนเท่ากันมากกว่าเศษส่วนที่มีตัวส่วนไม่เท่ากัน ทั้งการบวกลบเศษส่วนแต่ละจำนวนคล่องลงมา ได้แก่ เป็นกลุ่มนักเรียนที่ไม่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนในทุกเรื่อง และเป็นกลุ่มนักเรียนที่มีความรู้ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนในทุกเรื่องซึ่งมีจำนวนน้อยที่สุด โดยมีจำนวนน้อยกว่าร้อยละ 5 ของนักเรียนทั้งหมด

เมื่อพิจารณาผลจากการประเมินสำหรับกลุ่มนักเรียนที่ไม่มีความรู้ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วน พบว่า จำแนกนักเรียนได้ 3 กลุ่ม ประกอบด้วย กลุ่มที่มีแบบการคิดที่ผิด กลุ่มที่มีผลการประเมินวินิจฉัยไม่ได้และกลุ่มที่ไม่ได้รับการทดสอบในคุณลักษณะนั้น สำหรับผลการประเมินในคุณลักษณะที่ 1 และคุณลักษณะที่ 2 พบว่า ส่วนใหญ่

เป็นนักเรียนที่มีแบบการคิดที่ผิด ส่วนคุณลักษณะที่ 3 พบว่าส่วนใหญ่เป็นนักเรียนในกลุ่มนักเรียนที่ไม่ได้ทดสอบในคุณลักษณะนั้นซึ่งมีจำนวนเท่ากับนักเรียนกลุ่มดังกล่าวในคุณลักษณะที่ 2 เนื่องจากเป็นคุณลักษณะที่อยู่ในลำดับชั้นเดียวกัน ส่วนคุณลักษณะที่ 4 พบว่า ส่วนใหญ่เป็นนักเรียนที่ไม่ได้รับการทดสอบในลำดับชั้นนั้นเช่นเดียวกับคุณลักษณะที่ 3 แต่มีจำนวนมากกว่า เนื่องจากเป็นคุณลักษณะในลำดับชั้นสุดท้ายที่ต้องผ่านคุณลักษณะอื่น ๆ ตามลักษณะความสัมพันธ์เชิงลำดับชั้นจึงจะได้รับการทดสอบในลำดับชั้นนี้

เมื่อพิจารณาแบบการคิดที่ผิดที่ใช้ พบว่า คุณลักษณะที่ 1 นักเรียนส่วนใหญ่ไม่เข้าใจวิธีการบวกเศษส่วนแท้ที่มีตัวส่วนเท่ากัน โดยใช้วิธีเศษบวกลบเศษ ตัวส่วนบวกลบตัวส่วนมากกว่าใช้วิธีเศษส่วนตัวตั้งบวกลบกันเป็นเศษ เศษส่วนตัวบวกลบ บวกลบกันเป็นตัวส่วน สำหรับคุณลักษณะที่ 2 พบว่า นักเรียนส่วนใหญ่ไม่เข้าใจวิธีการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน โดยใช้วิธีเศษบวกลบเศษ ตัวส่วนบวกลบตัวส่วนมากกว่าวิธีอื่น ๆ ส่วนคุณลักษณะที่ 3 พบว่า นักเรียนส่วนใหญ่ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน และคำตอบไม่สมบูรณ์ โดยใช้วิธีตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เศษบวกลบเศษและไม่ทำเศษเกินให้เป็นจำนวนคละหรือไม่ทำให้เป็นเศษส่วนอย่างต่ำมากกว่าวิธีอื่น ๆ สำหรับคุณลักษณะที่ 4 พบว่า นักเรียนส่วนใหญ่ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน ไม่เข้าใจวิธีการทำตัวส่วนให้เท่ากันและคำตอบไม่สมบูรณ์โดยใช้วิธีตัวส่วนบวกจำนวนนับ บวกเศษเป็นเศษ เอาตัวส่วนคูณกันเป็นตัวส่วน เศษบวกลบเศษหรือหา ค.ร.น เป็นตัวส่วน เศษบวกลบเศษ ไม่ทำเศษเกินให้เป็นจำนวนคละหรือไม่ทำให้เป็นเศษส่วนอย่างต่ำมากกว่าวิธีอื่น ๆ

## 2. ผลการศึกษาคุณภาพของวิธีการประเมินเชิงวินิจฉัยขั้นที่พัฒนาขึ้น

ผลการศึกษาคุณภาพของวิธีการประเมินเชิงวินิจฉัยขั้นที่พัฒนาขึ้น พิจารณาจากจำนวนข้อสอบและเวลาที่ใช้ในการทดสอบ ค่าความเที่ยงและค่าความตรงของวิธีการประเมินเชิงวินิจฉัยขั้นที่พัฒนาขึ้น พบว่านักเรียนใช้เวลาในการทดสอบเฉลี่ย จำนวน 17.0 นาที โดยใช้เวลาดำสุด จำนวน 1 นาที และสูงสุด จำนวน 58 นาที ส่วนข้อสอบที่ทำทั้งหมด มีจำนวนเฉลี่ย 40.7 ข้อ โดยมีจำนวนข้อต่ำสุด 3 ข้อ และสูงสุด 160 ข้อ ด้านค่าความเที่ยงและค่าความตรง พบว่า วิธีการประเมินเชิงวินิจฉัยขั้นที่พัฒนาขึ้นมีค่าความเที่ยงซึ่งพิจารณาจากค่าสัดส่วนของความสอดคล้องระหว่างผลการประเมินเชิงวินิจฉัยโดยใช้การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง สำหรับการประเมินเชิงวินิจฉัยสถานะความรู้ เท่ากับ 0.94 และสำหรับการประเมินเชิงวินิจฉัยแบบการคิดที่ผิด เท่ากับ 0.86 ส่วนค่าความตรงซึ่งพิจารณาจากค่าสัดส่วนของความสอดคล้อง

ระหว่างผลการประเมินเชิงวิจักษ์จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 2 ครั้ง กับผลการประเมินเชิงวิจักษ์โดยใช้เทคนิคการคิดออกเสียง สำหรับการประเมินเชิงวิจักษ์สถานะความรู้ เท่ากับ 0.83 และ 0.81 ส่วนการประเมินเชิงวิจักษ์แบบการคิดที่ผิด เท่ากับ 0.86 และ 0.84

### 3. ผลการเปรียบเทียบคุณภาพของวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น

ผลการเปรียบเทียบคุณภาพของวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น เมื่อเกณฑ์ในการตัดสินความรอบรู้แตกต่างกันโดยพิจารณาเปรียบเทียบระหว่างโปรแกรมย่อย 6 โปรแกรม พบว่า ค่าความเที่ยง ค่าความตรงและจำนวนข้อสอบที่ใช้ในการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 ส่วนเวลาที่ใช้ในการทดสอบพบว่าแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 เมื่อศึกษาเปรียบเทียบรายคู่พบว่าเกณฑ์การตัดสินความรอบรู้ที่ใช้เวลาแตกต่างกัน คือ เกณฑ์ที่ 1 และเกณฑ์ที่ 3

### 4. ผลการศึกษาความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น

ผลการศึกษาความคิดเห็นของครูและนักเรียนที่มีต่อวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น พบว่า ครูและนักเรียนส่วนใหญ่เห็นว่าวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้นสามารถนำไปใช้ได้ง่าย มีความเป็นไปได้ในการนำไปใช้ ได้รับประโยชน์จากผลการประเมินมาก ชอบและเลือกที่จะได้รับการประเมินหรือใช้วิธีการประเมินที่พัฒนาขึ้นมากกว่าการประเมินแบบเดิม

### อภิปรายผลการวิจัย

การวิจัยครั้งนี้เป็นการพัฒนาวิธีการประเมินเชิงวิจักษ์โดยการประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จากการดำเนินการวิจัยและผลการวิจัยพบประเด็นที่น่าสนใจและน่าสนใจนำมาอภิปรายหลายประเด็น โดยจำแนกการนำเสนอเป็น 2 ส่วน ส่วนที่ 1 เป็นประเด็นเกี่ยวกับการพัฒนาวิธีการประเมินเชิงวิจักษ์โดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ และส่วนที่ 2 เป็นประเด็นเกี่ยวกับคุณภาพของวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น มีรายละเอียดดังนี้



## 1. การพัฒนาวิธีการประเมินเชิงวิจักษ์โดยการประยุกต์ใช้โมเดลลำดับชั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

เครื่องมือการประเมินสำหรับวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้น คือ โปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะเพื่อการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน ซึ่งเป็นการประยุกต์ใช้แนวคิดโมเดลลำดับชั้นของคุณลักษณะในการพัฒนาค้างข้อสอบและการกำหนดลำดับชั้นในการเลือกข้อสอบจากคลังข้อสอบ โปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะสำหรับการประเมินเชิงวิจักษ์ในครั้งนี้จึงแตกต่างจากโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะโดยทั่วไป กล่าวคือ คลังข้อสอบจำแนกออกเป็นชุดของข้อสอบซึ่งมีจำนวนชุดของข้อสอบตามจำนวนของคุณลักษณะที่กำหนดได้ในขั้นตอนแรก ชุดของข้อสอบแต่ละชุดจะมีความสัมพันธ์กันเชิงลำดับชั้น นอกจากนี้ยังมีการเพิ่มสารสนเทศในคลังข้อสอบในส่วนของคุณลักษณะคำตอบผิดสำหรับแต่ละแบบการคิดที่ผิด เนื่องจากเป็นโปรแกรมคอมพิวเตอร์ที่มีวัตถุประสงค์เพื่อใช้ในการทดสอบแบบปรับเหมาะ ซึ่งสามารถให้ผลการทดสอบเพื่อการประเมินเชิงวิจักษ์ได้ทั้งจากแบบแผนการตอบถูกและแบบแผนการตอบผิด ด้านลำดับชั้นในการเลือกข้อสอบจากคลังข้อสอบมีลักษณะเฉพาะตามแผนภาพลำดับชั้นของคุณลักษณะที่กำหนดได้ จากลักษณะของโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะดังกล่าวจึงทำให้วิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้นให้สารสนเทศการวิจักษ์ที่สามารถอธิบายได้ด้วยทฤษฎีทางพุทธิปัญญา ทั้งสถานะความรู้ที่เป็นจุดเด่นและจุดที่ควรพัฒนา รวมถึงแบบการคิดที่ผิดซึ่งบ่งบอกถึงมโนทัศน์ที่คลาดเคลื่อนของนักเรียนอันเป็นประโยชน์ต่อการพัฒนาการจัดการเรียนรู้และการพัฒนาผู้เรียนสอดคล้องกับงานวิจัยของ Wang and Gierl (2007) ซึ่งได้ใช้โมเดลลำดับชั้นของคุณลักษณะในการประเมิน The SAT Critical Reading Subtest และผลการวิจัยพบว่าโมเดลลำดับชั้นของคุณลักษณะมีประโยชน์มากกว่าวิธีการให้คะแนนจากการวัดทางจิตแบบดั้งเดิม (traditional psychometric scoring methods) โดยสามารถให้ทั้งคะแนนรวมและสารสนเทศการวิจักษ์ทางพุทธิปัญญาแก่ผู้สอบซึ่งเป็นสารสนเทศที่ช่วยให้ผู้สอบเห็นถึงจุดที่ควรแก้ไขในแต่ละคุณลักษณะที่ไม่ผ่าน โมเดลลำดับชั้นของคุณลักษณะจึงเป็นเครื่องมือในการพัฒนาการเรียนรู้ของผู้เรียน นอกจากนี้ยังช่วยเพิ่มความตรงเชิงโครงสร้างของแบบทดสอบเนื่องจากในขั้นตอนการสร้างแบบทดสอบจะต้องมีการศึกษาวิเคราะห์อย่างละเอียดเพื่อกำหนดลำดับชั้นของคุณลักษณะ รวมทั้งสามารถให้ข้อมูลแก่นักวิจัยในการศึกษาธรรมชาติของความสามารถในการอ่านและการพัฒนาทฤษฎีทางพุทธิปัญญาในการอ่านได้อีกด้วย

ในส่วนของวิธีการทดสอบซึ่งใช้การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จะเห็นได้ว่าแผนภาพลำดับชั้นของคุณลักษณะที่กำหนดขึ้นช่วยให้การดำเนินการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มีแนวทางการคัดเลือกข้อสอบที่ชัดเจนและครอบคลุมเนื้อหา มากกว่าการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบดั้งเดิมซึ่งไม่ได้นำโมเดลการวินิจฉัยทางพุทธิปัญญามาใช้ แม้ว่าจะจะเป็นแบบทดสอบแบบปรับเหมาะบนพื้นฐานของเนื้อหาโดยใช้คอมพิวเตอร์ (computerized content-based adaptive test) ซึ่งให้ความสำคัญกับความครอบคลุมเนื้อหาแต่ไม่ได้จัดลำดับชั้นของเนื้อหาจึงทำให้ต้องทดสอบทุกเนื้อหา ส่วนโปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะที่พัฒนาขึ้นมีการกำหนดลำดับชั้นของคุณลักษณะ ก่อนการพัฒนาโปรแกรมคอมพิวเตอร์ ดังนั้นในการทดสอบจึงไม่จำเป็นต้องทดสอบทุกเนื้อหา เนื่องจากคุณลักษณะที่กำหนดขึ้นมีความเกี่ยวข้องกับเนื้อหาที่มีความสัมพันธ์กันเป็นลำดับชั้น คุณลักษณะในลำดับชั้นแรกเป็นคุณลักษณะที่ต้องมีมาก่อนลำดับชั้นต่อไป กรณีที่ไม่ผ่านคุณลักษณะที่มีมาก่อนจึงสันนิษฐานได้ว่าคุณลักษณะในลำดับชั้นต่อไปก็จะเป็นเช่นนั้น

เมื่อพิจารณาด้านจำนวนข้อสอบในคลังข้อสอบจะเห็นได้ว่าจำนวนข้อสอบในคลังข้อสอบสำหรับโปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะที่พัฒนาขึ้นขึ้นอยู่กับจำนวนคุณลักษณะและจำนวนข้อสอบที่กำหนดสำหรับแต่ละชุดของข้อสอบ ซึ่งจากการดำเนินการพบว่าข้อสอบในคลังข้อสอบมีจำนวน 438 ข้อ เป็นไปตามข้อเสนอของ Weiss (1988, อ้างถึงใน รั้งสรรพ มณีเล็ก, 2540) ทั้งจำนวนข้อสอบในคลังข้อสอบทั้งหมดและจำนวนข้อสอบในคลังข้อสอบ จำแนกตามชุดของข้อสอบ ส่วนด้านค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบ พบว่ามีความสอดคล้องกับแผนภาพลำดับชั้นของคุณลักษณะ กล่าวคือ เมื่อจำแนกค่าพารามิเตอร์ตามชุดของข้อสอบจะเห็นได้ว่า ค่าความยากของข้อสอบคุณลักษณะที่ 4 มีค่ามากที่สุด รองลงมา คือคุณลักษณะที่ 3 คุณลักษณะที่ 2 และคุณลักษณะที่ 1 แสดงให้เห็นว่าข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ 1 ซึ่งเป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะอื่น ๆ เป็นข้อสอบที่ง่ายกว่าข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะอื่น ๆ ซึ่งเป็นคุณลักษณะในลำดับชั้นต่อไปและข้อสอบสำหรับคุณลักษณะที่ 4 ซึ่งเป็นคุณลักษณะที่อยู่ลำดับชั้นสูงสุดจะเห็นได้ว่าเป็นข้อสอบที่ยากกว่าข้อสอบในชุดของข้อสอบสำหรับคุณลักษณะที่ต้องมีมาก่อนอื่น ๆ ส่วนชุดของข้อสอบสำหรับคุณลักษณะที่ 2 และคุณลักษณะที่ 3 ซึ่งเป็นคุณลักษณะในลำดับชั้นเดียวกันจะเห็นได้ว่ามีค่าที่ใกล้เคียงกัน ทั้งนี้อาจเนื่องจากในขั้นของการกำหนดลำดับชั้นของคุณลักษณะมีการดำเนินการอย่างละเอียดและมีการตรวจสอบความถูกต้องทุกขั้นตอน ตั้งแต่ขั้นตอนการกำหนดคุณลักษณะ ขั้นตอนการจัดเรียงลำดับชั้นของคุณลักษณะและขั้นตอนการจัดทำแผนภาพลำดับชั้น

ของคุณลักษณะ โดยเฉพาะวิธีการที่ใช้ในการกำหนดคุณลักษณะจะเห็นได้ว่าใช้ทั้งการวิเคราะห์จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง (document analysis) จากการทำแบบทดสอบแบบเติมคำของนักเรียน (student-based analysis from protocol) และจากการวิเคราะห์จำแนกของครูผู้สอนวิชาคณิตศาสตร์ (expert-based analysis) ซึ่งเป็นไปตามผลการวิจัยของ Leighton, Cui and Cor (2008) ซึ่งเปรียบเทียบความเหมาะสมของโมเดลทางพุทธิปัญญาระหว่างโมเดลที่พัฒนาโดยนักเรียน (student-based model) จากการวิเคราะห์การรายงานด้วยถ้อยคำ (verbal analysis) และโมเดลที่พัฒนาโดยผู้เชี่ยวชาญ (expert-based model) จากการวิเคราะห์ข้อสอบโดยผู้ที่มีความเชี่ยวชาญในด้านสถิติ คณิตศาสตร์และการวัดและประเมินผล พบว่าทั้งโมเดลที่พัฒนาโดยนักเรียนและโมเดลที่พัฒนาโดยผู้เชี่ยวชาญสามารถทำนายการทดสอบการปฏิบัติงานของนักเรียนที่มีความสามารถปานกลางและสูงได้ดี แต่ต่างกันว่าโมเดลซึ่งพัฒนาโดยผู้เชี่ยวชาญทำนายกระบวนการตอบโดยทั่วไปได้ดีกว่า ส่วนโมเดลที่พัฒนาโดยนักเรียนทำนายระดับความสามารถเฉพาะได้ดีกว่า การใช้วิธีการที่หลากหลายสำหรับการกำหนดลำดับชั้นของคุณลักษณะในการพัฒนาวิธีการประเมินเชิงวินิจฉัยในครั้งนี้จึงส่งผลให้แผนภาพลำดับชั้นของคุณลักษณะมีความถูกต้องและความครอบคลุมดังที่กล่าวมาข้างต้น

สำหรับแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน พบว่า คุณลักษณะที่ 4 มีจำนวนแบบการคิดผิดที่จำแนกได้มากที่สุด (5 ข้อหลัก 26 ข้อย่อย) รองลงมา ได้แก่ คุณลักษณะที่ 3 คุณลักษณะที่ 2 และคุณลักษณะที่ 1 โดยมีจำนวนแบบการคิดที่ผิด จำนวน 5 ข้อหลัก 14 ข้อย่อย, 3 ข้อหลัก 8 ข้อย่อย และ 2 ข้อหลัก 4 ข้อย่อย ตามลำดับ จะเห็นได้ว่าจำนวนแบบการคิดที่ผิดที่จำแนกได้สอดคล้องกับแผนภาพลำดับชั้นของคุณลักษณะ กล่าวคือแบบการคิดที่ผิดจะมีจำนวนเพิ่มมากขึ้นตามลำดับจากคุณลักษณะที่ต้องมีมาก่อนไปถึงคุณลักษณะที่ต้องมีต่อ ๆ ไป ทั้งนี้อาจเป็นเพราะคุณลักษณะที่ต้องมีมาก่อนเป็นคุณลักษณะที่มีความง่ายมากกว่าและมีความซับซ้อนของความรู้ ทักษะและกระบวนการที่ใช้น้อยกว่าคุณลักษณะที่ต้องมีในลำดับชั้นต่อ ๆ ไป สอดคล้องกับเกณฑ์การสร้างลำดับชั้นเนื้อหาวิชาตามทฤษฎีการเรียนรู้ของกาญเย่ (Gagne' ,1965 และ 1967 อ้างถึงใน สุวัฒนา อุทัยรัตน์, สุชาวดี เขียมอรพรรณ และธนภรณ์ พิษณานนท์, 2541) ที่กล่าวว่าลำดับชั้นเนื้อหาวิชาที่อยู่ต่ำกว่าต้องมีการเรียนรู้ก่อนเพื่อใช้เป็นพื้นฐานในการเรียนรู้ลำดับชั้นเนื้อหาวิชาลำดับชั้นต่อไป โดย กมล ภูประเสริฐ (2520) ได้ให้ความหมายของความสัมพันธ์เชิงลำดับชั้นในเชิงวัดผล ไว้ว่า ขึ้นงานใด ๆ มีความสัมพันธ์เชิงลำดับชั้นกันเมื่อบุคคลใดสามารถปฏิบัติชิ้นงานเฉพาะที่ซับซ้อนกว่าได้แล้ว เขาต้องสามารถปฏิบัติชิ้นงานเฉพาะที่ง่ายกว่าได้ด้วยและบุคคลใดที่ปฏิบัติงานเฉพาะขั้นต่ำกว่าไม่ได้ย่อมไม่สามารถปฏิบัติชิ้นงาน

เฉพาะขั้นสูงกว่าได้ นอกจากนี้ยังสอดคล้องกับวิธีการจัดลำดับชั้นการเรียนรู้คณิตศาสตร์ของ Causey ซึ่งใช้ความยากง่ายของข้อคำถามในการจัดลำดับชั้นของการเรียนรู้ โดยเริ่มจากข้อที่ง่ายที่สุดไปข้อที่ยากและซับซ้อนที่สุด (Causey, 1977 อ้างถึงใน จินดา ลิ้มถาวรศิริพงศ์, 2525)

## 2. คุณภาพของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น

เมื่อพิจารณาผลการศึกษาคคุณภาพของวิธีการประเมินเชิงวินิจฉัยในภาพรวมจะเห็นได้ว่าวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นให้ผลการวินิจฉัยที่มีความเที่ยงและความตรงสูง รวมทั้งมีค่าเฉลี่ยของจำนวนข้อสอบและเวลาที่ใช้ในการทดสอบไม่มาก โดยใช้เวลาเฉลี่ยข้อละ 25 วินาที ( $(17 \times 60) / 40.7$ ) แสดงให้เห็นว่าวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นมีคุณภาพเป็นไปตามวัตถุประสงค์ที่ตั้งไว้สอดคล้องกับผลการวิจัยของ นันทิยา พึ่งคำ (2531), สุพัฒน์ สุขมลสันต์ (2536), ชูเกียรติ ละอองแก้ว (2538), ชูใจ ทรงเมฆ (2538), อำไพ วรณเกษม (2539), ละมัย พรหมศร (2540), นุจรี สุภิสุทธิ์ (2541), พงษ์พิชิต พรหมสิทธิ์ (2544), Mislevy (1999), Drasgow and Buchana (1999 อ้างถึงใน Hol, Vorst and Mellenbergh, 2005) และ Weiss (2006) ซึ่งพบว่าการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มีประสิทธิภาพมากกว่าการทดสอบแบบดั้งเดิมหรือการทดสอบแบบเขียนตอบบนกระดาษคำตอบ โดยมีความค่าความเที่ยงและความตรงสูง รวมทั้งช่วยลดจำนวนข้อสอบและเวลาในการทดสอบ

เมื่อพิจารณาค่าการกระจายของข้อมูล พบว่า มีค่าพิสัยและค่าเบี่ยงเบนมาตรฐานมาก ทั้งจำนวนข้อสอบและเวลาที่ใช้ในการทดสอบ ซึ่งสอดคล้องกับค่าการกระจายของค่าความยากของข้อสอบในคลังและค่าการกระจายของระดับความสามารถของผู้สอบ แสดงให้เห็นว่าผู้สอบใช้จำนวนข้อสอบและเวลาในการทดสอบมากขึ้นแตกต่างกันตามค่าความยากของข้อสอบและระดับความสามารถของผู้สอบ สอดคล้องกับผลการวิจัยของ Bridgeman and Cline (2004) ซึ่งพบว่าเวลาที่ใช้ในการทดสอบมีความสัมพันธ์กับค่าความยากของข้อสอบและผลการวิจัยของ เกียรติศักดิ์ สองแสง (2547) ซึ่งพบว่าการกำหนดเวลาในการทำข้อสอบแต่ละข้อตามระดับความยากของข้อสอบทำให้จำนวนข้อสอบและเวลาที่ใช้ในการทดสอบน้อยที่สุด

เมื่อพิจารณาผลการเปรียบเทียบคุณภาพของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น จำแนกตามข้อกำหนดในการตัดสินความรอบรู้ที่แตกต่างจากการเปรียบเทียบค่าความเที่ยง ค่าความตรง จำนวนข้อสอบและเวลาที่ใช้ในการทดสอบจำแนกตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้ พบว่า เกณฑ์ที่ใช้ในการตัดสินความรอบรู้ทั้ง 6 เกณฑ์ ส่วนใหญ่มีคุณภาพของผลการประเมินเชิงวินิจฉัยจากวิธีการที่พัฒนาขึ้นไม่แตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ยกเว้นด้านเวลาที่ใช้ในการทดสอบ พบว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

สำหรับเกณฑ์ที่ 1 และเกณฑ์ที่ 3 แสดงให้เห็นว่าการกำหนดจำนวนข้อสอบที่ต้องทำ จำนวนข้อสอบที่ตอบถูกและจำนวนแบบการคิดที่ผิดที่แตกต่างกันมีผลต่อเวลาที่ใช้ในการทดสอบที่แตกต่างกัน โดยเฉพาะเมื่อกำหนดจำนวนข้อสอบแตกต่างกัน 2 เท่า (2 ใน 3 ข้อ และ 4 ใน 5 ข้อ) แต่เมื่อใช้ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าร่วมกับการกำหนดจำนวนข้อสอบกลับพบว่าไม่มีผลต่อเวลาที่ใช้ในการทดสอบจะเห็นได้จากนักเรียนที่ใช้เกณฑ์ในการตัดสินความรอบรู้เกณฑ์ที่ 4 เกณฑ์ที่ 5 และเกณฑ์ที่ 6 ใช้เวลาในการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ทั้งนี้เนื่องจากการคำนวณค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าขึ้นอยู่กับค่าสารสนเทศที่ได้รับจากแบบทดสอบสำหรับผู้สอบแต่ละระดับความสามารถ จำนวนข้อสอบจึงไม่ได้เป็นปัจจัยเดียวที่มีผลต่อเวลาที่ใช้ในการทดสอบ แต่รวมถึงปัจจัยที่เป็นระดับความสามารถของผู้สอบด้วย ซึ่งเมื่อเปรียบเทียบค่าเฉลี่ยของระดับความสามารถของผู้สอบจำแนกตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้พบว่าไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ( $F = 2.187$ ) เมื่อพิจารณาค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพบว่าจะมีค่าลดลงเมื่อมีจำนวนข้อสอบเพิ่มขึ้นสอดคล้องกับผลการวิจัยของ รังสรรค์ มณีเล็ก (2540) ที่พบว่าจำนวนข้อสอบที่ใช้ในการทดสอบเพิ่มขึ้นถ้ากำหนดค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าลดลงและผลการวิจัยของเกียรติศักดิ์ สองแสง (2547) ที่พบว่ากำหนดจำนวนข้อสอบที่ให้ทำเพิ่มขึ้นส่งผลให้สารสนเทศของแบบทดสอบเพิ่มขึ้นซึ่งแสดงให้เห็นว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าจะลดลง เนื่องจากค่าสารสนเทศของแบบทดสอบแปรผกผันกับค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า

### **ข้อเสนอแนะเพื่อการพัฒนาการจัดการเรียนรู้เรื่องการवलบเศษส่วน**

ผลการประเมินเชิงวินิจฉัยแบบการคิดที่ผิดที่ใช้ในการवलบเศษส่วน ระดับการศึกษาขั้นพื้นฐานช่วงชั้นที่ 2 (ชั้นประถมศึกษาที่ 4 - 6) พบว่า นักเรียนส่วนใหญ่มีแบบแผนการตอบข้อสอบที่สังเกตได้สอดคล้องกับแบบแผนการตอบข้อสอบที่คาดหวังซึ่งแสดงให้เห็นว่ามีความรู้ทักษะหรือกระบวนการที่ใช้ในการवलบเศษส่วนเพียงบางส่วน โดยเฉพาะการवलบเศษส่วนแท้หรือจำนวนคละที่มีตัวส่วนเท่ากันเท่านั้น ซึ่งจากผลการประเมินเชิงวินิจฉัยแบบการคิดที่ผิดจะเห็นได้ว่านักเรียนส่วนใหญ่ไม่เข้าใจวิธีการवलบเศษส่วนแท้ ทั้งที่มีตัวส่วนเท่ากันและมีตัวส่วนไม่เท่ากัน โดยใช้วิธีเศษवलบเศษ ตัวส่วนवलบตัวส่วนมากที่สุด เมื่อพิจารณาตามลำดับขั้นของคุณลักษณะที่กำหนดได้จะเห็นได้ว่าการवलบเศษส่วนแท้ที่มีตัวส่วนเท่ากันเป็นคุณลักษณะขั้นพื้นฐานที่ต้องมีมาก่อนการवलบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน และการवलบจำนวนคละ

ทั้งที่มีตัวส่วนเท่ากันและที่มีตัวส่วนไม่เท่ากัน ดังนั้นเพื่อพัฒนาการจัดการเรียนรู้เรื่องการบวกลบเศษส่วนให้เกิดประโยชน์สูงสุด ครูผู้สอนควรวางแผนการจัดการเรียนรู้เพื่อให้ผู้เรียนมีความรู้ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากันก่อนการจัดการเรียนรู้เรื่องการบวกลบเศษส่วนเรื่องอื่น ๆ โดยเฉพาะการสร้างความรู้ความเข้าใจเรื่องวิธีการบวกลบเศษส่วนแท้ ทั้งที่มีตัวส่วนเท่ากันและที่มีตัวส่วนไม่เท่ากัน ทั้งนี้ครูผู้สอนอาจสร้างแบบทดสอบเรื่องการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากันและการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากันเพื่อวัดความรู้พื้นฐานและจำแนกแบบการคิดที่ผิดของนักเรียนก่อนจัดการเรียนรู้เพื่อให้การจัดการเรียนรู้สอดคล้องกับสภาพและความต้องการของนักเรียนมากที่สุดซึ่งจะช่วยให้ นักเรียนเกิดการเรียนรู้และการพัฒนาได้อย่างเต็มตามศักยภาพต่อไป

### ข้อเสนอแนะเพื่อนำวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นไปใช้

วิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นเป็นการประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะซึ่งเป็นโมเดลการวินิจฉัยทางพุทธิปัญญา (Cognitive Diagnostic Model: CDM) ในการพัฒนาแบบทดสอบและการแปลผลการทดสอบเพื่อให้แบบทดสอบและผลการทดสอบมีความหมายมากขึ้นจากการอธิบายได้ด้วยกระบวนการทางพุทธิปัญญา นอกจากนี้ยังใช้การทดสอบแบบปรับเหมาะซึ่งเป็นการประยุกต์ใช้ทฤษฎีการทดสอบแนวใหม่ (Modern Test theory) ในกระบวนการทดสอบเพื่อให้การบริหารจัดการการทดสอบมีประสิทธิภาพยิ่งขึ้น เนื้อหาที่ใช้เป็นเนื้อหาวิชาคณิตศาสตร์ เรื่องการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐานช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) ผลผลิตที่ได้เป็นโปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะจำนวน 6 โปรแกรม ซึ่งมีลักษณะเฉพาะทั้งสารสนเทศในคลังข้อสอบและขั้นตอนการคัดเลือกข้อสอบจากคลังข้อสอบ การนำวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นไปใช้ดำเนินการได้ 2 แนวทางแนวทางแรก คือ การนำผลผลิตที่ได้ไปใช้ ในกรณีที่ต้องการประเมินเชิงวินิจฉัยคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐานช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) ซึ่งเป็นเนื้อหาวิชาคณิตศาสตร์ ส่วนแนวทางที่ 2 คือ การนำแนวคิดและกระบวนการในการพัฒนาวิธีการประเมินเชิงวินิจฉัยไปใช้ กรณีที่ต้องการประเมินเชิงวินิจฉัยคุณลักษณะในเนื้อหาอื่น ๆ มีข้อเสนอแนะในการนำวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นไปใช้ในแต่ละแนวทาง ดังนี้

1. การนำวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นไปใช้เพื่อประเมินคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐานช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) อาจดำเนินการได้ทั้งการใช้คอมพิวเตอร์และไม่ใช้คอมพิวเตอร์ การดำเนินการโดยใช้คอมพิวเตอร์

สามารถเลือกโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะโปรแกรมใดโปรแกรมหนึ่ง ซึ่งแตกต่างกันตามเกณฑ์ที่ใช้ในการตัดสินความรอบรู้ไปใช้ได้ เนื่องจากผลการวิจัยพบว่าทั้ง 6 เกณฑ์ให้ผลการประเมินเชิงวินิจฉัยที่มีคุณภาพไม่แตกต่างกัน แต่เพื่อให้เกิดประโยชน์และความคุ้มค่าสูงสุดควรเลือกใช้เกณฑ์ที่ 1 หรือเกณฑ์ที่ 4 ซึ่งกำหนดจำนวนข้อสอบที่ทำให้ทำน้อยกว่าเกณฑ์อื่น ๆ โดยเกณฑ์ที่ 4 จะให้ความมั่นใจในผลการประเมินมากกว่าเกณฑ์ที่ 1 จากการกำหนดค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3 ร่วมด้วย ส่วนการดำเนินการโดยไม่ใช้คอมพิวเตอร์สามารถดำเนินการได้ 2 แนวทาง แนวทางแรก คือ จัดทำแบบทดสอบแบบเขียนตอบ สำหรับแต่ละคุณลักษณะโดยการคัดเลือกข้อสอบจากคลังข้อสอบ กำหนดเกณฑ์การผ่านและเกณฑ์การหยุดสอบสำหรับแต่ละคุณลักษณะ หลังจากนั้นจึงนำไปให้นักเรียนทำทีละข้อ เมื่อนักเรียนมีผลการสอบตามเกณฑ์การตัดสินความรอบรู้ที่กำหนด จึงให้นักเรียนทำข้อสอบในลำดับขั้นต่อไป จนกระทั่งถึงเกณฑ์การหยุดสอบจึงสรุปสถานะความรู้ตามแบบแผนการตอบข้อสอบที่คาดหวังและสรุปแบบการคิดที่ผิดที่ใช้ พร้อมทั้งให้ข้อมูลย้อนกลับแก่ผู้ที่เกี่ยวข้อง แนวทางที่ 2 คือ ดำเนินการประเมินเชิงวินิจฉัยโดยใช้เทคนิคการคิดออกเสียงตามแนวทางการดำเนินการในคู่มือการประเมินเชิงวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน โดยใช้เทคนิคการคิดออกเสียง มีข้อเสนอแนะในการนำวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นไปใช้ ดังนี้

1.1 วิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นสามารถนำไปใช้ในการประเมินเชิงวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนของนักเรียนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้น 2 (ชั้นประถมศึกษาปีที่ 4-6) ได้ทุกระดับชั้น เนื่องจากลักษณะข้อสอบมีความครอบคลุมความรู้ ทักษะและกระบวนการที่ใช้ในการบวกลบเศษส่วนในระดับชั้นนั้น ๆ

1.2 เพื่อให้ผลการประเมินมีความถูกต้องแม่นยำผู้ได้รับการประเมินควรผ่านการเรียนรู้เรื่องการบวกลบเศษส่วนมาแล้ว

1.3 วิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นมีข้อจำกัดที่ไม่สามารถวินิจฉัยคำตอบผิดซึ่งเกิดจากความไม่รอบคอบของผู้สอบได้ ดังนั้นเพื่อให้ผลที่ได้จากการประเมินเชิงวินิจฉัยเกิดประโยชน์สูงสุด ผู้ประเมินควรวิเคราะห์ผลการตอบข้อสอบรายบุคคล กรณีที่วิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นให้ผลการประเมินว่า “ไม่สามารถวินิจฉัยได้”

1.4 วิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้นให้ผลการประเมินเชิงวินิจฉัยทั้งสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนซึ่งเป็นประโยชน์ต่อการพัฒนาการจัดการเรียนรู้และการพัฒนาผู้เรียนมากกว่าการตัดสินระดับความสามารถของผู้เรียน ดังนั้นวิธีการ

ประเมินเชิงวิจักษ์ที่พัฒนาขึ้นจึงเหมาะสำหรับการประเมินเพื่อการพัฒนา (formative assessment) มากกว่าการประเมินผลสรุป (summative assessment)

1.5 แม้ว่าวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้นจะมีวัตถุประสงค์เพื่อใช้ในการประเมิน แต่จากการนำไปใช้พบว่าสามารถประยุกต์ใช้ระหว่างจัดการเรียนรู้เพื่อสร้างและตรวจสอบความคิดรวบยอดและการใช้กระบวนการทางคณิตศาสตร์ในการดำเนินการบวกลบเศษส่วนของนักเรียนได้ โดยสามารถดำเนินการได้ทั้งรายบุคคลและรายกลุ่ม รวมถึงการประเมินหลังการสอนซ่อมเสริมด้วย

1.6 เพื่อตรวจสอบยืนยันผลการประเมิน ผู้ประเมินอาจดำเนินการประเมินซ้ำ ด้วยวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้นหรือวิธีการประเมินที่เชื่อถือได้วิธีการอื่น

1.7 ในการนำผลการประเมินจากวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้นไปใช้ควรคำนึงถึงการให้ข้อมูลย้อนกลับแก่ผู้ที่เกี่ยวข้องอย่างเหมาะสมและเกิดประโยชน์สูงสุด

1.8 ผู้ที่จะนำวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้นไปใช้ อาจเพิ่ม ลดหรือแก้ไขปรับปรุงข้อสอบในคลังข้อสอบได้ตามหลักการแนวทางการจัดทำคลังข้อสอบ

2. การนำวิธีการประเมินเชิงวิจักษ์ที่พัฒนาขึ้นไปใช้เพื่อประเมินคุณลักษณะในเนื้อหาอื่น ๆ อาจเป็นวิชาคณิตศาสตร์หรือวิชาอื่น ๆ แต่ควรเป็นคุณลักษณะที่สามารถกำหนดลำดับชั้นของคุณลักษณะและแบบการคิดที่ผิดได้อย่างชัดเจน ทั้งนี้เพื่อให้การพัฒนาข้อสอบ การดำเนินการทดสอบและการกำหนดผลการประเมินเชิงวิจักษ์จากผลการทดสอบมีความถูกต้องโดยมีขั้นตอนการดำเนินการ ดังนี้

*ขั้นตอนที่ 1 การพัฒนาเครื่องมือการประเมิน มีแนวทางการดำเนินการ ดังนี้*

1) กำหนดลำดับชั้นของคุณลักษณะ เริ่มจากการสำรวจคุณลักษณะที่ใช้ที่เป็นไปได้ทั้งหมดสำหรับเนื้อหานั้น ๆ อาจใช้การวิเคราะห์และสังเคราะห์จากเอกสารที่เกี่ยวข้อง จากผลการตอบแบบทดสอบของนักเรียนหรือจากผู้เชี่ยวชาญด้านเนื้อหา หลังจากนั้นจึงนำคุณลักษณะที่ได้มาจัดเรียงลำดับชั้นของคุณลักษณะโดยพิจารณาว่าคุณลักษณะใดเป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะใด คุณลักษณะเหล่านั้นมีความสัมพันธ์เชิงลำดับชั้นแบบใด แล้วจึงนำมาจัดเรียงจากคุณลักษณะขั้นพื้นฐานที่ต้องมีมาก่อนทุกคุณลักษณะไปยังคุณลักษณะขั้นสูงสุดที่ไม่ต้องมีมาก่อนคุณลักษณะใดเลยตามลักษณะความสัมพันธ์เชิงลำดับชั้นที่กำหนดได้ เพื่อให้ลำดับชั้นของคุณลักษณะมีความถูกต้อง ควรมีการตรวจสอบความครอบคลุม ครบถ้วนและความถูกต้องของคุณลักษณะทั้งในขั้นตอนการสำรวจและการจัดเรียงลำดับชั้นของคุณลักษณะ ทั้งนี้อาจแสดงลำดับชั้นของคุณลักษณะเป็นแผนภาพความสัมพันธ์เชิงลำดับชั้นเพื่อให้ง่ายต่อการทำความเข้าใจ



2) พัฒนาค้างข้อสอบ เริ่มจากการเขียนเมทริกซ์แทนความสัมพันธ์เชิงลำดับชั้นของคุณลักษณะตามที่กำหนดได้ (รายละเอียดเมทริกซ์ต่าง ๆ ศึกษาได้จากการดำเนินการวิจัย บทที่ 3) ทั้งนี้เพื่อกำหนดลักษณะเฉพาะของข้อสอบ แบบแผนการตอบข้อสอบที่คาดหวัง ลำดับชั้นการคัดเลือกข้อสอบและแนวทางการวินิจฉัยผลที่ได้จากการทดสอบ หลังจากนั้นจึงเขียนข้อสอบเพื่อใช้ประเมินในแต่ละคุณลักษณะ ประกอบด้วยข้อคำถาม คำตอบถูก คำตอบผิดสำหรับแต่ละแบบการคิดที่ผิดกรณีที่ต้องการวินิจฉัยแบบการคิดที่ผิด (แนวทางการกำหนดแบบการคิดที่ผิดศึกษาได้จากการดำเนินการวิจัย บทที่ 3) นำข้อสอบที่ได้ไปทดลองใช้เพื่อหาคุณภาพของข้อสอบ คัดเลือกข้อสอบที่มีคุณภาพตามเกณฑ์ที่กำหนดสำหรับแต่ละคุณลักษณะ

3) จัดทำแบบทดสอบหรือโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะ การจัดทำแบบทดสอบควรมีลักษณะและจำนวนข้อสอบเป็นไปตามหลักการการวัดและประเมินผล ส่วนการจัดทำโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะมีแนวทางการดำเนินการ เหมือนกับการจัดทำโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะทั่วไป แต่ต่างกันที่การคัดเลือกข้อสอบซึ่งเป็นไปตามลักษณะความสัมพันธ์เชิงลำดับชั้นที่กำหนดได้

### *ขั้นตอนที่ 2 การดำเนินการประเมิน*

การประเมินอาจดำเนินการโดยใช้คอมพิวเตอร์หรือไม่ใช้คอมพิวเตอร์ การดำเนินการโดยใช้คอมพิวเตอร์ดำเนินการโดยใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะที่พัฒนาขึ้นเป็นเครื่องมือในการประเมิน ส่วนการดำเนินการโดยไม่ใช้คอมพิวเตอร์มีแนวทางและวิธีการดำเนินการประเมินเช่นเดียวกับการประเมินเชิงวินิจฉัยโดยใช้วิธีการที่พัฒนาขึ้นในการประเมินเชิงวินิจฉัยคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐานช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) ในข้อ 1

### *ขั้นตอนที่ 3 การแปลผลการประเมิน*

การแปลผลการประเมินดำเนินการโดยนำผลการทดสอบที่ได้มาสรุปเป็นแบบแผนการตอบข้อสอบที่สังเกตได้และนำไปตรวจสอบความสอดคล้องกับแบบแผนการตอบข้อสอบที่คาดหวังซึ่งกำหนดได้ในขั้นตอนการพัฒนาค้างข้อสอบ หลังจากนั้นจึงแปลผลการประเมินตามแบบแผนการตอบข้อสอบที่คาดหวังนั้น ๆ กรณีที่มีการวินิจฉัยแบบการคิดที่ผิดให้ตรวจสอบความสอดคล้องระหว่างแบบการคิดที่ผิดที่นักเรียนใช้ตามเกณฑ์ที่กำหนดกับแบบการคิดที่ผิดที่กำหนดได้หลังจากนั้นจึงแปลผลจากแบบการคิดที่ผิดที่สอดคล้องกัน

## ข้อเสนอแนะเพื่อการวิจัยต่อไป

1. การประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะในการประเมินเชิงวิจจัยทางการศึกษาช่วยให้การพัฒนาแบบทดสอบและการแปลผลการทดสอบมีความชัดเจนมากขึ้น ซึ่งจากการดำเนินการวิจัยครั้งนี้พบว่าขั้นตอนการประยุกต์ใช้ที่สำคัญ ก็คือ การกำหนดและจัดเรียงลำดับขั้นของคุณลักษณะ ดังนั้นเพื่อให้การกำหนดและจัดเรียงลำดับขั้นของคุณลักษณะเกิดประโยชน์สูงสุด จึงควรพัฒนาวิธีการกำหนดและจัดเรียงลำดับขั้นของคุณลักษณะที่มีคุณภาพหรือศึกษาเปรียบเทียบคุณภาพเมื่อใช้วิธีการกำหนดและจัดเรียงลำดับขั้นของคุณลักษณะที่แตกต่างกัน

2. การวิจัยครั้งนี้กำหนดเกณฑ์การตัดสินความรอบรู้ในแต่ละลำดับขั้นจากจำนวนข้อสอบที่ทำได้ถูกต้องหรือไม่ถูกต้อง รวมทั้งระดับของความคลาดเคลื่อนมาตรฐานในการประมาณค่า ซึ่งผลการวิจัยพบว่าส่วนใหญ่ส่งผลต่อคุณภาพของการประเมินเชิงวิจจัยไม่แตกต่างกัน และเมื่อศึกษาจำแนกตามระดับความสามารถของผู้สอบพบว่าผู้สอบที่มีระดับความสามารถสูงและต่ำมีผลการประเมินเชิงวิจจัยที่มีคุณภาพสูงกว่าผู้สอบที่มีระดับความสามารถปานกลาง ทั้งด้านความเที่ยง ความตรง จำนวนเวลาและจำนวนข้อสอบที่ใช้ ไม่ว่าจะใช้เกณฑ์การตัดสินความรอบรู้เกณฑ์ใด แต่การวิจัยครั้งนี้ไม่ได้มีวัตถุประสงค์เพื่อศึกษาเปรียบเทียบคุณภาพของการประเมินเชิงวิจจัยเมื่อผู้สอบมีระดับความสามารถที่แตกต่างกันจึงทำให้ไม่สามารถสรุปข้อค้นพบได้อย่างชัดเจน ดังนั้นเพื่อให้ได้ข้อสรุปเกี่ยวกับผลการศึกษากฎเกณฑ์การตัดสินความรอบรู้ที่ชัดเจนขึ้นจึงควรศึกษาเปรียบเทียบคุณภาพของการประเมินเชิงวิจจัยเมื่อใช้เกณฑ์การตัดสินความรอบรู้ที่แตกต่างกันในแต่ละระดับความสามารถของผู้สอบที่แตกต่างกัน

3. จากการพัฒนาโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะพบว่าใช้เวลาในการพัฒนาโปรแกรมคอมพิวเตอร์ค่อนข้างมาก แต่โปรแกรมคอมพิวเตอร์ที่พัฒนาขึ้นมีข้อจำกัดที่สามารถใช้ได้กับลำดับขั้นของคุณลักษณะที่กำหนดได้เฉพาะการวิจัยครั้งนี้เท่านั้น ดังนั้นเพื่อให้การพัฒนาโปรแกรมคอมพิวเตอร์ดังกล่าวมีความคุ้มค่าและเกิดประโยชน์สูงสุดจึงควรพัฒนาโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบแบบปรับเหมาะซึ่งสามารถรองรับการกำหนดลำดับขั้นของคุณลักษณะที่หลากหลายได้

## รายการอ้างอิง

### ภาษาไทย

- กนกวรรณ รัตนธน. (2544). การเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์ เมื่อระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกัน. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชามัธยมศึกษา. คณะครุศาสตร์. จุฬาลงกรณ์มหาวิทยาลัย.
- กมล ภูประเสริฐ. (2520). การเปรียบเทียบวิธีการสองวิธีในการทดสอบความเที่ยงตรงของลำดับชั้นการเรียนรู้. วิทยานิพนธ์ปริญญาโทศึกษาศาสตร์. ภาควิชาการวัดผลการศึกษา. บัณฑิตวิทยาลัย. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- เกรียงศักดิ์ สุวรรณภาค. (2532). ความสัมพันธ์ระหว่างรูปแบบของแบบทดสอบแยกกลุ่มกับความสามารถทางการเรียนในวิชาคณิตศาสตร์. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชาการวัดผลการศึกษา. บัณฑิตวิทยาลัย. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ขรรค์ชัย คงเสน่ห์. (2530). การสร้างแบบทดสอบสองขั้นตอน โดยใช้ทฤษฎีการตอบข้อคำถาม. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชาการวัดผลการศึกษา. บัณฑิตวิทยาลัย. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- จินดา ลิ้มถาวรศิริพงศ์. (2525). การสร้างลำดับชั้นเนื้อหาวิชาจากการวินิจฉัยการเรียนรู้คณิตศาสตร์เรื่องจำนวนเต็ม ของนักเรียนชั้นมัธยมศึกษาปีที่ 2. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชามัธยมศึกษา. บัณฑิตวิทยาลัย. จุฬาลงกรณ์มหาวิทยาลัย.
- จิราพร ไกรสรศิริเวท. (2529). ความสัมพันธ์ระหว่างรูปแบบบางประการของแบบทดสอบรูปปริมาตรกับความสามารถทางการเรียนในวิชาคณิตศาสตร์. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชาการวัดผลการศึกษา. บัณฑิตวิทยาลัย. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ชูเกียรติ ลอยงแก้ง. (2538). การสร้างแบบทดสอบสองขั้นตอนด้วยคอมพิวเตอร์ ในวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 5. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชาการบริหารการศึกษา. บัณฑิตวิทยาลัย. มหาวิทยาลัยขอนแก่น.
- ชูใจ ทรงเมฆ. (2538). การเปรียบเทียบผลการประมาณค่าความสามารถด้วยวิธีการทดสอบแบบเทเลอร์และวิธีการทดสอบแบบดั้งเดิม. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชาวิจัยและประเมินผลการศึกษา. บัณฑิตวิทยาลัย. มหาวิทยาลัยเกษตรศาสตร์.

ดวงเดือน อ่อนนวม. (2533). การสอนซ่อมเสริมคณิตศาสตร์. กรุงเทพมหานคร: โรงพิมพ์  
จุฬาลงกรณ์มหาวิทยาลัย.

ตรึงใจ พูลผลอำนวย. (2534). การพัฒนาแบบสอบเฉพาะบุคคลในวิชาวิทยาศาสตร์สำหรับ  
นักเรียน ชั้นมัธยมศึกษาปีที่ 3. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต.  
ภาควิชาวิจัยการศึกษา. บัณฑิตวิทยาลัย. จุฬาลงกรณ์มหาวิทยาลัย.

ต่าย เชิญฉวี. (2534). การศึกษาเปรียบเทียบประสิทธิภาพในการประมาณค่าความสามารถของ  
ผู้สอบจากการทดสอบเทเลอรรูปปิรามิดที่มีรูปแบบ จำนวนชั้นและวิธีการให้คะแนนที่  
แตกต่างกัน โดยใช้วิธีมอนติคาร์โล. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต.

ภาควิชาการวิจัยและพัฒนาหลักสูตร. บัณฑิตวิทยาลัย. มหาวิทยาลัยศรีนครินทรวิโรฒ.

นันทิยา พึ่งคำ. (2531). การเปรียบเทียบคุณภาพการทดสอบแบบ ซี เอ ทีและแบบประเพณีนิยม  
ในการวัดความสามารถด้านคำศัพท์ภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3.  
วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชาวิจัยการศึกษา. บัณฑิตวิทยาลัย.  
จุฬาลงกรณ์มหาวิทยาลัย.

นุจรี สุภิสุทธิ์. (2541). การสร้างชุดทดสอบแบบเทเลอรรูปและดำเนินการสอบโดยใช้ไมโครคอมพิวเตอร์  
ในวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต.

ภาควิชาการวัดและประเมินผลการศึกษา. บัณฑิตวิทยาลัย. มหาวิทยาลัยเชียงใหม่.

ปราโมทย์ แสนหาญ. (2542). การสร้างแบบทดสอบปรับระดับชั้นที่ดำเนินการสอบด้วยคอมพิวเตอร์  
วิชาคณิตศาสตร์ เรื่อง จำนวนเชิงซ้อน ชั้นมัธยมศึกษาปีที่ 5. วิทยานิพนธ์ปริญญา  
มหาบัณฑิต. ภาควิชาการวัดผลการศึกษา. บัณฑิตวิทยาลัย. มหาวิทยาลัยมหาสารคาม.

ปราริษฐ์ นามคุณ. (2541). การเปรียบเทียบความเที่ยงตรงในการจำแนกความรอบรู้และการประมาณ  
ค่าความสามารถของผู้สอบ ในวิชาคณิตศาสตร์ของแบบทดสอบอิงเกณฑ์และแบบทดสอบ  
ปิรามิดบางรูปแบบ. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชาการวัดผลการศึกษา.  
บัณฑิตวิทยาลัย. มหาวิทยาลัยมหาสารคาม.

ปิยาพร ขาดสะอาด. (2541). การพัฒนากระบวนการสอบความก้าวหน้าวิชาภาษาอังกฤษโดยใช้  
คอมพิวเตอร์ที่ให้ผลการวินิจฉัยย้อนกลับ. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต.  
ภาควิชาวิจัยการศึกษา. คณะครุศาสตร์. จุฬาลงกรณ์มหาวิทยาลัย.

พงษ์พิชิต พรหมสิทธิ์. (2544). การดำเนินการสอบแบบทดสอบเทเลอรรูป ตามยุทธวิธีของเบย์  
ในวิชาคณิตศาสตร์ ค 014 โดยเว็บเพจ. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชาการ  
วัดและประเมินผลการศึกษา. บัณฑิตวิทยาลัย. มหาวิทยาลัยขอนแก่น.

- พินิจา สังข์เพชร. (2538). การพัฒนาดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบ  
รายบุคคล. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต. ภาควิชาวิจัยการศึกษา. บัณฑิตวิทยาลัย.  
จุฬาลงกรณ์มหาวิทยาลัย.
- ละมัย พรหมศร. (2540). การสร้างแบบทดสอบปิรามิดสำหรับใช้ทดสอบด้วยคอมพิวเตอร์ วิชา  
คณิตศาสตร์ เรื่องภาคตัดกรวย สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 4 จังหวัดนครพนม.  
วิทยานิพนธ์ปริญญาามหาบัณฑิต. ภาควิชาการวัดผลการศึกษา. บัณฑิตวิทยาลัย.  
มหาวิทยาลัยมหาสารคาม.
- วรรณวิภา จัตตุชัย. (2529). รูปแบบโครงสร้างของลำดับขั้นการเรียนรู้ด้านพุทธิพิสัยในผลสัมฤทธิ์  
วิชาคณิตศาสตร์ของนักเรียนระดับมัธยมศึกษาตอนต้น. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต.  
ภาควิชาวิจัยการศึกษา. บัณฑิตวิทยาลัย. จุฬาลงกรณ์มหาวิทยาลัย.
- วลี เฉลยสมัย. (2538). การพัฒนาวิธีการวินิจฉัยการแก้ไขโจทย์ปัญหาคณิตศาสตร์ที่คำนึงถึง  
สภาพที่เป็นจริงของนักเรียนชั้นประถมศึกษาปีที่ 6. วิทยานิพนธ์ปริญญาามหาบัณฑิต.  
ภาควิชาวิจัยการศึกษา. บัณฑิตวิทยาลัย. จุฬาลงกรณ์มหาวิทยาลัย.
- วสันต์ ทองไทย. (2539). การเปรียบเทียบผลการประมาณค่าความสามารถด้วยวิธีการทดสอบ  
แบบเทเลอร์รูปพีระมิดขนาดชั้นคงที่และรูปพีระมิดข้างตัด : การทดสอบทักษะ  
กระบวนการทางวิทยาศาสตร์. วิทยานิพนธ์ปริญญาามหาบัณฑิต. ภาควิชาวิจัยและ  
ประเมินผลการศึกษา. บัณฑิตวิทยาลัย. มหาวิทยาลัยเกษตรศาสตร์.
- วิชาการ,กรม. (2545). แนวทางการประเมินผลการเรียน ตามหลักสูตรการศึกษาขั้นพื้นฐาน  
พุทธศักราช 2544. กรุงเทพฯ: โรงพิมพ์คุรุสภา.
- ศลิษา สัตถาผล. (2542). การเปรียบเทียบแบบทดสอบสองขั้นตอนและแบบทดสอบสามขั้นตอนที่  
ดำเนินการสอบโดยใช้คอมพิวเตอร์ ในวิชาภาษาอังกฤษ ชั้นประถมศึกษาปีที่ 6.  
วิทยานิพนธ์ปริญญาามหาบัณฑิต. ภาควิชาการวัดและประเมินผลการศึกษา.  
บัณฑิตวิทยาลัย. มหาวิทยาลัยขอนแก่น.
- ศิริชัย กาญจนวาสี. (2538). การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ. กรุงเทพฯ:  
ศูนย์ตำราและเอกสารทางวิชาการ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2545). ทฤษฎีการทดสอบแนวใหม่. พิมพ์ครั้งที่ 2. กรุงเทพฯ:  
โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

- ศิริเดช สุชีวะ. (2538). การพัฒนาวิธีวินิจฉัยสำหรับตรวจสอบมโนทัศน์ที่คลาดเคลื่อนทางคณิตศาสตร์. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต. ภาควิชาวิจัยการศึกษา. บัณฑิตวิทยาลัย. จุฬาลงกรณ์มหาวิทยาลัย.
- ศิริเดช สุชีวะ. (2538). การวินิจฉัยมโนทัศน์ที่คลาดเคลื่อน. กรุงเทพฯ: ศูนย์ทดสอบทางการศึกษา และศูนย์ตำราและเอกสารทางวิชาการ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- สงวน บัญปลุก. (2533). การศึกษาความเที่ยงตรงในการจำแนกความรอบรู้และประมาณค่าความสามารถ ในวิชาคณิตศาสตร์จากการใช้แบบทดสอบอิงเกณฑ์และแบบทดสอบแบบเทเลอร์บางรูปแบบ. วิทยานิพนธ์ปริญญามหาบัณฑิต. สาขาวิชาการวัดผลการศึกษา. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- สายชล อบทม. (2539). การพัฒนาโปรแกรมสำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์. วิทยานิพนธ์ปริญญามหาบัณฑิต. ภาควิชาวิจัยการศึกษา. บัณฑิตวิทยาลัย. จุฬาลงกรณ์มหาวิทยาลัย.
- สำราญ มีแจ้ง. (2533). การพัฒนาดัชนีชี้ความบกพร่องของผู้ตอบแบบทดสอบโดยการวิเคราะห์แบบแผนการตอบข้อกระทง. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต. ภาควิชาวิจัยการศึกษา. บัณฑิตวิทยาลัย. จุฬาลงกรณ์มหาวิทยาลัย.
- สุวรรณ สมพงศ์พาณิชย์. (2538). การวิเคราะห์ลำดับขั้นเนื้อหาวิชาคณิตศาสตร์เรื่อง ระบบจำนวนจริง ในหลักสูตรมัธยมศึกษาตอนปลาย. วิทยานิพนธ์ปริญญามหาบัณฑิต. ภาควิชามัธยมศึกษา. บัณฑิตวิทยาลัย. จุฬาลงกรณ์มหาวิทยาลัย.
- สุวัฒนา อุทัยรัตน์, สุชาวดี เอี่ยมอรพวรรณ และ ธนภรณ์ พิษณานนท์. (2541) รายงานการวิจัย เรื่องการวิเคราะห์ลำดับขั้นโครงสร้างเนื้อหาวิชาคณิตศาสตร์ในหลักสูตรมัธยมศึกษาตอนต้นเพื่อพัฒนาแบบฝึกหัดเสริมการเรียนรู้. คณะครุศาสตร์. จุฬาลงกรณ์มหาวิทยาลัย.
- สุวิทยา พรหมอ่อน. (2540). การสร้างแบบทดสอบเทเลอร์รูปปริมาตรชนิดขนาดขั้นคงที่ วิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6. วิทยานิพนธ์ปริญญามหาบัณฑิต. ภาควิชาการวัดผลการศึกษา. บัณฑิตวิทยาลัย. มหาวิทยาลัยนเรศวร.
- เสวี ปรมชวลิตโรจน์. (2531). การศึกษาความเที่ยงตรงในการจำแนกความรอบรู้และการประมาณค่าความสามารถ ในวิชาคณิตศาสตร์ของแบบทดสอบเฟล็กซีเลเวลและแบบทดสอบรูปปริมาตรที่ใช้ขนาดขั้นแปรผัน. วิทยานิพนธ์ปริญญามหาบัณฑิต. ภาควิชาการวัดผลการศึกษา. บัณฑิตวิทยาลัย. มหาวิทยาลัยศรีนครินทรวิโรฒ.

- อนุชิต กอศักดิ์. (2542). การสร้างแบบทดสอบเทเลอร์ ชนิดยุทธวิธีสองขั้นตอน วิชาคณิตศาสตร์ เรื่อง ฟังก์ชันเอกซ์โพเนนเชียลและฟังก์ชันลอการิทึม ชั้นมัธยมศึกษาปีที่ 5.  
 วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชาการวัดผลการศึกษา. บัณฑิตวิทยาลัย.  
 มหาวิทยาลัยมหาสารคาม.
- อมรรัตน์ คำแดง. (2539). การประยุกต์ใช้การวิเคราะห์แบบพอลท์ทรีสำหรับการวิเคราะห์สาเหตุในการประเมินความต้องการจำเป็นเพื่อการพัฒนาประสิทธิภาพการใช้ครู วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชาวิจัยการศึกษา. บัณฑิตวิทยาลัย. จุฬาลงกรณ์มหาวิทยาลัย.
- อาตุณ ช้างแก้วมณี. (2529). ความสัมพันธ์ระหว่างลำดับชั้นการเรียนรู้ด้านพุทธิพิสัยกับลำดับที่ของค่าพารามิเตอร์ที่ได้จากการประมาณตามโมเดลโลจิสติกในแบบสอบวิชาคณิตศาสตร์.  
 วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชาวิจัยการศึกษา. บัณฑิตวิทยาลัย.  
 จุฬาลงกรณ์มหาวิทยาลัย.
- อำไพ วรรณเกษม. (2539). การสร้างแบบทดสอบเฉพาะบุคคลแบบ ซี เอ ที. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชาการวัดและประเมินผลการศึกษา. บัณฑิตวิทยาลัย. มหาวิทยาลัยรามคำแหง
- ไอลดา คล้ายสำริด. (2538). การเปรียบเทียบความคลาดเคลื่อนในการประมาณค่า ประสิทธิภาพสัมพัทธ์ในการประมาณค่าและความสามารถของผู้สอบ ในแบบทดสอบวิชาคณิตศาสตร์ที่ใช้วิธีต่างกัน. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชาการวัดผลการศึกษา. บัณฑิตวิทยาลัย. มหาวิทยาลัยศรีนครินทรวิโรฒ

## ภาษาอังกฤษ

- Baron. D.; and Bernard, H. W. (1958). Evaluation techniques for classroom teachers.  
 New York: McGraw-Hill.
- Bejar, I. I. (1984). Educational diagnostic assessment. Journal of Educational Measurement, 2: 175-189.
- Bělohávek, R.; Vladimír, S. and Jirí, Z. (2004). Formal concept analysis with hierarchically ordered attributes[Online]. International Journal of General Systems, 33: 383-395. Abstract from: Academic Search Premier: ISSN: 0308-1079
- Brown, F. G. (1983). Educational and psychological testing. 3<sup>rd</sup> ed. New York: CBS College Publishing.

- Chen, C. H.; Khoo, L.; and Yan, W. (2002). A strategy for acquiring customer requirement patterns using laddering technique and ART2 neural network [Online]. Advanced Engineering Informatics, 16, 229-241. Abstract from: Academic Search Premier: ISSN: 1474-0346
- Cohen, R. J.; Swerdlik, M. E.; and Smith, D. K. (1992). Psychological testing and assessment: An introduction to tests & assessment. 2<sup>nd</sup> ed. California: Mayfield Publishing Company.
- Cui, Y. and Leighton, J.P. (2008). The Hierarchy Consistency Index: Evaluating Person Fit for Cognitive Diagnostic Assessment [Online]. Available from: <http://www.education.ualberta.ca/educ/psych/crame/files/AHM.pdf>. [2008, September 19]
- Cui, Y.; Leighton, J. P.; Zheng, Y. (2006). Simulation studies for evaluation the performance of the two classification methods in the AHM. Annual Meeting of The National Council on Measurement in Education (NCME), pp. 1-18. California.
- Cui, Y. et al. (2006). The hierarchy consistency index: A person-fit statistic for the attribute hierarchy method. Annual Meeting of The National Council on Measurement in Education (NCME), pp. 1-30. California.
- Gardner, W et al. Computerized adaptive measurement of depression : A simulation study [Online]. Available from: <http://pubmedcentral.nih.gov/articlerender.fcgi?artid=416483> . [2006, August 1]
- Gierl, M. J.; Cui, Y. and Hunka, S. (2007). Using Connectionist Models to Evaluate Examinees' Response Patterns on Tests: An Application of the Attribute Hierarchy Method to Assessment Engineering. Annual Meeting of The National Council on Measurement in Education (NCME), pp. 1-22. Chicago.
- Gierl, M. J.; Leighton, J. P. and Hunka, S. P. (2000). Exploring the logic of Tatsuoka's rule-space model for test development and analysis. Educational Measurement: Issue and Practice, 19: 34-44.



- Gierl, M. J.; Zheng, Y. and Cui, Y. (2008). Using the Attribute Hierarchy Method to Identify and Interpret Cognitive Skills that Produce Group Differences. Journal of Educational Measurement, 45: 65-89.
- Gipps, C. V. (1995). Beyond testing : Towards a theory of educational assessment. 2<sup>nd</sup> ed. London: Burgess Science.
- Good, C. V. (1973). Dictionary of education. 3<sup>rd</sup> ed. New York: McGraw-Hill.
- Gregory, R. J. (1992). Psychological testing: History, principle and applications. Boston: Allyn and Bacon.
- Guzmán, E. and Conejo, R. (2004). A model for student knowledge diagnosis through Adaptive e testing. In Jame, G; Lester, R. M. and Vicari, F (eds.), Intelligent Tutoring System: 7 th international c onferences, ITS 2004 , pp. 12 - 21. Maceio, Alagoas, Brazil: Springer.
- Hambleton, R. K. and Swaminathan, H. (1985). Item response theory principles and applications. Boston: Kluwer Nijhoff.
- Han, C. H. (2004). Tools for establishing dominance with linear partial information and attribute hierarchy [Online]. International Journal of Systems Science, 35, 123-133. Abstract from: Academic Search Premier: ISSN: 0020-7721
- Harnisch, D. L.; and Linn, R. L. (1981). Analysis of item response patterns: Questionable test data and dissimilar curriculum practices. Journal of Educational Measurement, 18: 133-146.
- Hopkins, C. D. and Antes, R. L. (1990). Classroom measurement and evaluation. 3<sup>rd</sup> ed. Illinois: F. E. PEACOCK.
- Latu, E. and Chapman, E. (2002). Computerized adaptive testing . British Journal of Educational Technology, 33: 619-622.
- Leighton, J. P.; Cui, Y. and Cor, M. K. (2008). Testing Expert-based and Student-based Cognitive Models: An Application of the Attribute Hierarchy Method and Hierarchy Consistency Index [Online]. Available from : <http://www.education.ualberta.ca/educ/psych/crame/files/AHM.pdf>. [2008, September 19]

- Leighton, J. P.; Gierl, M. J.; and Hunka, S. P. (2002). The attribute hierarchy model for Assessment. Annual Meeting of The National Council on Measurement in Education (NCME), pp. 1-50. Louisiana.
- Leighton, J. P.; Gierl, M. J.; and Hunka, S. P. (2004). The attribute hierarchy method for cognitive assessment: A variation on Tatsuoka' s rule-space approach. Journal of Education Measurement, 41: 205-237.
- Linacre, J. M. (1999). Individualized testing in the classroom. In Masters, G. N. and Keeves, J. P. (eds.), Advances in Measurement in Educational Research and Measurement, pp. 138-150. Amsterdam: PERGAMON.
- Lohman, D. (2002). Complex information processing and intelligence[Online]. Available from: [http://www.faculty.education.uiowa.edu/dlohman/pdf/complex\\_information.processing.pdf](http://www.faculty.education.uiowa.edu/dlohman/pdf/complex_information.processing.pdf) [2005, June 24]
- McGlohon, M.K.; Chang, H. H.; and Wills, J. T. (2003). Combining computer adaptive testing technology with cognitively diagnostic assessment [Online]. Available from: <http://www.psych.umn.edu/psylabs/catcentral/pdf%20files/mcg/04-01.pdf>. [2006, August 1]
- Nichols, P. D. (1994). A framework for developing cognitively diagnostic assessment. Review of Educational research, 64: 575-603.
- Noye, J. and Fritz, J. (2003) Computerized adaptive testing for intelligent tutoring systems. World Conference on E-Learning in Corp.,Govt.,Health.,&Higher Ed. 2003(1) [Online]. Available from: <http://dl.aace.org/13662> [2006, August 1]
- Sadesky, G. S.; and Gushta, M. G. (2004). Standard setting using the attribute hierarchy model [Online]. Available from: <http://www.education.ualberta.ca/educ/psych/crame/files/AHM.pdf>. [2005, June 17]
- Shermis, M. D. and Wolting, M. (1996). Computerized adaptive testing for reading placement and diagnostic assessment [Online]. Available from: <http://search.epnet.com/login.aspx?direct=true&db=aph&an=10250385> [2005, June 17]

- Tatsuoka, K. K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. Journal of Educational Measurement, 20: 345-354.
- Tatsuoka, K. K. (1984). Caution indices based on item response theory. Psychometrika, 49: 95-110.
- Tatsuoka, K. K. (2002). The rule-space model for educational measurement [Online]. Available from: <http://www.tc.columbia.edu/news/artical.htm?id=4271&tid=103> [2005, June 17]
- Tatsuoka, K. K.; Corter, J. E. and Tatsuoka, C. (2004). Patterns of diagnosed mathematical content and process skill in TIMSS-R across a sample of 20 countries. American Educational Research Journal, 41: 901-926.
- Tatsuoka, K. K.; and Tatsuoka, M. M. (1983). Spotting erroneous rules of operation by the individual consistency index. Journal of Educational Measurement, 20: 221-230.
- Tatsuoka, K. K.; and Tatsuoka, M. M. (1988). Rule space. Encyclopedia of Statistical Sciences 8: 217-220.
- Tatsuoka, K. K.; and Tatsuoka, M. M. (1997). Computerized cognitive diagnostic adaptive testing: Effect on remedial instruction as empirical validation. Journal of Educational Measurement, 34: 3-20.
- Tatsuoka, K. K.; Xin, T.; and Xu, Z. (2004). Linkage between teacher quality, student achievement and cognitive skills: A rule-space model. Studies in Educational Evaluation, 30: 205-223.
- Thissen, D. and Mislevy, R. J. (1990). Testing algorithms. In Wainer, H. (eds.), Computerized Adaptive Testing: A Primer, pp. 103-135. New Jersey: Lawrence Erlbaum Associates.
- Van der Linden, W. J. (1999). Computerized educational testing. In Masters, G. N. and Keeves, J. P. (eds.), Advances in Measurement in Educational Research and Measurement, pp. 138-150. Amsterdam: PERGAMON.
- Wainer, H. (1990). Introduction and history. In Wainer, H. (eds.), Computerized Adaptive Testing: A Primer, pp. 1-21. New Jersey: Lawrence Erlbaum Associate.

- Wang, C. and Gierl, M. J. (2007). Investigating the Cognitive Attributes Underlying Student Performance on the SAT Critical Reading Subtest: An Application of the Attribute Hierarchy Method. Annual Meeting of The National Council on Measurement in Education (NCME), pp. 1-34. Chicago.
- Wang, C.; Gierl, M. J. and Leighton, J. P. (2006). Investigating the cognitive attribute underlying student performance on a foreign language reading test: An application of the attribute hierarchy method. Annual Meeting of The National Council on Measurement in Education (NCME), pp. 1-23. California.
- Weiss, D. J. (1983). Introduction. In Weiss, D. J. (eds.), New horizons in testing : latent trait test theory and computerized adaptive testing, pp. 1-7. New York: ACADEMIC.
- Weiss, D. J. (2004). Computerized adaptive testing for effective and efficient measurement in counseling and education. Measurement and Evaluation in Counseling and Development, 37: 70-84.
- Weiss, D. J. and Schleisman, J. L. (1999). Adaptive testing. In Masters, G. N. and Keeves, J. P. (eds.), Advances in Measurement in Educational Research and Measurement, pp. 138-150. Amsterdam: PERGAMON.
- Wen, J. (2003). Application of the rule space model in computerized adaptive testing for diagnostic assessment. Doctoral Dissertation, The Chinese University of Hong Kong (People's Republic of China).
- Yan, D.; Almond, R.; and Mislevy, R. (2003). Empirical comparisons of cognitive diagnostic model [Online]. Available from: <http://www.ets.org/legal/copyright> [2005, June 17]



## ภาคผนวก ก

รายชื่องานวิจัยที่เกี่ยวกับการวิจัยทางการศึกษาและแหล่งที่มา

ตารางที่ 1 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการสร้างหรือพัฒนาแบบทดสอบเชิงวินิจจัย

ที่	ชื่อ - สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปี ที่ทำ
1	นงคราญ ตันตะละ	การสร้างแบบทดสอบวินิจจัยเรื่องสมการ ชั้นมัธยมศึกษาปีที่ 2	ม.เชียงใหม่	2534
2	ฤตินันท์ สมุทรทัย	การสร้างแบบทดสอบวินิจจัยการแก้โจทย์ปัญหาคณิตศาสตร์	ม.เชียงใหม่	2534
3	แสงทอง สุภา	การสร้างแบบทดสอบวินิจจัยเรื่องไฟฟ้า วิชาวิทยาศาสตร์ สำหรับชั้นมัธยมศึกษาปีที่ 3	ม.เชียงใหม่	2540
4	นันทนา สิงห์วัฒนาศิริ	การพัฒนาแบบทดสอบวินิจจัย วิชาคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1	จุฬาลงกรณ์มหาวิทยาลัย	2534
5	สุกัลยา ฉายสุวรรณ	การสร้างแบบทดสอบวินิจจัยความรู้พื้นฐานทางพีชคณิตของนักเรียน ชั้นมัธยมศึกษาปีที่ 3 ในเขตกรุงเทพมหานคร	ม.เกษตรศาสตร์	2539
6	วรรณรัตน์ วิบูลสุข	การวินิจจัยข้อบกพร่องทางการเรียนคณิตศาสตร์ เรื่อง กำหนดการเชิงเส้น ของนักศึกษามหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ	ม.เกษตรศาสตร์	2539
7	กาญจนา ทัดสวน	การวินิจจัยข้อบกพร่องทางการเรียนวิทยาศาสตร์ เรื่อง เครื่องใช้ไฟฟ้าในบ้านของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในโรงเรียนโครงการขยายโอกาสทางการศึกษาสังกัดสำนักงานการประถมศึกษาจังหวัดสุพรรณบุรี	ม.เกษตรศาสตร์	2541
8	สุภาวดี กิตติวิศิษฎ์	การสร้างแบบทดสอบวินิจจัย วิชาคณิตศาสตร์ เรื่องสมการ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ในจังหวัดนครราชสีมา	ม.ขอนแก่น	2537
9	สมศักดิ์ อ้วนสาเล	การสร้างแบบทดสอบวินิจจัย วิชาคณิตศาสตร์ เรื่องฟังก์ชันตรีโกณมิติ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 4 ในจังหวัดหนองบัวลำภู	ม.ขอนแก่น	2540

ตารางที่ 1 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการสร้างหรือพัฒนาแบบทดสอบเชิงวินิจจัย (ต่อ)

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปี ที่ทำ
10	สุรพันธ์ พันธุ์มณี	การสร้างแบบทดสอบวินิจจัย วิชา คณิตศาสตร์ เรื่อง ฟังก์ชันลอการิทึม สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5 ใน จังหวัดมหาสารคาม	ม.ขอนแก่น	2541
11	อภิวัฒน์ ไชยศร	การสร้างแบบทดสอบวินิจจัยในการเรียนวิชา คณิตศาสตร์ เรื่องทศนิยม สำหรับนักเรียนชั้น ประถมศึกษาปีที่ 6 ในจังหวัดนครราชสีมา	ม.ขอนแก่น	2542
12	ณัฐพร ศรีบุรณ์	การสร้างแบบทดสอบอัตรันัยเพื่อวินิจจัยการ แก้โจทย์ปัญหาคณิตศาสตร์ เรื่องเศษส่วน ระดับชั้นมัธยมศึกษาปีที่ 2	ม.ขอนแก่น	2543
13	ประสงค์ คีนดี	การสร้างแบบทดสอบวินิจจัย วิชา คณิตศาสตร์ เรื่องเลขยกกำลัง สำหรับ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ในจังหวัด อุบลราชธานี	ม.มหาสารคาม	2535
14	วณิดา ปาณินิจ	การสร้างแบบทดสอบวินิจจัยในการเรียน เรื่องสมการและอสมการเชิงเส้นตัวแปรเดียว สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ใน จังหวัดร้อยเอ็ด	ม.มหาสารคาม	2535
15	สุมาลี ไชคสมุท	การสร้างแบบทดสอบวินิจจัยในการเรียนเรื่อง ตัวประกอบของจำนวนนับ สำหรับนักเรียนชั้น ประถมศึกษาปีที่ 6 สังกัดสำนักงานการ ประถมศึกษาจังหวัดมหาสารคาม	ม.มหาสารคาม	2535
16	ศุภวิทย์ จันทร์หอม	การสร้างแบบทดสอบวินิจจัยข้อบกพร่องนา การเรียนวิชาคณิตศาสตร์ (ค 012) เรื่อง วงกลมพาราโบลา วงรีไฮเพอร์โบลา สำหรับ นักเรียนชั้นมัธยมศึกษาปีที่ 4 สังกัดกรม สามัญศึกษา จังหวัดยโสธร	ม.มหาสารคาม	2536



ตารางที่ 1 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการสร้างหรือพัฒนาแบบทดสอบเชิงวินิจัย (ต่อ)

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปี ที่ทำ
17	สมจิต เมืองนาม	การสร้างแบบทดสอบวินิจัยการเรียนวิชาเคมี เรื่องพันธะเคมี สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 4 สังกัดกรมสามัญศึกษา จังหวัดมหาสารคาม	ม.มหาสารคาม	2536
18	สมชาย รบไพร่	การสร้างแบบทดสอบวินิจัยวิชาคณิตศาสตร์ เรื่อง ทศนิยม สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ในจังหวัดมหาสารคาม	ม.มหาสารคาม	2536
19	วุฒิชัย เหล่าพิเดช	การสร้างแบบทดสอบวินิจัยการเรียน เรื่อง เสียงและการได้ยินวิชาฟิสิกส์ สำหรับนักเรียน ชั้นมัธยมศึกษาปีที่ 4 สังกัดกรมสามัญศึกษา จังหวัดมหาสารคาม	ม.มหาสารคาม	2537
20	ดวงฤดี สิงคนุต	การสร้างแบบทดสอบวินิจัยการเรียนวิชาคณิตศาสตร์ ด้านทักษะการแก้โจทย์ปัญหา เรื่องการคูณและการหาร สำหรับนักเรียนชั้นประถมศึกษาปีที่ 4 ในสังกัดสำนักงานการประถมศึกษา จังหวัดอุบลราชธานี	ม.มหาสารคาม	2540
21	ราตรี ศรีนา	การสร้างแบบทดสอบวินิจัยวิชาคณิตศาสตร์ เรื่อง บทประยุกต์สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ในจังหวัดขอนแก่น	ม.มหาสารคาม	2540
22	สุริยะพงษ์ พงศ์สิทธิศักดิ์	การพัฒนาแบบทดสอบวินิจัย วิชาคณิตศาสตร์ เรื่องระบบสมการเชิงเส้น สำหรับนักเรียน ชั้นมัธยมศึกษาปีที่ 3 สังกัดกรมสามัญศึกษา จังหวัดมหาสารคาม	ม.มหาสารคาม	2540
23	อาภรณ์ เวียงวิเศษ	การพัฒนาแบบทดสอบวินิจัยในการเรียนเรื่องสมการสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ในเขตจังหวัดบุรีรัมย์	ม.มหาสารคาม	2540

ตารางที่ 1 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการสร้างหรือพัฒนาแบบทดสอบเชิงวินิจัย (ต่อ)

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปี ที่ทำ
24	มงคล ไชยประดิษฐ์	การสร้างแบบทดสอบวินิจัย วิชา คณิตศาสตร์ เรื่องเศษส่วนสำหรับนักเรียน ชั้นประถมศึกษาปีที่ 4 ในจังหวัดมหาสารคาม	ม.มหาสารคาม	2541
25	เรืองยศ เรืองแห่	การสร้างแบบทดสอบวินิจัย วิชา คณิตศาสตร์ เรื่องอัตราส่วนตรีโกณมิติ สำหรับนักเรียน ชั้นมัธยมศึกษาปีที่ 3 โรงเรียนขยายโอกาสทางการศึกษา สังกัด สำนักงานการประถมศึกษาจังหวัดขอนแก่น	ม.มหาสารคาม	2541
26	จินตนา สีนกิง	การพัฒนาแบบทดสอบวินิจัยทักษะพื้นฐาน ของความเข้าใจในการอ่านภาษาอังกฤษ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ใน จังหวัดสกลนคร	ม.มหาสารคาม	2542
27	มัญญ น้อยนาแสง	การสร้างแบบทดสอบวินิจัยในการเรียน เรื่อง การอินทิเกรตเบื้องต้น สำหรับ นักศึกษา ประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) ของวิทยาลัยเทคนิค สังกัดกรม อาชีวศึกษา ในเขตการศึกษา 10	ม.มหาสารคาม	2542
28	สมประสงค์ เกิดชัย	การสร้างแบบทดสอบวินิจัยวิชา คณิตศาสตร์ สำหรับนักเรียนชั้น ประถมศึกษาปีที่ 2 เรื่องโจทย์ปัญหาการ บวก ลบ คูณและหาร	ม.นเรศวร	2536
29	พิสมัย รัศมีดารา	การสร้างแบบทดสอบวินิจัยวิชา คณิตศาสตร์ เรื่องเศษส่วน สำหรับนักเรียน ชั้นประถมศึกษาปีที่ 5	ม.นเรศวร	2538
30	เบญจา เขียวสม	การสร้างแบบทดสอบวินิจัยตามพฤติกรรม การแก้โจทย์ปัญหา วิชาคณิตศาสตร์ ชั้น มัธยมศึกษาปีที่ 1	ม.ศรีนครินทรวิโรฒ	2533

ตารางที่ 1 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการสร้างหรือพัฒนาแบบทดสอบเชิงวินิจัย (ต่อ)

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปี ที่ทำ
31	ประพนธ์ จำเริญญ	การสร้างแบบทดสอบวินิจัยการเขียนสะกด คำในภาษาไทย สำหรับนักเรียนชั้น ประถมศึกษาปีที่ 2 ในโรงเรียนสาธิต สังกัด ทบวงมหาวิทยาลัย เขตกรุงเทพมหานคร	ม.ศรีนครินทรวิโรฒ	2533
32	สมชาย บุญรักษา	การสร้างแบบทดสอบวินิจัยในการเรียน คณิตศาสตร์ เรื่องสมการและอสมการเชิง เส้นตัวแปรเดียวและระบบสมการเชิงเส้น สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ใน จังหวัดพังงา	ม.ศรีนครินทรวิโรฒ	2535
33	สุชาติ สิริมีนนท์	การสร้างแบบทดสอบวินิจัยจุดบกพร่องใน การเรียนวิชาคณิตศาสตร์ เรื่องภาคตัดกรวย สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 4 ใน กรุงเทพมหานคร	ม.ศรีนครินทรวิโรฒ	2541
34	พิมสิริ ชัยจิตบริบูรณ์	การสร้างแบบทดสอบวินิจัยความเข้าใจใน การอ่านภาษาไทย สำหรับนักเรียนชั้น มัธยมศึกษาปีที่ 1	ม.รามคำแหง	2539
35	ธเนศ เต็งชู	การสร้างแบบทดสอบวินิจัยทางการเรียน คณิตศาสตร์	ม.รามคำแหง	2539
36	สมจิตร หามาลี	การสร้างแบบทดสอบวินิจัยความเข้าใจใน การอ่านภาษาอังกฤษ ระดับชั้นมัธยมศึกษา ปีที่ 3	ม.รามคำแหง	2540
37	พรสวรรค์ ศรีสุเทพ	การสร้างแบบทดสอบวินิจัยในการเรียน คณิตศาสตร์ เรื่องการแยกตัวประกอบของ พหุนาม สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3	ม.รามคำแหง	2540
38	อุมาวดี มะคำไก่อ	การสร้างแบบทดสอบวินิจัยความเข้าใจใน การอ่านภาษาไทย สำหรับนักเรียนชั้น ประถมศึกษาปีที่ 4	ม.รามคำแหง	2541

ตารางที่ 1 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการสร้างหรือพัฒนาแบบทดสอบเชิงวินิจัย (ต่อ)

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปี ที่ทำ
39	สุนทรี่ สิงห์พันธ์	การสร้างและการใช้แบบทดสอบวินิจัยทางการเรียน ด้านพุทธิพิสัยและทักษะกระบวนการทางวิทยาศาสตร์ เรื่องพลังงานและสารเคมี สำหรับนักเรียน ชั้นประถมศึกษาปีที่ 5	ม.รามคำแหง	2543
40	บุญสำรวย กฤตานุพงศ์	การวินิจัยข้อบกพร่องในการแก้ไขโจทย์ปัญหาคณิตศาสตร์ ของนักเรียนชั้นประถมศึกษาปีที่ 4	ม.สงขลานครินทร์	2536
41	มนตรี จรรย์านุกวัฒน์	การสร้างแบบทดสอบวินิจัยทางคณิตศาสตร์ เรื่องเศษส่วน สำหรับนักเรียนชั้นประถมศึกษาปีที่ 3 ในจังหวัดสตูล	ม.สงขลานครินทร์	2537
42	สมนึก ศรีนคร	การวินิจัยข้อบกพร่องในการเรียนคณิตศาสตร์เรื่องเศษส่วน ของนักเรียนชั้นประถมศึกษาปีที่ 6 ในจังหวัดสงขลา	ม.สงขลานครินทร์	2539
43	เจริญ กู้เขียว	การวินิจัยข้อบกพร่องความเข้าใจในการอ่านของนักเรียนชั้นประถมศึกษาปีที่ 6 ในจังหวัดพังงา	ม.สงขลานครินทร์	2540
44	สกาเว สันติเทวกุล	การพัฒนาแบบทดสอบวินิจัยวิชาฟิสิกส์ว 021 โดยใช้ทฤษฎีการตอบสนองข้อสอบ	ม.สงขลานครินทร์	2540
45	ประดิษฐ์ นวลจันทร์	การสร้างแบบทดสอบวินิจัยข้อบกพร่องในการเรียนวิชาฟิสิกส์ เรื่องการเคลื่อนที่แบบวงกลมและการเคลื่อนที่แบบซิมเปิลฮาร์โมนิก สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 6 ในจังหวัดนครปฐม	ม.สุโขทัยธรรมกิติราช	2539
46	ดวงพร เพ็ชรดอน	การสร้างแบบทดสอบวินิจัยทักษะการเขียนภาษาไทย สำหรับนักเรียนชั้นประถมศึกษาปีที่ 4	ม.ขอนแก่น	2545

ตารางที่ 1 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการสร้างหรือพัฒนาแบบทดสอบเชิงวินิจัย (ต่อ)

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปี ที่ทำ
47	นงลักษณ์ เสมอภาพ	การวินิจัยการเรียนรู้คณิตศาสตร์เรื่องการ หาร ของนักเรียนชั้นประถมศึกษาปีที่ 5 ใน โรงเรียนสังกัดสำนักงานการประถมศึกษา เขตการศึกษา 12	จุฬาลงกรณ์ มหาวิทยาลัย	2533
48	ประดิษฐ์ เรื่องตระกูล	การใช้โมเดลโลจิสติกเพื่อพัฒนาแบบสอบ วินิจัยวิชาคณิตศาสตร์สำหรับนักเรียนชั้น มัธยมศึกษาปีที่ 1	จุฬาลงกรณ์ มหาวิทยาลัย	2528
49	สุภาพ ละอองวิจิตร	การพัฒนาแบบทดสอบวินิจัยวิชา คณิตศาสตร์ เรื่องฟังก์ชัน โดยใช้ทฤษฎีการ ตอบสนองข้อสอบ สำหรับนักเรียนชั้น มัธยมศึกษาปีที่ 4	ร.ร.แพมดุงวิทยา	2543
50	หทัยรัตน์ ทองแน่น	การวินิจัยในห้องเรียนกลุ่มทักษะ คณิตศาสตร์ชั้นประถมศึกษาปีที่ 6	ร.ร.เมืองสุรินทร์	2534
51	แดนสวรรค์ สุดทน	การสร้างแบบทดสอบวินิจัยวิชา คณิตศาสตร์ ค 011	สนง.ศึกษาธิการ เขต 10 อุบลราชธานี	2534
52	พรรณผกา มณีวัง	การใช้แบบทดสอบวินิจัยวิชาภาษาอังกฤษ ชั้นประถมศึกษาปีที่ 6	ม.เชียงใหม่	2542
53	มาลา ไซติเพิ่ม	การสร้างแบบทดสอบวินิจัยทักษะการอ่าน ภาษาไทย สำหรับนักเรียนชั้นประถมศึกษา ปีที่ 6 ในจังหวัดสกลนคร	ม.มหาสารคาม	2536
54	สามารถ จ้อยทอง	การสร้างแบบทดสอบวินิจัยทางการเรียน คณิตศาสตร์ สำหรับนักเรียนชั้น ประถมศึกษาปีที่ 4 จากการศึกษา ข้อบกพร่องทางการเรียนของนักเรียน	สพจ.อ่างทอง	2537
55	ทวิกา แก้วมกระโทก	การพัฒนาแบบสอบวินิจัยข้อบกพร่องใน การคิดคำนวณ	ม.เกษตรศาสตร์	2544

ตารางที่ 1 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการสร้างหรือพัฒนาแบบทดสอบเชิงวินิจัย (ต่อ)

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปี ที่ทำ
56	สุรางค์ ชูสิงห์	สภาพและปัญหาการเรียนคณิตศาสตร์ เรื่อง การหาร ชั้นประถมศึกษาปีที่ 4 โดยใช้แบบทดสอบวินิจัยข้อบกพร่องทางการเรียนกับนักเรียนสังกัดสำนักงานการประถมศึกษาอำเภอชุมแสง	สพจ.นครสวรรค์	2543
57	สุภาพ วชิรศิริ	การสร้างและการใช้แบบทดสอบวินิจัยวิชาคณิตศาสตร์ ด้านการแก้โจทย์ปัญหา เรื่อง การบวก ลบ คูณ หาร โดยใช้สมการสำหรับนักเรียนชั้นประถมศึกษาปีที่ 6	ม.รามคำแหง	2545
58	สุพิศ ฝักแต่	การพัฒนาแบบทดสอบวินิจัยในการเรียนวิชาคณิตศาสตร์ ค 101 เรื่องทศนิยม สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1	สรม. อุบลราชธานี	2545
59	กิตติวัฒน์ ภูริพัฒน์	การพัฒนาแบบทดสอบวินิจัยในการเรียนวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 4 เรื่อง ฟังก์ชันตรีโกณมิติ	สรม. อุบลราชธานี	2545
60	ภูรินาถ ไชคากรณ์	การสร้างแบบทดสอบวินิจัยตามพฤติกรรม การแก้โจทย์ปัญหาเรื่องร้อยละ วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2	ม.ทักษิณ	2545
61	วีณา วโรดมะวิษญ์	การสร้างแบบทดสอบวินิจัยการเรียนคณิตศาสตร์เรื่องการหาร สำหรับนักเรียนชั้นประถมศึกษาปีที่ 4	ม.เชียงใหม่	2541

ตารางที่ 2 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการสร้างหรือพัฒนาแบบทดสอบเชิงวินิจฉัยและการ  
แก้ไขข้อบกพร่อง

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปีที่ทำ
1	สุมาลี เวียงรัตน์	การวินิจฉัยและแก้ไขข้อบกพร่องทางการ เรียน เรื่องสมการของนักเรียน ชั้น มัธยมศึกษาปีที่ 2 โรงเรียนบ้านตระแสง จังหวัดสุรินทร์	มสุโขทัยธรรมดิราช	2541
2	รินฤดี เทพขุน	รายงานผลการศึกษาเกี่ยวกับการวินิจฉัย ข้อบกพร่องและแก้ไขข้อบกพร่องในการแต่ง ประโยคความเดียว ประโยคความรวมและ ประโยคความซ้อน ชั้นมัธยมศึกษาปีที่ 2 โรงเรียนชินโรรสวิทยาลัย	ร.ร.ชินโรรส วิทยาลัย	2532

ตารางที่ 3 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการเขียนตัวลง

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปีที่ทำ
1	ศักดิ์สิทธิ์ ฤทธิสิน	การเปรียบเทียบคุณภาพของแบบทดสอบ วินิจฉัยวิชาคณิตศาสตร์ที่สร้างตัวลง โดยวิธี ต่างกัน	ม.มหาสารคาม	2541

ตารางที่ 4 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการพัฒนาดัชนีเพื่อป้องกันความผิดปกติของแบบแผน  
การตอบข้อสอบ

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปีที่ทำ
1	สำราญ มีแจ้ง	การพัฒนาดัชนีชี้ความบกพร่องของผู้ตอบ แบบทดสอบโดยการวิเคราะห์แบบแผนการ ตอบข้อกระทง	จุฬาลงกรณ์ มหาวิทยาลัย	2533
2	พนินชา สังข์เพชร	การพัฒนาดัชนีชี้ความสอดคล้องของแบบ แผนการตอบข้อสอบรายบุคคล	จุฬาลงกรณ์ มหาวิทยาลัย	2535

ตารางที่ 5 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการใช้ดัชนีบ่งชี้ความผิดปกติของแบบแผนการตอบข้อสอบ

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปีที่ทำ
1	องอาจ วิชัยสุชาติ	การวิจัยข้อบกพร่องทางการเรียน คณิตศาสตร์ เรื่องทศนิยม ของนักเรียนชั้น ประถมศึกษาปีที่ 5 จังหวัดนครราชสีมา โดย ใช้แผนภูมิเอส-พีและดัชนีบ่งชี้ของซาโต้	ม.เกษตรศาสตร์	2536
2	สุทธิ กระจะจ่าง	การวิจัยข้อบกพร่องทางการเรียนจากแบบ แผนการตอบข้อสอบวิชาคณิตศาสตร์ เรื่อง อัตราส่วนและร้อยละ ของนักเรียนชั้น มัธยมศึกษาปีที่ 1 โรงเรียนบางช้างวิทยา โดยใช้ทฤษฎีเส้นโค้งเอส-พีและดัชนีบ่งชี้ของ ซาโต้	ม.เกษตรศาสตร์	2538

ตารางที่ 6 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการพัฒนาโปรแกรมคอมพิวเตอร์เพื่อการวิเคราะห์  
แบบแผนการตอบข้อสอบ

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปีที่ทำ
1	ชัยณรงค์ วิริยะศรีสุวัฒนา	การพัฒนาโปรแกรมไมโครคอมพิวเตอร์เพื่อ วิเคราะห์แบบแผนการตอบข้อสอบโดยใช้สูตร ดัชนีชี้เตือนที่ปรับปรุงโดยฮาร์นิสและลินน์	จุฬาลงกรณ์ มหาวิทยาลัย	2534

ตารางที่ 7 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับแนวคิดการประเมินกฎในการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อน

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปีที่ทำ
1	ศิริเดช สุชีวะ	การพัฒนาวิธีการวินิจฉัยสำหรับตรวจสอบ มโนทัศน์ที่คลาดเคลื่อนทางคณิตศาสตร์	จุฬาลงกรณ์ มหาวิทยาลัย	2538
2	อภิญา ตังประพฤทิกุล	วิธีการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อน ทางการเรียนคณิตศาสตร์ เรื่องการหาร ใน ระดับชั้นประถมศึกษาปีที่ 3	ม.ขอนแก่น	2543



ตารางที่ 8 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการคำนึงถึงสภาพจริงในการแก้ไขภัยปัญหา

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปีที่ทำ
1	วลี เฉลยสมัย	การพัฒนาวิธีการวินิจฉัยการแก้ไขภัยปัญหา คณิตศาสตร์ที่คำนึงถึงสภาพที่เป็นจริง ของ นักเรียนชั้นประถมศึกษาปีที่ 6	จุฬาลงกรณ์ มหาวิทยาลัย	2538

ตารางที่ 9 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับการวินิจฉัยและแก้ไขข้อบกพร่องโดยนักเรียน

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปีที่ทำ
1	อรนุช เขียมธรรม	ผลการสอนวิธีการวินิจฉัยและแก้ไขงานเขียนความ เรียงที่ไม่ชัดเจนที่มีต่อความสามารถในการเขียนความ เรียงภาษาไทยของนักเรียนชั้นมัธยมศึกษาปีที่ 4	จุฬาลงกรณ์ มหาวิทยาลัย	2543

ตารางที่ 10 รายชื่อและแหล่งที่มาของงานวิจัยเกี่ยวกับเทคนิคการคิดออกเสียง

ที่	ชื่อ – สกุลผู้วิจัย	ชื่อเรื่อง	สถาบัน/ หน่วยงาน	ปีที่ทำ
1	สุดารัตน์ มนต์นิมิตร	การใช้เทคนิคการคิดออกเสียงเป็นเครื่องมือ ในการวินิจฉัยความสามารถในการแก้ปัญห าโจทย์คณิตศาสตร์เพื่อจัดสอนซ่อมเสริม สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3	จุฬาลงกรณ์ มหาวิทยาลัย	2545

ภาคผนวก ข

เครื่องมือที่ใช้ในการวิจัย

**แบบสำรวจคุณลักษณะที่ใช้ในการบวกลบเศษส่วน  
ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6)**

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสำรวจ

คำชี้แจง โปรดทำเครื่องหมาย ✓ ลงใน  หน้าข้อความหรือเติมข้อความ  
ที่เป็นข้อมูลเกี่ยวกับตัวท่านตามสภาพที่เป็นจริง

- |  |                          |                  |                          |             |
|--|--------------------------|------------------|--------------------------|-------------|
| 1. เพศ   | <input type="checkbox"/> | หญิง             | <input type="checkbox"/> | ชาย         |
| 2. อายุ  | <input type="checkbox"/> | ต่ำกว่า 25 ปี    | <input type="checkbox"/> | 25 – 30 ปี  |
|  | <input type="checkbox"/> | 31 – 35 ปี       | <input type="checkbox"/> | 36 – 40 ปี  |
|  | <input type="checkbox"/> | 41 – 45 ปี       | <input type="checkbox"/> | 56 ปีขึ้นไป |
| 3. การศึกษาสูงสุด  | <input type="checkbox"/> | ต่ำกว่าปริญญาตรี | <input type="checkbox"/> | ปริญญาตรี   |
|  | <input type="checkbox"/> | ปริญญาโท         | <input type="checkbox"/> | ปริญญาเอก   |
| 4. สาขาวิชา/วิชาเอกที่สำเร็จการศึกษา.....                        |                          |                  |                          |             |
| 5. ประสบการณ์ในการสอนวิชาคณิตศาสตร์ ระดับประถมศึกษา จำนวน.....ปี |                          |                  |                          |             |

ตอนที่ 2 ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วน

คำชี้แจง โปรดจำแนกความรู้ ทักษะหรือกระบวนการที่นักเรียนต้องใช้เพื่อให้สามารถ  
ดำเนินการบวกลบเศษส่วนแต่ละประเภทได้อย่างถูกต้อง (การบวกลบเศษส่วนแต่ละประเภท  
อาจใช้ความรู้ ทักษะหรือกระบวนการมากกว่า 1 อย่างและอาจจะซ้ำกับประเภทอื่น ๆ ได้)

เช่น ในการบวกเศษส่วนแท้ที่มีส่วนเท่ากัน นักเรียนจะต้อง

1. เข้าใจความหมายของเศษส่วนแท้ที่มีส่วนเท่ากัน
2. เข้าใจวิธีบวกลบเศษส่วนแท้ที่มีส่วนเท่ากัน
3. บวกลบจำนวนนับ

ความรู้ ทักษะหรือกระบวนการที่ใช้

1. การรวบรวมเศษส่วนเท่าที่มีส่วนเท่ากัน

 .....	 .....
.....	.....
.....	.....
.....	.....

2. การรวบรวมเศษส่วนเท่าที่มีส่วนไม่เท่ากัน

 .....	 .....
.....	.....
.....	.....
.....	.....


3. การรวบรวมจำนวนคละที่มีส่วนเท่ากัน

 .....	 .....
.....	.....
.....	.....
.....	.....

4. การรวบรวมจำนวนคละที่มีส่วนไม่เท่ากัน

 .....	 .....
.....	.....
.....	.....
.....	.....

ข้อเสนอแนะ/ความคิดเห็น

 .....
.....

**แบบประเมินความถูกต้องของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน**

**คำชี้แจง** โปรดพิจารณาและตัดสินความถูกต้องของความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแต่ละข้อ และทำเครื่องหมาย ✓ ลงในช่อง “ถูกต้อง” หรือ “ไม่ถูกต้อง” เพื่อประเมินความถูกต้องของความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วน โปรดกำหนดความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนเพิ่มเติม หรือชี้ค่าความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนออกกรณีเห็นว่าความรู้ ทักษะหรือกระบวนการที่กำหนดไม่ใช่ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนนั้น ๆ

ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วน	ผลการประเมิน		ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนที่ควรกำหนดเพิ่ม
	ถูกต้อง	ไม่ถูกต้อง	
1. ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีส่วนเท่ากัน 1.1 เข้าใจความหมายของเศษส่วนแท้ที่มีตัวส่วนเท่ากัน 1.2 เข้าใจวิธีบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน 1.3 บวกลบจำนวนนับ 1.4 เข้าใจวิธีทำผลลัพธ์ให้เป็นเศษส่วนอย่างต่ำ 1.5 คูณจำนวนนับ 1.6หารจำนวนนับ			
2. ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีส่วนไม่เท่ากัน 2.1 เข้าใจความหมายของเศษส่วนแท้ที่มีตัวส่วนเท่ากัน 2.2 เข้าใจความหมายของเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน 2.3 เข้าใจวิธีบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน 2.4 เข้าใจวิธีบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน			

ความรู้ ทักษะหรือกระบวนการที่ใช้ในการ บวกลบเศษส่วน	ผลการประเมิน		ความรู้ ทักษะหรือกระบวนการ ที่ใช้ในการบวกลบเศษส่วนที่ ควรกำหนดเพิ่ม
	ถูกต้อง	ไม่ ถูกต้อง	
2.5 เข้าใจวิธีทำเศษส่วนแท้ 2 จำนวนให้มี ตัวส่วนเท่ากัน 2.6 บวกลบจำนวนนับ 2.7 เข้าใจวิธีทำผลลัพธ์ให้เป็นเศษส่วนอย่างต่ำ 2.8 คูณจำนวนนับ 2.9 หารจำนวนนับ			
3. ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบ จำนวนคละที่มีส่วนเท่ากัน 3.1 เข้าใจความหมายของเศษส่วนแท้ที่มี ตัวส่วนเท่ากัน 3.2 เข้าใจความหมายของจำนวนคละ 3.3 เข้าใจความหมายของเศษเกิน 3.4 เข้าใจวิธีบวกลบจำนวนคละที่มีตัวส่วน เท่ากัน <b>วิธีที่ 1</b> 1) บวกลบจำนวนนับ 2) เข้าใจวิธีบวกลบเศษส่วนที่มีส่วนเท่ากัน <b>วิธีที่ 2</b> 1) เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน 2) เข้าใจวิธีบวกลบเศษเกินที่มีตัวส่วน เท่ากัน 3.5 เข้าใจวิธีทำผลลัพธ์ให้เป็นเศษส่วนอย่างต่ำ 3.6 คูณจำนวนนับ 3.7 หารจำนวนนับ			

ความรู้ ทักษะหรือกระบวนการที่ใช้ในการ บวกลบเศษส่วน	ผลการประเมิน		ความรู้ ทักษะหรือกระบวนการ ที่ใช้ในการบวกลบเศษส่วนที่ ควรกำหนดเพิ่ม
	ถูกต้อง	ไม่ ถูกต้อง	
<p>4. ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบ จำนวนคละที่มีส่วนเท่ากัน</p> <p>4.1 เข้าใจความหมายของเศษส่วนแท้ที่มี ตัวส่วนเท่ากัน</p> <p>4.2 เข้าใจความหมายของเศษส่วนแท้ที่มี ตัวส่วนไม่เท่ากัน</p> <p>4.3 เข้าใจความหมายของจำนวนคละ</p> <p>4.4 เข้าใจความหมายของเศษเกิน</p> <p>4.5 เข้าใจวิธีบวกลบจำนวนคละที่มีตัวส่วน ไม่เท่ากัน</p> <p><b>วิธีที่ 1</b></p> <p>1) บวกลบจำนวนนับ</p> <p>2) เข้าใจวิธีบวกลบเศษส่วนที่มีส่วนเท่ากัน</p> <p>3) เข้าใจวิธีทำเศษส่วนแท้ 2 จำนวน ให้มี ตัวส่วนเท่ากัน</p> <p><b>วิธีที่ 2</b></p> <p>1) เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน</p> <p>2) เข้าใจวิธีบวกลบเศษเกินที่มีตัวส่วนไม่เท่ากัน</p> <p>3) เข้าใจวิธีทำเศษเกิน 2 จำนวน ให้มีตัวส่วนเท่ากัน</p> <p>4) เข้าใจวิธีบวกลบเศษเกินที่มีตัวส่วนเท่ากัน</p> <p>5) บวกลบจำนวนนับ</p> <p>4.6 เข้าใจวิธีทำผลลัพธ์ให้เป็นเศษส่วนอย่างต่ำ</p> <p>4.7 คูณจำนวนนับ</p> <p>4.8 หหารจำนวนนับ</p>			

**ข้อเสนอแนะ/ความคิดเห็น**

.....

.....

.....

**แบบเปรียบเทียบลำดับขั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน**

**คำชี้แจง** โปรดเปรียบเทียบลำดับขั้นของความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วน แต่ละข้อต่อไปนี่ทีละคู่ และพิจารณาว่าความรู้ ทักษะหรือกระบวนการทั้งสองมีความเกี่ยวข้องสัมพันธ์กันหรือไม่ ความรู้ ทักษะหรือกระบวนการใดเป็นความรู้ ทักษะหรือกระบวนการที่ต้องมีมาก่อน โดยทำเครื่องหมายล้อมรอบหรือทับตัวเลข ดังนี้

- เครื่องหมาย ✕ ทับตัวเลขทั้งคู่ กรณีเป็นความรู้ ทักษะหรือกระบวนการที่ไม่เกี่ยวข้องสัมพันธ์กัน
- เครื่องหมาย ○ ล้อมรอบตัวเลขที่เป็นความรู้ ทักษะหรือกระบวนการที่ต้องมีมาก่อน
- เครื่องหมาย ○ ล้อมรอบตัวเลขทั้งคู่ กรณีเป็นความรู้ ทักษะหรือกระบวนการที่อยู่ในลำดับขั้นเดียวกัน

**ตัวอย่างเช่น** เมื่อเปรียบเทียบ 1 กับ 2 พบว่า ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกจำนวนนับไม่เกี่ยวข้องกับความรู้ ทักษะหรือกระบวนการที่ใช้ในการลบจำนวนนับ และเมื่อเปรียบเทียบ 1 กับ 3 พบว่า ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกเป็นความรู้ ทักษะหรือกระบวนการที่ต้องมีมาก่อน ความรู้ ทักษะหรือกระบวนการที่ใช้ในการคูณ ให้ทำเครื่องหมาย ดังนี้

1. ความรู้ความสามารถที่ใช้ในการบวกจำนวนนับ
2. ความรู้ความสามารถที่ใช้ในการลบจำนวนนับ
3. ความรู้ความสามารถที่ใช้ในการคูณจำนวนนับ

✕	①
✕	3
	2
	3

1. ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีส่วนเท่ากัน	1   1   1
2. ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีส่วนไม่เท่ากัน	2   3   4
3. ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบจำนวนคละที่มีส่วนเท่ากัน	2   2
4. ความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบจำนวนคละที่มีส่วนไม่เท่ากัน	3   4
	3
	4

**ข้อเสนอแนะ/ความคิดเห็น**

✍️ .....

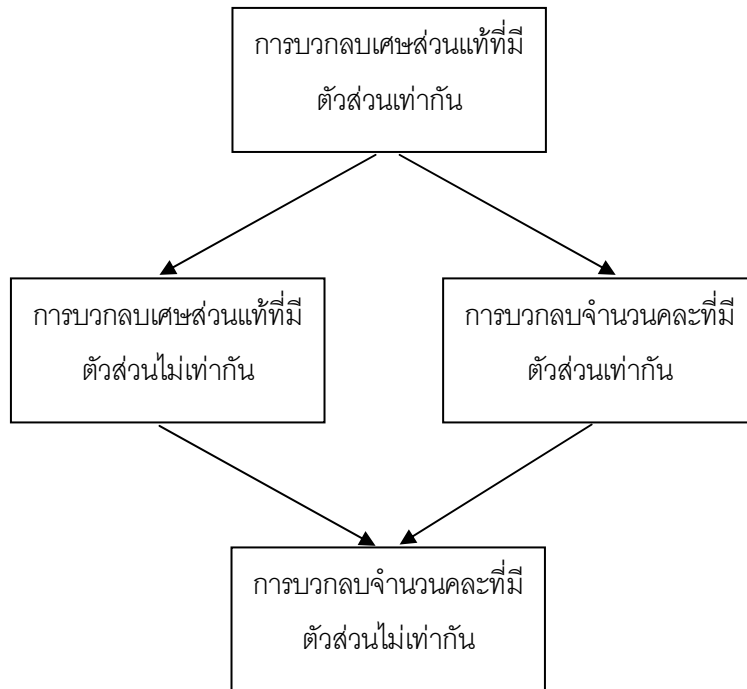
.....

.....



**แบบประเมินความถูกต้องของแผนภาพลำดับชั้นของคุณลักษณะที่ใช้ในการควบคุมพิเศษ**

คำชี้แจง โปรดพิจารณาและตัดสินความถูกต้องของแผนภาพลำดับชั้นของคุณลักษณะในการควบคุมพิเศษส่วนข้างล่างนี้ และทำเครื่องหมาย ✓ ลงในช่อง  เพื่อประเมินความถูกต้องหรือไม่ถูกต้องของแผนภาพลำดับชั้นของคุณลักษณะในการควบคุมพิเศษ โปรดให้เหตุผลและ/หรือข้อเสนอแนะเพื่อการปรับปรุงแก้ไข กรณีเห็นว่าเป็นคุณลักษณะที่ไม่ถูกต้อง



คำอธิบาย การควบคุมส่วนแท้ที่มีตัวส่วนเท่ากันเป็นคุณลักษณะที่อยู่ในลำดับชั้นที่ 1 และเป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะที่อยู่ในลำดับชั้นที่ 2 ได้แก่ การควบคุมพิเศษแท้ที่มีตัวส่วนไม่เท่ากันและการควบคุมจำนวนคละที่มีตัวส่วนเท่ากัน ซึ่งเป็นคุณลักษณะที่ต้องมีมาก่อนคุณลักษณะที่อยู่ในลำดับชั้นที่ 3 คือ การควบคุมจำนวนคละที่มีตัวส่วนไม่เท่ากัน

ผลการประเมิน  ถูกต้อง  ไม่ถูกต้อง

เหตุผลและ/หรือข้อเสนอแนะเพื่อการปรับปรุงแก้ไข.....  
 .....  
 .....

แบบประเมินความครอบคลุมและความครบถ้วนของแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบประเมิน

คำชี้แจง โปรดทำเครื่องหมาย ✓ ลงใน  หน้าข้อความหรือเติมข้อความที่เป็นข้อมูลเกี่ยวกับตัวท่านตามสภาพที่เป็นจริง

1. เพศ  หญิง  ชาย
2. อายุ  ต่ำกว่า 25 ปี  25 – 30 ปี  
 31 – 35 ปี  36 – 40 ปี  
 41 – 45 ปี  46 ปีขึ้นไป
3. การศึกษาสูงสุด  ต่ำกว่าปริญญาตรี  ปริญญาตรี  
 ปริญญาโท  ปริญญาเอก

6. สาขาวิชา/วิชาเอกที่สำเร็จการศึกษา.....


7. ประสบการณ์ในการสอนวิชาคณิตศาสตร์ ระดับประถมศึกษา จำนวน.....ปี

ตอนที่ 2 ความครอบคลุมและความครบถ้วนของแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน

คำชี้แจง โปรดพิจารณาและตัดสินความครอบคลุมและความครบถ้วนของแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนจากเอกสารแสดงแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน และทำเครื่องหมาย ✓ ลงในช่อง  เพื่อประเมินความครอบคลุมและความครบถ้วนหรือความไม่ครอบคลุมและไม่ครบถ้วนของแบบการคิดที่ผิด โปรดให้รายละเอียดแบบการคิดที่ผิดที่ควรกำหนดเพิ่มเติม กรณีเห็นว่าไม่ครอบคลุมและไม่ครบถ้วน

แบบการคิดที่ผิด หมายถึง หลักการหรือวิธีการที่ใช้ในการแก้ปัญหาที่ไม่ได้ให้คำตอบที่ถูกต้องในทุกเงื่อนไขของปัญหาในเรื่องการบวกลบเศษส่วน

ผลการประเมิน	แบบการคิดที่ผิดที่ใช้ในการบวกเศษส่วนที่ควรกำหนดเพิ่มเติม
<b>1.การบวกเศษส่วน แท้ที่มีตัวส่วนเท่ากัน</b> <input type="checkbox"/> ครอบคลุม/ครบถ้วน <input type="checkbox"/> ไม่ครอบคลุม/ ไม่ครบถ้วน	<hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/>
<b>2.การบวกเศษส่วน แท้ที่มีตัวส่วนไม่เท่ากัน</b> <input type="checkbox"/> ครอบคลุม/ครบถ้วน <input type="checkbox"/> ไม่ครอบคลุม/ ไม่ครบถ้วน	<hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/>
<b>3.การบวกจำนวน คละที่มีตัวส่วนเท่ากัน</b> <input type="checkbox"/> ครอบคลุม/ครบถ้วน <input type="checkbox"/> ไม่ครอบคลุม/ ไม่ครบถ้วน	<hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/>
<b>4.การบวกจำนวน คละที่มีตัวส่วนไม่เท่ากัน</b> <input type="checkbox"/> ครอบคลุม/ครบถ้วน <input type="checkbox"/> ไม่ครอบคลุม/ ไม่ครบถ้วน	<hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/>

ความคิดเห็น/ข้อเสนอแนะเพิ่มเติม  .....

.....

แบบสอบถามปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์  
ในการทดสอบแบบปรับเหมาะ เพื่อการประเมินเชิงวินิจฉัยสถานะความรู้  
และแบบการคิดที่ผิดในการบวกลบเศษส่วน

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม


คำชี้แจง โปรดทำเครื่องหมาย ✓ ลงใน  หน้าข้อความหรือเติมข้อความที่เป็นข้อมูล  
เกี่ยวกับตัวท่านตามสภาพที่เป็นจริง

1. เพศ  หญิง  ชาย
2. อายุ  ต่ำกว่า 25 ปี  25 – 30 ปี  
 31 – 35 ปี  36 – 40 ปี  
 41 – 45 ปี  46 ปีขึ้นไป
3. การศึกษาสูงสุด  ต่ำกว่าปริญญาตรี สาขาวิชา/วิชาเอก.....  
 ปริญญาตรี สาขาวิชา/วิชาเอก.....  
 ปริญญาโท สาขาวิชา/วิชาเอก.....  
 ปริญญาเอก สาขาวิชา/วิชาเอก.....

ตอนที่ 2 ปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์

คำชี้แจง โปรดทำเครื่องหมาย ✓ ลงใน ช่อง “สภาพปัญหา” เพื่อระบุว่า “มีปัญหา” หรือ  
“ไม่มีปัญหา” กรณีมีปัญหา โปรดระบุปัญหาที่เกิดขึ้น โดยเขียนสภาพปัญหา  
ลงในช่อง “โปรดระบุสภาพปัญหา” และประเมินระดับความมากน้อยของปัญหา  
โดยทำเครื่องหมาย ✓ ลงใน ช่อง “ระดับปัญหา” ตามเกณฑ์ต่อไปนี้

- 3 หมายถึง มีปัญหานั้น ๆ เกิดขึ้นในระดับมาก  
2 หมายถึง มีปัญหานั้น ๆ เกิดขึ้นในระดับปานกลาง  
1 หมายถึง มีปัญหานั้น ๆ เกิดขึ้นในระดับน้อย

มีต่อหน้า 2 

	สภาพ		ระดับปัญหา
--	------	--	------------

ด้าน	ปัญหา		โปรดระบุสภาพปัญหา			
	มี	ไม่มี		3	2	1
1. การเข้าสู่โปรแกรม						
2. การแสดงผลทางหน้าจอ						
2.1 ตัวอักษร (สี, ความชัดเจน..)						
2.2 พื้นหลัง						
2.3 การจัดรูปแบบ (ส่วนคำถาม, ส่วนคำตอบ, ข้อมูลอื่น ๆ .....						
2.4 อื่น ๆ						
3. การพิมพ์ข้อมูลของผู้สอบ						
4. การคัดเลือกและนำเสนอข้อคำถาม						
5. การพิมพ์คำตอบ						
6. การส่งพิมพ์ผลการทดสอบทางเครื่องพิมพ์						
7. อื่น ๆ						

ตอนที่ 3 ข้อสังเกต/ ข้อคิดเห็น/ ข้อเสนอแนะ

✍️.....

.....

.....

.....

.....

.....

.....

.....

.....



แบบสอบถามปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์  
ในการทดสอบแบบปรับเหมาะ เพื่อการประเมินเชิงวินิจฉัยสถานะความรู้  
และแบบการคิดที่ผิดในการบวกลบเศษส่วน

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม

คำชี้แจง โปรดทำเครื่องหมาย ✓ ลงใน  หน้าข้อความหรือเติมข้อความที่เป็นข้อมูล  
เกี่ยวกับนักเรียนตามสภาพที่เป็นจริง

1. เพศ  หญิง  ชาย

2. ชั้น .....

3. อายุ ..... ปี

ตอนที่ 2 ปัญหาที่เกิดจากการใช้โปรแกรมคอมพิวเตอร์

คำชี้แจง โปรดทำเครื่องหมาย ✓ ลงใน ช่อง “สภาพปัญหา” เพื่อระบุว่า “มีปัญหา” หรือ  
“ไม่มีปัญหา” กรณีมีปัญหา โปรดระบุปัญหาที่เกิดขึ้นโดยเขียนสภาพปัญหาลงใน  
ช่อง “โปรดระบุสภาพปัญหา” และประเมินระดับความมากน้อยของปัญหา โดยทำ  
เครื่องหมาย ✓ ลงใน ช่อง “ระดับปัญหา” ทับภาพใบหน้าเพื่อแสดงความหมาย  
ดังนี้

☹ หมายถึง มีปัญหานั้น ๆ เกิดขึ้นในระดับมาก

☺ หมายถึง มีปัญหานั้น ๆ เกิดขึ้นในระดับปานกลาง

😊 หมายถึง มีปัญหานั้น ๆ เกิดขึ้นในระดับน้อย

ด้าน	สภาพปัญหา		โปรดระบุสภาพปัญหา	ระดับปัญหา		
	มี	ไม่มี		☹	☺	😊
1. การเข้าสู่โปรแกรม				☹	☺	😊
2. การแสดงผลทางหน้าจอ				☹	☺	😊
2.1 ตัวอักษร (สี, ความชัดเจน..)				☹	☺	😊
2.2 พื้นหลัง				☹	☺	😊
2.3 การจัดรูปแบบ (ส่วนคำถาม, ส่วนคำตอบ, ข้อมูลอื่น ๆ .....				☹	☺	😊
2.4 อื่น ๆ				☹	☺	😊
3. การพิมพ์ข้อมูลของผู้สอบ				☹	☺	😊
4. การคัดเลือกและนำเสนอข้อคำถาม				☹	☺	😊
5. การพิมพ์คำตอบ				☹	☺	😊
6. การส่งพิมพ์ผลการทดสอบทางเครื่องพิมพ์				☹	☺	😊
7. อื่น ๆ				☹	☺	😊

ตอนที่ 3 ข้อสังเกต/ ข้อคิดเห็น/ ข้อเสนอแนะ

✍.....

.....

.....

.....

.....

.....

.....

.....

.....

.....



## แบบสอบถาม

ความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจิตร โดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะ และการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม

คำชี้แจง โปรดทำเครื่องหมาย ✓ ลงใน  หน้าข้อความหรือเติมข้อความที่เป็นข้อมูลเกี่ยวกับตัวท่านตามสภาพที่เป็นจริง

- |                   |                          |  |                          |             |
|-------------------|--------------------------|--|--------------------------|-------------|
| 1. เพศ            | <input type="checkbox"/> | หญิง                                   | <input type="checkbox"/> | ชาย         |
| 2. อายุ           | <input type="checkbox"/> | ต่ำกว่า 25 ปี                          | <input type="checkbox"/> | 25 – 30 ปี  |
|                   | <input type="checkbox"/> | 31 – 35 ปี                             | <input type="checkbox"/> | 36 – 40 ปี  |
|                   | <input type="checkbox"/> | 41 – 45 ปี                             | <input type="checkbox"/> | 46 ปีขึ้นไป |
| 3. การศึกษาสูงสุด | <input type="checkbox"/> | ต่ำกว่าปริญญาตรี สาขาวิชา/วิชาเอก..... |                          |             |
|                   | <input type="checkbox"/> | ปริญญาตรี                              | สาขาวิชา/วิชาเอก.....    |             |
|                   | <input type="checkbox"/> | ปริญญาโท                               | สาขาวิชา/วิชาเอก.....    |             |
|                   | <input type="checkbox"/> | ปริญญาเอก                              | สาขาวิชา/วิชาเอก.....    |             |

ตอนที่ 2 ความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจิตร โดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

วิธีการประเมินเชิงวิจิตร โดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ หมายถึงการประเมินที่ให้นักเรียนทำข้อสอบผ่านหน้าจอคอมพิวเตอร์ โปรแกรมคอมพิวเตอร์จะคัดเลือกข้อสอบที่มีความยากง่ายปานกลางให้นักเรียนทำเป็นข้อแรก กรณีที่นักเรียนตอบถูกโปรแกรมจะคัดเลือกข้อสอบข้อต่อไปที่ยากขึ้น กรณีที่นักเรียนตอบผิดโปรแกรมจะคัดเลือกข้อสอบข้อต่อไปที่ง่ายลง การทดสอบจะดำเนินการไปจนกระทั่งถึงเกณฑ์การหยุดสอบ โปรแกรมจะหยุดคัดเลือกข้อสอบและรายงานผลการทดสอบของนักเรียนเป็นรายบุคคล นักเรียนแต่ละคนจึงได้ทำข้อสอบในข้อที่แตกต่างกัน จำนวนที่แตกต่างกันและใช้เวลาในการทำที่แตกต่างกัน



**คำชี้แจง** โปรดแสดงความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจจัย โดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ โดยเขียนตอบคำถามต่อไปนี้ตามความคิดเห็นของท่าน

1. ท่านคิดว่าวิธีการประเมินเชิงวิจจัยโดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์นำไปใช้ได้ยากหรือง่ายเพียงใด เพราะอะไร

✍.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

2. ถ้านำวิธีประเมินเชิงวิจจัยโดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ไปใช้ในโรงเรียน ท่านคิดว่าเป็นไปได้มากน้อยเพียงใด เพราะอะไร


✍.....  
.....  
.....  
.....  
.....  
.....

3. ท่านคิดว่าโรงเรียน ครู นักเรียน และผู้ปกครองจะได้รับประโยชน์จากผลการประเมินเชิงวิจจัยโดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มากน้อยเพียงใด อย่างไรบ้าง


- ประโยชน์ที่มีต่อโรงเรียน

✍.....  
.....


- ประโยชน์ที่มีต่อครู

.....  
.....


ประโยชน์ที่มีต่อนักเรียน

.....  
.....

ประโยชน์ที่มีต่อผู้ปกครอง

.....  
.....

ท่านขอทราบประเมินเชิงวิจิจฉัยโดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบ  
ปรับเหมาะโดยใช้คอมพิวเตอร์หรือไม่ เพราะอะไร

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

5. ถ้าให้ท่านใช้วิธีการประเมินเชิงวิจิจฉัยโดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะ  
และการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เพื่อประเมินนักเรียนท่านจะใช้หรือไม่ใช้  
เพราะอะไร

.....  
.....  
.....  
.....  
.....  
.....

แบบสอบถาม

ความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจจัย โดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะ และการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม


1. เพศ  หญิง  ชาย

2. ชั้น .....

3. อายุ ..... ปี

ตอนที่ 2 ความคิดเห็นที่มีต่อวิธีการประเมินเชิงวิจจัย โดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

1. นักเรียนคิดว่าการทำแบบทดสอบโดยใช้คอมพิวเตอร์ทำได้ยากหรือง่ายเพียงใด เพราะอะไร

 .....

.....

.....

.....

.....

.....

.....

2. ถ้าครูจะประเมินนักเรียน โดยให้นักเรียนทำแบบทดสอบโดยใช้คอมพิวเตอร์ นักเรียนคิดว่าเป็นไปได้หรือไม่ เพราะอะไร

 .....

.....


.....

.....

.....

.....

3. นักเรียนได้รับประโยชน์จากผลการประเมินที่ได้จากการทำแบบทดสอบโดยใช้คอมพิวเตอร์มากน้อยเพียงใด อย่างไรบ้าง

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

4. นักเรียนชอบการทำแบบทดสอบโดยใช้คอมพิวเตอร์หรือไม่ เพราะอะไร

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

5. ถ้าให้นักเรียนทำแบบทดสอบโดยเขียนคำตอบบนกระดาษกับทำโดยใช้คอมพิวเตอร์ นักเรียนจะเลือกอะไร เพราะอะไร

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....







ภาคผนวก ค

ตัวอย่างรายงานผลการประเมินเชิงวิจจัยจากการทดสอบแบบปรับเหมาะ  
โดยใช้คอมพิวเตอร์

## ตัวอย่างรายงานผลการประเมินโปรแกรมย่อยที่ 1

แบบทดสอบ: การบวกลบเศษส่วน ช่วงชั้นปีที่ 2 (ป.4-6)

ชื่อผู้ทดสอบ: ฝน

ชั้น/ห้อง: ป.4

โรงเรียน: บ้านปรือน้อย

วันที่: 8/10/2008

เวลา: 1:48:56

ผลการทดสอบ:

เงื่อนไขการทดสอบ: ถูก/ผิด 2 ใน 3 ข้อ

เวลาที่ใช้ในการทดสอบ: 0:09:11

ตำแหน่งเปอร์เซ็นต์ไทล์: 0

ความสามารถในการสอบ: -2.4190 (ต่ำ)

ความคลาดเคลื่อนมาตรฐาน: 0.4161

ข้อที่ ชูต / ข้อในข้อสอบ คำตอบ ผล แบบการคิดที่ผิด

---

1	1 / 25	- /7/22	ผิด	N/D
2	1 / 40	- /7/0	ผิด	N/D
3	1 / 33	- /18/12	ผิด	N/D
4	1 / 42	- /9/8	ผิด	N/D
5	1 / 43	- /12/12	ผิด	N/D
6	1 / 88	- /5/2	ผิด	N/D
7	1 / 100	- /6/3	ผิด	N/D
8	1 / 28	- /12/15	ผิด	1.1.2
9	1 / 91	- /13/16	ผิด	1.1.2

---



## ตัวอย่างรายงานผลการประเมินโปรแกรมย่อยที่ 2

แบบทดสอบ: การบวกลบเศษส่วน ช่วงชั้นปีที่ 2 (ป.4-6)

ชื่อผู้ทดสอบ: ทรัพย์ทวี ชวนชม

ชั้นห้อง: ป6

โรงเรียน: บ้านหนองยิงหมี่

วันที่: 3/10/2008

เวลา: 15:11:26

ผลการทดสอบ:

-----

เงื่อนไขการทดสอบ: ถูก/ผิด 3 ใน 4 ข้อ

เวลาที่ใช้ในการทดสอบ: 0:02:55

ตำแหน่งเปอร์เซ็นต์ไทล์: 4

ความสามารถในการสอบ: -1.5949 (ต่ำ)

ความคลาดเคลื่อนมาตรฐาน: 0.5661

ข้อที่ ชูต / ข้อในข้อสอบ คำตอบ ผล แบบการคิดที่ผิด

-----

1	1 / 25	- /3/0	ผิด	1.1.1
2	1 / 40	- /7/22	ผิด	1.1.1
3	1 / 33	- / 2 / 9	ผิด	N/D
4	1 / 42	- /3/14	ผิด	1.1.1

-----

### ตัวอย่างรายงานผลการประเมินโปรแกรมย่อยที่ 3

แบบทดสอบ: การบวกลบเศษส่วน ช่วงชั้นปีที่ 2 (ป.4-6)

ชื่อผู้ทดสอบ: ศุภวิชญ์ ศรีนวลจำปา

ชั้นห้อง: ป.6/1

โรงเรียน: อนุบาลชัยภูมิ

วันที่: 10/10/2008

เวลา: 13:27:09

ผลการทดสอบ:

เงื่อนไขการทดสอบ: ถูก/ผิด 4 ใน 5 ข้อ

เวลาที่ใช้ในการทดสอบ: 0:03:58

ตำแหน่งเปอร์เซ็นต์: 36

ความสามารถในการสอบ: -0.3392 (ปานกลาง)

ความคลาดเคลื่อนมาตรฐาน: 0.3453

ข้อที่ ชูด/ข้อในข้อสอบ คำตอบ ผล แบบการคิดที่ผิด

1	1 / 25	-/3/7	ถูก	
2	1 / 85	-/2/7	ถูก	
3	1 / 30	-/5/6	ถูก	
4	1 / 93	-/1/8	ถูก	
5	1 / 32	-/6/7	ถูก	
6	2 / 172	-/9/9	ผิด	2.1.1
7	2 / 146	-/6/7	ผิด	2.1.1
8	2 / 212	-/5/13	ผิด	2.2.1
9	2 / 131	-/1/5	ผิด	N/D
10	2 / 137	-/2/2	ผิด	2.1.2
11	2 / 169	-/1/6	ผิด	2.2.1
12	2 / 186	-/8/9	ผิด	2.1.1
13	2 / 185	-/0/9	ผิด	2.1.1
14	3 / 258	1/2/7	ถูก	
15	3 / 248	1/2/7	ถูก	
16	3 / 280	2/4/8	ผิด	N/D
17	3 / 244	2/4/10	ผิด	N/D
18	3 / 228	1/2/5	ถูก	
19	3 / 299	2/5/9	ถูก	

## ตัวอย่างรายงานผลการประเมินโปรแกรมย่อยที่ 4

แบบทดสอบ: การบวกลบเศษส่วน ช่วงชั้นปีที่ 2 (ป.4-6)

ชื่อผู้ทดสอบ: เด็กชาย ศุภกิจ จันทร์ทร

ชั้นห้อง: ป.6/4

โรงเรียน: อนุบาลชัยภูมิ

วันที่: 10/10/2008

เวลา: 14:04:47

ผลการทดสอบ:

เงื่อนไขการทดสอบ: ถูก/ผิด 2 ใน 3 ข้อและ SEE<0.3

เวลาที่ใช้ในการทดสอบ: 0:16:10

ตำแหน่งเปอร์เซ็นต์: 11

ความสามารถในการสอบ: -1.1886 (ต่ำ)

ความคลาดเคลื่อนมาตรฐาน: 0.2948

ข้อที่ ชูด/ข้อในข้อสอบ คำตอบ ผล แบบการคิดที่ผิด

---

1	1 / 25	- /3/7	ถูก	
2	1 / 40	- /7/11	ถูก	
3	1 / 85	- /2/7	ถูก	
4	2 / 139	- /1/8	ถูก	
5	2 / 162	- /11/10	ผิด	N/D
6	2 / 136	- /22/16	ผิด	N/D
7	2 / 131	- /3/10	ถูก	
8	3 / 288	- /46/9	ผิด	N/D
9	3 / 294	- /56/10	ผิด	3.2.1
10	3 / 254	- /10/9	ผิด	N/D
11	3 / 252	- /8/7	ผิด	N/D
12	3 / 270	- /13/5	ผิด	N/D
13	3 / 248	- /9/7	ผิด	N/D
14	3 / 305	- /131/10	ผิด	N/D
15	3 / 231	- /58/9	ผิด	N/D
16	3 / 261	- /43/7	ผิด	N/D
17	3 / 242	- /22/7	ผิด	N/D
18	3 / 304	- /70/7	ผิด	N/D
19	3 / 303	- /56/5	ผิด	N/D
20	3 / 325	- /53/7	ผิด	N/D
21	3 / 264	- /11/10	ผิด	3.2.1

---

## ตัวอย่างรายงานผลการประเมินโปรแกรมย่อยที่ 5

แบบทดสอบ: การบวกลบเศษส่วน ช่วงชั้นปีที่ 2 (ป.4-6)

ชื่อผู้ทดสอบ: เด็กหญิงกนกพร เสนโปพาน

ชั้น/ห้อง: ชั้น ป.5/1

โรงเรียน: อนุบาลชัยภูมิ

วันที่: 10/10/2008

เวลา: 9:17:27

ผลการทดสอบ:

เงื่อนไขการทดสอบ: ถูก/ผิด 3 ใน 4 ข้อและ SEE<0.3

เวลาที่ใช้ในการทดสอบ: 0:20:18

ตำแหน่งเปอร์เซ็นต์: 9

ความสามารถในการสอบ: -1.2920 (ต่ำ)

ความคลาดเคลื่อนมาตรฐาน: 0.2763

ข้อที่ ชัด / ข้อในข้อสอบ คำตอบ ผล แบบการคิดที่ผิด

---

1	1 / 25	-/3/7	ถูก	
2	1 / 33	-/6/11	ถูก	
3	1 / 85	-/2/7	ถูก	
4	1 / 30	-/5/6	ถูก	
5	2 / 144	-/25/75	ผิด	N/D
6	2 / 212	-/5/13	ผิด	2.2.1
7	2 / 187	-/3/6	ผิด	N/D
8	2 / 169	-/1/9	ถูก	
9	2 / 137	-/1/2	ผิด	N/D
10	2 / 200	-/3/11	ผิด	2.2.1
11	2 / 201	-/2/8	ผิด	N/D
12	2 / 165	-/5/1	ผิด	2.2.1
13	3 / 228	-/3/5	ผิด	3.1.2
14	3 / 231	6/4/9	ถูก	
15	3 / 275	1/4/5	ผิด	3.1.2
16	3 / 242	2/6/7	ผิด	N/D
17	3 / 304	10/ - / -	ถูก	
18	3 / 314	-/6/8	ผิด	N/D
19	3 / 326	1/6/8	ผิด	N/D
20	3 / 230	2/5/7	ผิด	N/D
21	3 / 325	9/ - / -	ผิด	N/D
22	3 / 264	-/9/10	ผิด	N/D
23	3 / 301	4/5/7	ผิด	N/D
24	3 / 260	-/5/9	ผิด	3.1.2

---

## ตัวอย่างรายงานผลการประเมินโปรแกรมย่อยที่ 6

แบบทดสอบ: การบวกลบเศษส่วน ช่วงชั้นปีที่ 2 (ป.4-6)

ชื่อผู้ทดสอบ: ธนระวีชาติ กว่างสาสดี

ชั้นห้อง: ห้อง ป6

โรงเรียน: บ้านหนองยิงหมี่

วันที่: 3/10/2008

เวลา: 15:21:42

ผลการทดสอบ:

เงื่อนไขการทดสอบ: ถูก/ผิด 4 ใน 5 ข้อและ SEE<0.3

เวลาที่ใช้ในการทดสอบ: 0:02:55

ตำแหน่งเปอร์เซ็นต์: 0

ความสามารถในการสอบ: -3.4294 (ต่ำ)

ความคลาดเคลื่อนมาตรฐาน: 0.2610

ข้อที่ ชูค/ข้อในข้อสอบ คำตอบ ผล แบบการคิดที่ผิด

1	1 / 25	3/-/-	ผิด	N/D
2	1 / 40	12/-/-	ผิด	N/D
3	1 / 33	6/-/-	ผิด	N/D
4	1 / 42	8/-/-	ผิด	N/D
5	1 / 43	6/-/-	ผิด	N/D
6	1 / 88	4/-/-	ผิด	N/D
7	1 / 100	6/-/-	ผิด	N/D
8	1 / 28	12/-/-	ผิด	N/D
9	1 / 91	13/-/-	ผิด	N/D
10	1 / 27	5/-/-	ผิด	N/D
11	1 / 64	0/-/-	ผิด	N/D
12	1 / 5	13/-/-	ผิด	N/D
13	1 / 92	14/-/-	ผิด	N/D
14	1 / 68	1/-/-	ผิด	N/D
15	1 / 83	6/-/-	ผิด	N/D
16	1 / 48	-/4/0	ผิด	1.1.1
17	1 / 3	11/-/-	ผิด	N/D
18	1 / 4	10/-/-	ผิด	N/D
19	1 / 10	-/2/0	ผิด	1.1.1
20	1 / 89	-/6/22	ผิด	1.1.1
21	1 / 14	5/-/-	ผิด	N/D
22	1 / 16	-/3/0	ผิด	1.1.1

## ภาคผนวก ง

คู่มือการประเมินเชิงวินิจฉัยสถานะความรู้

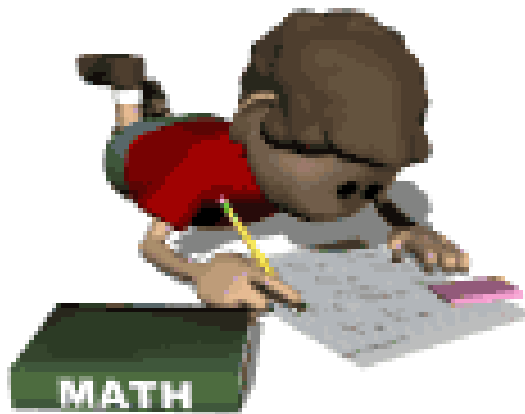
และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน

ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6)

โดยใช้เทคนิคการคิดออกเสียง

# คู่มือ

การประเมินเชิงวินิจฉัยสถานะความรู้  
และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน  
ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6)  
โดยใช้เทคนิคการคิดออกเสียง



จัดทำโดย

นางอมรรัตน์ สร้อยสังวาลย์

คู่มือนี้เป็นส่วนหนึ่งของการศึกษาระดับปริญญาโท  
สาขาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา  
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

## คำนำ

คู่มือการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) โดยใช้เทคนิคการคิดออกเสียงฉบับนี้ จัดทำขึ้นเพื่อสร้างความรู้ความเข้าใจให้กับผู้ประเมินในการใช้เทคนิคการคิดออกเสียงเป็นเครื่องมือเพื่อการประเมินเชิงวิจักษ์ ซึ่งเป็นส่วนหนึ่งของการดำเนินการจัดทำวิทยานิพนธ์ระดับปริญญาเอก สาขาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เรื่อง การพัฒนาวิธีการประเมินเชิงวิจักษ์โดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ผลการประเมินในครั้งนี้จะนำไปใช้เพื่อศึกษาความตรงของวิธีการประเมินที่พัฒนาขึ้น โดยศึกษาจากความสอดคล้องระหว่างผลการประเมินจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์กับผลการประเมินจากการทดสอบโดยใช้เทคนิคการคิดออกเสียง

เนื้อหาที่นำเสนอในคู่มือจำแนกเป็น 3 ส่วน ส่วนที่ 1 บทนำ จะกล่าวถึงการประเมินเชิงวิจักษ์และความเป็นมาของการนำเทคนิคการคิดออกเสียงไปใช้เพื่อการประเมินเชิงวิจักษ์ในครั้งนี้ ส่วนที่ 2 มโนทัศน์เบื้องต้นเกี่ยวกับเทคนิคการคิดออกเสียง จะกล่าวถึงความเป็นมาของเทคนิคการคิดออกเสียง รวมทั้งลักษณะและรูปแบบการนำเทคนิคการคิดออกเสียงไปใช้ ส่วนที่ 3 การดำเนินการประเมินเชิงวิจักษ์โดยใช้เทคนิคการคิดออกเสียง จะกล่าวถึงวิธีการดำเนินการประเมินเชิงวิจักษ์โดยใช้เทคนิคการคิดออกเสียง ได้แก่ การกำหนดผู้ได้รับการประเมิน การเตรียมความพร้อมในการประเมินและขั้นตอนการดำเนินการประเมิน

ผู้จัดทำหวังเป็นอย่างยิ่งว่าคู่มือฉบับนี้จะช่วยสร้างความรู้ความเข้าใจเกี่ยวกับการใช้เทคนิคการคิดออกเสียงในการประเมินเชิงวิจักษ์ให้กับผู้ประเมิน และช่วยให้การประเมินในครั้งนี้ดำเนินการไปอย่างมีประสิทธิภาพและเกิดประโยชน์สูงสุด

นางอมรรัตน์ สร้อยสังวาลย์

ผู้จัดทำ

กรกฎาคม 2551



## สารบัญ

	หน้า
บทนำ.....	1
มโนทัศน์เบื้องต้นเกี่ยวกับเทคนิคการคิดออกเสียง.....	1
การดำเนินการประเมินเชิงวินิจฉัย โดยใช้เทคนิคการคิดออกเสียง.....	3
บรรณานุกรม.....	9
ภาคผนวก ก (1) แบบทดสอบ เรื่อง การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน.....	11
ภาคผนวก ก (2) แบบทดสอบ เรื่อง การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน.....	12
ภาคผนวก ก (3) แบบทดสอบ เรื่อง การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน.....	13
ภาคผนวก ก (4) แบบทดสอบ เรื่อง การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน.....	14
ภาคผนวก ข   แบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษา ปีที่ 4 – 6).....	15
ภาคผนวก ค   แบบบันทึกกระบวนการคิดแก้ปัญหาในการประเมินเชิงวินิจฉัยสถานะความรู้ และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษา ปีที่ 4 – 6).....	22

## บทนำ

การประเมินเชิงวินิจฉัย (diagnostic assessment) เป็นการรวบรวมข้อมูลเพื่อตัดสินจุดเด่น จุดด้อย ปัญหาอุปสรรค รวมทั้งสาเหตุที่ทำให้เกิดปัญหาอุปสรรคเหล่านั้น โดยมีเป้าหมายเพื่อการปรับปรุงแก้ไขและการพัฒนา สารสนเทศที่ได้จากการประเมินจะเป็นแนวทางให้ครูหาวิธีการในการจัดการเรียนรู้ หรือมาตรการในการจัดซ่อมเสริมให้กับผู้เรียนเพื่อให้ผู้เรียนประสบความสำเร็จในการเรียนรู้และสามารถพัฒนาตนเองได้อย่างเต็มตามศักยภาพ การประเมินเชิงวินิจฉัยสามารถดำเนินการได้หลายรูปแบบ ด้วยวิธีการที่หลากหลาย ทั้งรูปแบบที่เป็นทางการ เช่น การใช้แบบทดสอบเชิงวินิจฉัย การวินิจฉัยจากแบบแผนการตอบข้อสอบ หรือการใช้การย้อนรอยกระบวนการคิด เป็นต้น และรูปแบบที่ไม่เป็นทางการ เช่น การสังเกต การสัมภาษณ์ หรือการตรวจผลงาน เป็นต้น การประเมินเชิงวินิจฉัยที่มีประสิทธิภาพจะต้องให้สารสนเทศเพื่อการพัฒนาผู้เรียนและการจัดการเรียนรู้ที่ถูกต้อง ชัดเจน วิธีการหนึ่งซึ่งได้รับการพัฒนาจนเป็นที่ยอมรับว่าสามารถช่วยให้เห็นถึงกระบวนการคิดของนักเรียนที่ใช้ในการปฏิบัติงานได้อย่างชัดเจน โดยการแสดงออกทางคำพูด ก็คือ เทคนิคการคิดออกเสียง

การประเมินเชิงวินิจฉัยสถานะความรู้ และแบบการคิดที่ผิดในการบวกลบเศษส่วนระดับช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) ในครั้งนี้ จึงเลือกใช้เทคนิคการคิดออกเสียงเป็นเครื่องมือในการประเมิน โดยมีวัตถุประสงค์เพื่อนำผลการประเมินที่ได้ไปเปรียบเทียบกับผลการประเมินจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ เพื่อศึกษาความตรงของวิธีการประเมินเชิงวินิจฉัยที่พัฒนาขึ้น ซึ่งเป็นส่วนหนึ่งของการจัดทำวิทยานิพนธ์ ระดับปริญญาเอก สาขาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในหัวข้อเรื่อง การพัฒนาวิธีการประเมินเชิงวินิจฉัยโดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

## มโนทัศน์เบื้องต้นเกี่ยวกับเทคนิคการคิดออกเสียง

เทคนิคการคิดออกเสียง (think-aloud technique) เป็นวิธีการหรือกระบวนการที่จัดทำขึ้นเพื่อให้บุคคลรายงานความคิดของตนออกมาเป็นคำพูด มักใช้กับการปฏิบัติงานเชิงวิชาการที่มีเป้าหมายเฉพาะ (performance task) เช่น งานที่เกี่ยวข้องกับการแก้ปัญหา การตอบคำถาม การทำการทดลอง การอ่านบทประพันธ์หรือบทเรียน เป็นต้น เทคนิคการคิดออกเสียงพัฒนามาจากวิธีการทบทวนความคิดของตน (introspection method) ซึ่งเป็นวิธีการที่นักจิตวิทยาใช้ในการศึกษาสิ่งที่อยู่ภายในจิตของมนุษย์ เช่น กระบวนการทางพุทธิปัญญา ความรู้สึก ความสนใจ

รูปแบบการเรียนรู้ เป็นต้น โดยพยายามค้นหาเหตุการณ์ที่แสดงถึงสิ่งที่อยู่ภายในจิตเหล่านั้น แล้วนำมาอธิบาย ดีความและสรุป วิธีการทบทวนความคิดของตนดังกล่าวมีข้อจำกัดในเรื่องของความเชี่ยวชาญเฉพาะของผู้ศึกษาทำให้ไม่ได้รับการยอมรับเชื่อถือในประเด็นความเที่ยงและความตรง เทคนิคการคิดออกเสียงจึงพัฒนาขึ้นเพื่อแก้ข้อจำกัดดังกล่าว โดยหลีกเลี่ยงการตีความข้อมูลจากบุคคลใดบุคคลหนึ่ง ด้วยการทำความเข้าใจสิ่งที่บุคคลพูดเพื่อรายงานกระบวนการคิด ซึ่งส่วนใหญ่จะเป็นคำพูดง่าย ๆ ที่บุคคลทั่วไปสามารถเข้าใจได้ ไม่จำเป็นต้องเป็นผู้ที่มีความเชี่ยวชาญเฉพาะเท่านั้น

เทคนิคการคิดออกเสียงเป็นเทคนิคที่เหมาะสมกับงานการแก้ปัญหา (problem-solving task) ซึ่งถูกนำไปประยุกต์ใช้ในศาสตร์ต่าง ๆ ด้วยวัตถุประสงค์ที่แตกต่างกัน จำแนกได้ 2 ลักษณะดังนี้

ลักษณะที่ 1 นำไปใช้เป็นเทคนิคการจัดการเรียนรู้เพื่อพัฒนาผู้เรียน การนำไปใช้ในลักษณะนี้ โดยทั่วไปจะมีการแสดงตัวอย่างการคิดออกเสียงจากครู และให้ผู้เรียนฝึกปฏิบัติตามวิธีการนี้จะช่วยให้ผู้เรียนเห็นตัวอย่างกระบวนการคิดที่ถูกต้อง ซึ่งเมื่อผู้เรียนนำไปฝึกปฏิบัติในชั้นเรียน ผู้เรียนก็จะได้ฝึกเชื่อมโยงความคิดของตนกับการรายงานออกมาเป็นคำพูด รวมทั้งได้เห็นตัวอย่างการคิดทั้งที่ถูกและที่ผิดจากเพื่อน ๆ การฝึกปฏิบัติและการแลกเปลี่ยนเรียนรู้ระหว่างเพื่อนดังกล่าวจึงมีบทบาทสำคัญยิ่งต่อการพัฒนาทักษะการใช้ภาษาและการพัฒนากระบวนการคิดของผู้เรียน ให้มีความถูกต้องและเป็นระบบมากขึ้น

ลักษณะที่ 2 นำไปใช้เพื่อการประเมิน ทั้งประโยชน์ในการจัดการเรียนรู้และในการพัฒนาเครื่องมือการประเมินหรือเครื่องมือวิจัย ในกรณีของการนำไปใช้เพื่อการประเมินในการจัดการเรียนรู้ ครูจะออกแบบงานและให้ผู้เรียนบรรยายกระบวนการคิดของตนออกมาเป็นคำพูด ในทุกขั้นตอนของกระบวนการทำงานเพื่อให้งานนั้นสำเร็จ คำพูดดังกล่าวอาจเกี่ยวกับการปฏิบัติงาน การคาดเดาวิธีที่ใช้ในการหาคำตอบ หรือความรู้สึกนึกคิดต่าง ๆ ครูมีบทบาทเป็นผู้ฟัง ผู้เก็บรวบรวมข้อมูล ผู้กระตุ้นเตือนเมื่อผู้เรียนเงียบ หรือให้กำลังใจเมื่อผู้เรียนท้อแท้ เบื่อหน่าย แต่ต้องระวังไม่รบกวน สร้างความกดดันหรือถามคำถามที่ชี้แนะคำตอบหรือแนวทางการปฏิบัติให้กับผู้เรียน การเก็บรวบรวมข้อมูลอาจใช้การบันทึกบนกระดาษ (paper-pencil) หรือการบันทึกภาพและเสียง โดยใช้เทปบันทึกภาพและเสียงหรือคอมพิวเตอร์ ทั้งนี้เพื่อให้สามารถเก็บรวบรวมข้อมูลการประเมินได้มากที่สุด ครอบคลุมที่สุด การใช้เทคนิคการคิดออกเสียงเพื่อการประเมินในการจัดการเรียนรู้ นอกจากจะให้ผลการประเมินที่เป็นระดับความสำเร็จในการปฏิบัติงานของผู้เรียนแล้ว ยังทำให้ครูทราบถึงมโนทัศน์ที่คลาดเคลื่อนหรือแบบการคิดที่ผิดของผู้เรียน รวมทั้งความคิดเห็นและความรู้สึกอื่น ๆ ด้วย แต่ก็มีข้อจำกัดในประเด็นเวลาที่ใช้ในการประเมิน ซึ่งต้องใช้เวลามาก รวมทั้งไม่สามารถดำเนินการกับผู้เรียนกลุ่มใหญ่ได้ ส่วนกรณีของการนำไปใช้เพื่อการประเมินใน

การพัฒนาเครื่องมือการประเมินหรือการวิจัยนั้นจะใช้ในขั้นตอนการหาคุณภาพเครื่องมือ ในด้านความตรง โดยการนำเครื่องมือการประเมินหรือเครื่องมือวิจัยไปให้กลุ่มทดลองซึ่งมีลักษณะเหมือนกลุ่มตัวอย่างตอบคำถามหรือปฏิบัติงานตามรูปแบบโครงสร้างของเครื่องมือการประเมินหรือเครื่องมือการวิจัยนั้น ๆ ผู้พัฒนาเครื่องมือจะเก็บรวบรวมข้อมูลที่ได้จากการปฏิบัติดังกล่าวเพื่อนำไปปรับปรุงแก้ไขเครื่องมือให้มีความตรงต่อไป

การใช้เทคนิคการคิดออกเสียงดำเนินการได้ใน 2 รูปแบบ ดังนี้

รูปแบบที่ 1 การคิดออกเสียงเพื่อทบทวนสิ่งที่ผ่านไป (retrospective think-aloud) ดำเนินการโดยกำหนดให้บุคคลรายงานกระบวนการคิดที่ใช้ในการปฏิบัติงานด้วยคำพูด หลังจากปฏิบัติงานแล้ว วิธีการนี้มีจุดเด่นตรงที่สามารถจำกัดเวลาที่ใช้ในการปฏิบัติงานได้ ไม่มีการรบกวนระหว่างการปฏิบัติงานซึ่งช่วยลดข้อจำกัดในเรื่องอารมณ์ของผู้ปฏิบัติงานที่มีต่อผลของการปฏิบัติงาน เช่น ความวิตกกังวล ความกดดัน ความเบื่อหน่าย ความรำคาญใจ แต่รูปแบบนี้ก็มีข้อจำกัดตรงที่ผู้ปฏิบัติงานอาจมีความจำเกี่ยวกับกระบวนการคิดที่ใช้ในการปฏิบัติงานคลาดเคลื่อน สัมภาษณ์กระบวนการคิดในบางขั้นตอน หรือไม่ได้ใช้กระบวนการคิดนั้นระหว่างการปฏิบัติงานแต่เป็นกระบวนการคิดที่เกิดขึ้นจากการทบทวนการปฏิบัติงานเมื่อการปฏิบัติงานเสร็จสิ้นแล้ว

รูปแบบที่ 2 การคิดออกเสียงในระหว่างการปฏิบัติงาน (concurrent think-aloud) ดำเนินการโดยกำหนดให้บุคคลรายงานกระบวนการคิดด้วยคำพูดในทุกขั้นตอนระหว่างการปฏิบัติงาน วิธีการนี้มีจุดเด่นตรงที่ได้ข้อมูลเกี่ยวกับกระบวนการคิดที่ใช้ในการปฏิบัติงานที่แท้จริง ไม่มีความคลาดเคลื่อนที่เกิดจากการลืมหรือกระบวนการคิดที่สอดแทรกขึ้นหลังจากการปฏิบัติงาน แต่มีข้อจำกัดตรงที่การดำเนินการอาจมีผลทางอารมณ์ต่อผู้ปฏิบัติงาน เช่น คำถามต่าง ๆ อาจสร้างความวิตกกังวล ความกดดัน ความเบื่อหน่าย ความรำคาญใจให้กับผู้ปฏิบัติงาน หรืออาจชี้แนะคำตอบหรือแนวทางการปฏิบัติงานให้กับผู้ปฏิบัติงานได้ รวมทั้งไม่สามารถกำหนดเวลาที่ใช้ในการปฏิบัติงานที่แน่นอนได้

ที่กล่าวมาจะเห็นได้ว่าการนำเทคนิคการคิดออกเสียงไปใช้สามารถดำเนินการได้ในลักษณะและรูปแบบที่แตกต่างกัน การนำไปใช้เพื่อให้เกิดประโยชน์สูงสุดจึงต้องคำนึงถึงเป้าหมายและวัตถุประสงค์เฉพาะของการดำเนินการนั้น ๆ เป็นสำคัญ

## การดำเนินการประเมินเชิงวิจจัยโดยใช้เทคนิคการคิดออกเสียง

สำหรับการใช้เทคนิคการคิดออกเสียงเพื่อการประเมินเชิงวิจจัยสถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วนในครั้งนี้ ผู้วิจัยกำหนดให้เลือกใช้การดำเนินการในลักษณะเป็นเทคนิคที่ใช้ประโยชน์เพื่อการประเมินในการจัดการเรียนรู้ โดยเลือกใช้รูปแบบการคิดออกเสียงในระหว่างการปฏิบัติงาน มีการดำเนินการ ดังนี้

### 1. ผู้ได้รับการประเมิน

ผู้ได้รับการประเมินเป็นนักเรียนระดับการศึกษาขั้นพื้นฐาน (ชั้นประถมศึกษาปีที่ 4 – 6) ซึ่งได้จากการสุ่มตัวอย่างอย่างง่าย จากนักเรียนที่เป็นกลุ่มตัวอย่างในการประเมินเชิงวิจจัยสถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ จำนวน 60 คน

### 2. การเตรียมความพร้อมสำหรับการประเมิน

การเตรียมความพร้อมสำหรับการประเมินเป็นการดำเนินการก่อนการประเมินเพื่อให้การประเมินดำเนินการสำเร็จลุล่วงไปได้ด้วยดี และป้องกันปัญหาอุปสรรคที่อาจเกิดขึ้นระหว่างการดำเนินการให้ได้มากที่สุด การเตรียมความพร้อมสำหรับการประเมิน ประกอบด้วย การเตรียมความพร้อมของเครื่องมือการประเมิน การเตรียมความพร้อมของนักเรียน การเตรียมความพร้อมของผู้ประเมินและการเตรียมความพร้อมของสถานที่ที่ใช้ในการประเมิน มีรายละเอียด ดังนี้

#### 2.1 การเตรียมความพร้อมของเครื่องมือการประเมิน

เครื่องมือที่ใช้ในการประเมินครั้งนี้ ประกอบด้วย

##### 2.1.1 แบบทดสอบแบบเขียนตอบ จำนวน 4 ฉบับ ฉบับละ 10 ข้อ ดังนี้

ฉบับที่ 1 การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน

ฉบับที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน

ฉบับที่ 3 การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน

ฉบับที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน

2.1.2 แบบบันทึกกระบวนการคิดแก้ปัญหา เป็นแบบบันทึกที่ให้ผู้ประเมินใช้เพื่อบันทึกกระบวนการคิดแก้ปัญหาของนักเรียนในขณะที่นักเรียนทำแบบทดสอบและสรุปผลการวิจจัยสถานะความรู้หรือแบบการคิดที่ผิด รวมทั้งข้อสังเกต ข้อเสนอแนะและความคิดเห็นอื่น ๆ

2.1.3 เครื่องบันทึกเสียง ใช้สำหรับบันทึกเสียงพูดบรรยายกระบวนการคิดของนักเรียนในขณะที่นักเรียนทำแบบทดสอบ เพื่อเป็นข้อมูลประกอบการพิจารณาวินิจฉัยสถานะ ความรู้และแบบการคิดที่ผิด

## 2.2 การเตรียมความพร้อมของนักเรียน

เทคนิคการคิดออกเสียงเป็นเทคนิคที่พัฒนาขึ้นเพื่อใช้กับผู้ใหญ่ ซึ่งใช้ได้ดีกับเด็กอายุ 8-14 ปี แต่ผลจากการศึกษาการใช้เทคนิคการคิดออกเสียงกับนักเรียนในบริบทสังคมไทยพบว่า นักเรียนส่วนใหญ่ไม่กล้าพูดและไม่ได้รับการฝึกกระบวนการคิดเท่าที่ควร ทำให้พูดบรรยายกระบวนการคิดได้ไม่ตรงกับสิ่งที่คิด ดังนั้น การใช้เทคนิคการคิดออกเสียงในครั้งนี้ จึงควรฝึกการรายงานกระบวนการคิด โดยใช้เทคนิคการคิดออกเสียงแก่นักเรียนก่อนการดำเนินการจริง เพื่อให้ นักเรียนเข้าใจวิธีการบรรยายกระบวนการคิดโดยใช้คำพูดมากที่สุด เริ่มจากการชี้แจงวัตถุประสงค์และอธิบายวิธีการพูดบรรยายกระบวนการคิดให้กับนักเรียน จนนักเรียนเกิดความเข้าใจ โดยพยายามสร้างบรรยากาศเป็นกันเอง เพื่อไม่สร้างความเครียดหรือความวิตกกังวลให้กับนักเรียน

หลังจากนั้น จึงฝึกการรายงานกระบวนการคิด โดยใช้เทคนิคการคิดออกเสียง เริ่มจากให้นักเรียนดูตัวอย่างคำถามและอ่านออกเสียงดัง ๆ ชัด ๆ หลังจากนั้นให้นักเรียนคิดวิธีการหาคำตอบสำหรับคำถามนั้น และพูดสิ่งที่คิดออกมาให้มากที่สุด แม้ว่าจะไม่ใช่วิธีการหาคำตอบระหว่างที่นักเรียนคิดและพูด ผู้ประเมินจะต้องไม่กระทำการใด ๆ ซึ่งมีผลทางลบต่อการคิดและการพูดของนักเรียน ระหว่างใช้เทคนิคการคิดออกเสียงอาจมีกรณีต่าง ๆ เกิดขึ้น ซึ่งผู้ประเมินมีแนวดำเนินการ ดังนี้

2.2.1 กรณีที่นักเรียนอ่านคำถามเสร็จแล้ว แต่ไม่พูด ผู้ประเมินอาจกระตุ้นนักเรียนโดยการใช้คำถาม เช่น “นักเรียนอ่านโจทย์แล้วคิดอย่างไรบ้าง พูดให้ครูฟังบ้างสิ ไม่ต้องกลัวว่าจะถูกหรือผิด”

2.2.2 กรณีที่นักเรียนพูดไปแล้วหยุดหรือเงียบติดต่อกันเป็นเวลานาน (ประมาณ 5 – 10 วินาที) ครูอาจถามเพื่อกระตุ้นให้นักเรียนพูดสิ่งที่คิดอย่างต่อเนื่อง เช่น “แล้วยังไงอีก พูดต่อสิ ครูกำลังฟังอยู่” หรืออาจใช้สีหน้าหรือท่าทางแสดงให้นักเรียนรู้ว่าครูกำลังฟังอยู่ พูดต่อได้เลย แต่การใช้สีหน้าท่าทางจะต้องระวังไม่ให้เป็นการกดดันนักเรียนจนนักเรียนเกิดความกลัวหรือวิตกกังวล

2.2.3 กรณีที่นักเรียนพูดนอกประเด็นมากเกินไป จนไม่สามารถพิจารณากระบวนการคิดของนักเรียนได้ ครูอาจจะใช้คำพูดเพื่อให้นักเรียนหยุดพูดและทบทวนกระบวนการคิดอีกครั้ง เช่น “แล้วคำถามนี้จะหาคำตอบได้ยังไงนะ” หรือ “ขั้นตอนนี้นักเรียนคิดยังไง ทำยังไงได้บ้าง”

2.2.4 กรณีที่นักเรียนพูดเร็วเกินไป จนผู้ประเมินฟังไม่ทันหรือไม่สามารถจับประเด็นได้ ผู้ประเมินอาจขอให้นักเรียนพูดช้าลง เช่น “พูดช้าลงหน่อยนะ ตรงนี้ได้มายังไงนะ ครูฟังไม่ทัน” หรือ “ลองอธิบายขั้นตอนนี้ให้ครูฟังอีกครั้งซิ”

2.2.5 กรณีที่นักเรียนบอกว่า “คิดไม่ออก” “ทำไม่ได้” “ไม่รู้จะอธิบายอย่างไร” ผู้ประเมินจะต้องพูดให้กำลังใจนักเรียน เช่น “ลองคิดอีกทีซิ ค่อย ๆ คิด ไม่ต้องรีบ คิดยังไงก็พูดออกมา พูดอะไรก็ได้” หรือ “ลองดูคำถามอีกครั้งก็ได้ นักเรียนน่าจะเคยเรียนมาแล้วนะ ลองคิดดูซิ”

2.2.6 กรณีที่นักเรียนพูดบรรยายข้ามขั้นตอน ซึ่งอาจทำให้ผู้ประเมินไม่สามารถวินิจฉัยสถานะความรู้หรือแบบการคิดที่ผิดของนักเรียนได้หรือวินิจฉัยได้แต่ไม่ชัดเจน ผู้ประเมินอาจถามคำถามเพื่อให้นักเรียนย้อนกลับไปคิด หรือย้อนกลับไปพูดบรรยายกระบวนการคิดนั้นอีกครั้ง เช่น “ก่อนได้ผลลัพธ์นี้ ต้องทำยังไง” หรือ “ได้ผลลัพธ์นี้มาได้ยังไง”

2.2.7 กรณีที่นักเรียนมีกระบวนการคิดที่ถูกต้องทุกขั้นตอน แต่ได้คำตอบผิด ซึ่งอาจเกิดจากการคิดคำนวณ ครูอาจใช้คำถามเพื่อให้แน่ใจว่านักเรียนคิดคำนวณผิด เนื่องจากไม่มีความสามารถในการคิดคำนวณหรือเป็นความไม่ระมัดระวัง รีบคิดเพื่อให้ได้คำตอบ เช่น “ได้คำตอบเท่านี้เหรอ ลองคิดอีกครั้งซิ”

2.2.8 กรณีที่นักเรียนไม่สามารถพูดบรรยายกระบวนการคิดได้ ให้ผู้ประเมินพิจารณาว่าสาเหตุเกิดจากอะไรและพยายามแก้ไขปัญหามาจากสาเหตุเหล่านั้น หรือลองฝึกการรายงานกระบวนการคิด โดยใช้เทคนิคการคิดออกเสียงให้กับนักเรียนอีกครั้ง จนนักเรียนเข้าใจวิธีการและสามารถพูดบรรยายกระบวนการคิดได้

### 2.3 การเตรียมความพร้อมของผู้ประเมิน

ผู้ประเมินในครั้งนี้นี้ประกอบด้วยครูผู้สอนวิชาคณิตศาสตร์ของนักเรียนที่ได้รับการประเมินและผู้วิจัย เพื่อให้ผู้ประเมินทุกคนมีความเข้าใจในวิธีการประเมินตรงกัน จึงมีขั้นตอนในการเตรียมความพร้อมให้กับผู้ประเมิน ดังนี้

2.3.1 ให้ผู้ประเมินศึกษาคู่มือการประเมินเชิงวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน (ชั้นประถมศึกษาปีที่ 4 – 6) โดยใช้เทคนิคการคิดออกเสียง และแบบบันทึกกระบวนการคิดแก้ปัญหา ก่อนดำเนินการประเมินอย่างน้อย 1 สัปดาห์

2.3.2 ผู้ประเมินสุ่มเลือกนักเรียนที่ได้รับการประเมินเชิงวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน (ชั้นประถมศึกษาปีที่ 4 – 6) จากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ที่ไม่ได้เป็นกลุ่มตัวอย่างสำหรับการประเมินครั้งนี้

จำนวน 1 – 2 คน เพื่อทดลองประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิด โดยใช้เทคนิคการคิดออกเสียง

2.3.3 หลังจากผู้ประเมินดำเนินการทดลองประเมินแล้ว ผู้ประเมินและผู้วิจัยจะร่วมกันพิจารณาแบบบันทึกกระบวนการคิดแก้ปัญหาที่ผู้ประเมินบันทึกไว้ กรณีที่ผู้ประเมินไม่เข้าใจหรือเกิดปัญหาระหว่างการดำเนินการ ผู้ประเมินและผู้วิจัยจะทำความเข้าใจในขั้นตอนและวิธีการประเมิน พร้อมทั้งวางแผนการประเมินร่วมกัน

#### 2.4 การเตรียมสถานที่ที่ใช้ในการประเมิน

การใช้เทคนิคการคิดออกเสียงให้ได้ผลดี นักเรียนจะต้องมีสมาธิในการคิดและถ่ายถอดกระบวนการคิดออกมาเป็นคำพูด ดังนั้นสถานที่ที่ใช้ในการประเมินจึงต้องเป็นสถานที่ที่เงียบ ไม่มีสิ่งรบกวน รวมทั้งมีบรรยากาศที่ผ่อนคลาย เพื่อช่วยลดความกดดันและความวิตกกังวลของนักเรียน สถานที่ที่ใช้ในการประเมินจึงอาจเป็นห้องเรียน ห้องปฏิบัติการอื่น ๆ หรือที่นั่งใดก็ได้รวมทั้งนี้ขึ้นอยู่กับความพร้อมของแต่ละโรงเรียน

### 3. ขั้นตอนการดำเนินการประเมิน

การประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนระดับการศึกษาขั้นพื้นฐาน (ชั้นประถมศึกษาปีที่ 4 – 6) มีการดำเนินการตามขั้นตอนต่าง ๆ ดังนี้

3.1 ก่อนดำเนินการประเมิน จะต้องตรวจสอบเครื่องมือที่ใช้ในการประเมิน และสถานที่ที่ใช้ในการประเมิน รวมทั้งนัดหมายกับนักเรียนที่จะได้รับการประเมินให้พร้อมสำหรับการประเมิน

3.2 ชี้แจงวัตถุประสงค์และอธิบายวิธีการพูดบรรยายกระบวนการคิดให้กับนักเรียน จนนักเรียนเกิดความเข้าใจ

3.3 ฝึกการรายงานกระบวนการคิด โดยใช้เทคนิคการคิดออกเสียงให้แก่นักเรียนก่อนเริ่มการดำเนินการ

3.4 เริ่มบันทึกเสียงและดำเนินการประเมินตามขั้นตอนการรายงานความคิดที่ได้ฝึกให้กับนักเรียน โดยให้นักเรียนดูคำถามจากแบบทดสอบฉบับที่ 1 การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน ทั้ง 10 ข้อ และให้นักเรียนเลือกข้อคำถามที่นักเรียนคิดว่าสามารถหาคำตอบได้จำนวน 1 ข้อ โดยอ่านคำถามข้อนั้นดัง ๆ ซัด ๆ หลังจากนั้นจึงเริ่มดำเนินการในขั้นตอนต่อไป ตามที่ได้ฝึกการรายงานกระบวนการคิดให้กับนักเรียนก่อนการดำเนินการประเมิน และดำเนินการเฉพาะกรณี ดังนี้

3.4.1 กรณีที่นักเรียนตอบคำถามข้อนั้นได้ถูกต้อง และผู้ประเมินพิจารณาตัดสินได้ว่านักเรียนมีสถานะความรู้ในเรื่อง ๆ นั้น ให้ผู้ประเมินสรุปสถานะความรู้ของนักเรียน และ



บันทึกกระบวนการคิดแก้ปัญหาในแบบบันทึกกระบวนการคิดแก้ปัญหา แล้วจึงดำเนินการประเมินต่อไป โดยใช้แบบทดสอบ ฉบับที่ 2 ฉบับที่ 3 และฉบับที่ 4 ตามลำดับ

3.4.2 กรณีที่นักเรียนตอบคำถามข้อนั้นได้ไม่ถูกต้อง และผู้ประเมินพิจารณาตัดสินได้ว่านักเรียนใช้แบบการคิดที่ผิดแบบใด ให้ผู้ประเมินยุติการประเมินและสรุปแบบการคิดที่ผิดที่นักเรียนใช้ รวมทั้งบันทึกกระบวนการคิดที่นักเรียนใช้ในการแก้ปัญหาในแบบบันทึกกระบวนการคิดแก้ปัญหา

3.4.3 กรณีที่นักเรียนตอบคำถามข้อนั้นได้ถูกต้อง แต่ผู้ประเมินไม่สามารถพิจารณาตัดสินสถานะความรู้ของนักเรียนได้ หรือนักเรียนตอบคำถามได้ไม่ถูกต้องและผู้ประเมินไม่สามารถพิจารณาตัดสินแบบการคิดที่ผิดที่นักเรียนใช้ได้ ให้นักเรียนเลือกคำถามจากแบบทดสอบฉบับนั้นเพื่อคิดหาคำตอบอีก จำนวน 1 ข้อ และดำเนินการตามขั้นตอนที่กำหนดไว้อีกครั้ง จนกว่าจะสามารถพิจารณาตัดสินสถานะความรู้หรือแบบการคิดที่ผิดที่นักเรียนใช้ได้

3.5 เมื่อสามารถพิจารณาตัดสินสถานะความรู้หรือแบบการคิดที่ผิดของนักเรียนได้อย่างชัดเจน ให้ผู้ประเมินบันทึกกระบวนการคิดแก้ปัญหาและสรุปสถานะความรู้ รวมทั้งแบบการคิดที่ผิดที่นักเรียนใช้เหล่านั้น และบันทึกลงในแบบบันทึกกระบวนการคิดแก้ปัญหา

3.6 ดำเนินการประเมินกับนักเรียนคนอื่น ๆ ต่อไป จนกว่านักเรียนจะได้รับการประเมินครบทุกคน

3.7 รวบรวมแบบบันทึกกระบวนการคิดแก้ปัญหา และเทปบันทึกเสียงพูดที่ได้จากการรายงานกระบวนการคิดให้กับผู้วิจัย

## บรรณานุกรม

### ภาษาไทย

- กำจร มณีแก้ว. 2539. ผลของการสอนโดยใช้เทคนิคการคิดออกเสียงที่มีต่อความสามารถในการแก้ปัญหาทางคณิตศาสตร์และเจตคติต่อวิชาคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนสาธิตสังกัดสำนักงานสภาสถาบัน. วิทยานิพนธ์ปริญญาามหาบัณฑิต ภาควิชา มัธยมศึกษา ครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ดวงใจ จงธนากร. 2535. ผลของการสอนอ่านแบบแสดงการคิดเป็นถ้อยคำที่มีต่อผลสัมฤทธิ์ในการอ่านเพื่อการเข้าใจภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 4. วิทยานิพนธ์ปริญญาามหาบัณฑิต ภาควิชามัธยมศึกษา ครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- วีระชัย รัตนศิลป์. 2533. รายงานการวิจัยเรื่อง การทดลองใช้สอนในวิชาคณิตศาสตร์กับนักเรียนชั้นประกาศนียบัตรวิชาชีพปีที่ 1. โรงเรียนอัสสัมชัญพาณิชยการนครราชสีมา.
- สุดารัตน์ มนต์นิมิตร. 2545. การใช้เทคนิคการคิดออกเสียงเป็นเครื่องมือในการวินิจฉัยความสามารถในการแก้ปัญหาโจทย์คณิตศาสตร์เพื่อจัดซ่อมเสริม สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3. วิทยานิพนธ์ปริญญาามหาบัณฑิต ภาควิชาวิจัยการศึกษา ครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

### ภาษาอังกฤษ

- Baauw, E and Markopoulos, P.(n.d.) A Comparison of think-aloud and post-task interview for usability testing with children. [Online]. Available from [www.idemployee.id.tue.nl/p.markopoulos/downloadablePapers/BaauwMarkopoulosIDC2004.pdf](http://www.idemployee.id.tue.nl/p.markopoulos/downloadablePapers/BaauwMarkopoulosIDC2004.pdf) [2008, June 30]
- Crain-Threson, C. ; Lippman, M. Z. and McClendon-Magnuson, D. (1997). Windows on comprehension: reading comprehension processes as revealed by two think-aloud procedures. Journal of educational psychology, 89(4):579-591.
- DPhill, H. V. (2005). Thinking aloud a technique for uncovering multi - sensory learning in professional education. Paper presented at the society for research intobhigher education conference, University of Edinburgh, pp.1-10. [Online]. Available from <http://www.leeds.ac.uk/educol/documents/152807.htm> [2008, June 30]

- Ericsson, K. A. and Simon, H. A. (1980). Verbal report as data. Psychological Review, 87(3): 215-251.
- Guan, Z. et al. (2006) The validity of the stimulated retrospective think-aloud method as measured by eye tracking. [Online]. Available from <http://www.staff.washington.edu/zguan/publication/GuanRamey-CHI06-ValidityTAP.pdf> [2008, June 30]
- Johnston, J. C. ; Bottsford-Miller, N. A. and Thompson, S. D.(2006) Using the think aloud method (cognitive labs) to evaluate test design for students with disabilities and English language learners Technical report 44. [Online]. Available from <http://cehd.umn.edu/NCEO/OnlinePubs/Tech44/default.html> [2008, June 30]
- Lavandenz, M.(2003) Think aloud protocols: teaching reading processes to young bilingual students. [Online]. Available from <http://www.cal.org/resources/Digest/0314lavandez.html> [2008, June 30]
- Seng, G. H. (2007, August). The effects of think-aloud in a collaborative environment to improve comprehension of L2 texts. Reading matrix : an international online journal [On-line serial]. ]. Available from 1533242X, 7(2)
- van Den Haak, M. J.; De Jong M. D. T. And Schellens, P. T. (2003) Retrospective vs. Concurrent think-aloud protocols: testing the usability of an online library catalogue. [Online]. Available from <http://www.students.cs.uu.nl/people/jpwester/WO1/Artikelen/p339.pdf> [2008, June 30]
- van Someren, M. W.; Barnard, Y. M. and Sandberg, J. A. C. (1994). The think aloud method a practical guide to modeling cognitive processes. London: Academic Press.

ภาคผนวก

ภาคผนวก ก (1)

แบบทดสอบ เรื่อง การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน

คำชี้แจง ให้นักเรียนพิจารณาข้อคำถามต่อไปนี้ เลือกข้อคำถามที่นักเรียนสามารถหาคำตอบได้จำนวน 1 ข้อ อ่านข้อคำถามตั้ง ๆ ชัด ๆ คิดวิธีหาคำตอบ และดำเนินการหาคำตอบ พร้อมทั้งพูดสิ่งที่คิดออกมาให้มากที่สุด

1.  $\frac{1}{3} + \frac{1}{3}$

2.  $\frac{1}{11} + \frac{2}{11}$

3.  $\frac{6}{7} - \frac{2}{7}$

4.  $\frac{7}{9} - \frac{2}{9}$

5.  $\frac{2}{9} + \frac{6}{9}$

6.  $\frac{3}{8} - \frac{2}{8}$

7.  $\frac{5}{8} - \frac{2}{8}$

8.  $\frac{3}{8} + \frac{4}{8}$

9.  $\frac{2}{6} - \frac{1}{6}$

10.  $\frac{3}{15} + \frac{7}{15}$

ภาคผนวก ก (2)

แบบทดสอบ เรื่อง การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน

คำชี้แจง ให้นักเรียนพิจารณาข้อคำถามต่อไปนี้ เลือกข้อคำถามที่นักเรียนสามารถหาคำตอบได้จำนวน 1 ข้อ อ่านข้อคำถามตั้ง ๆ ชัด ๆ คิดวิธีหาคำตอบ และดำเนินการหาคำตอบ พร้อมทั้งพูดสิ่งที่คิดออกมาให้มากที่สุด

1.  $\frac{1}{2} + \frac{2}{3}$

2.  $\frac{4}{5} + \frac{5}{9}$

3.  $\frac{1}{2} + \frac{4}{5}$

4.  $\frac{1}{2} + \frac{8}{9}$

5.  $\frac{1}{2} - \frac{1}{3}$

6.  $\frac{2}{3} + \frac{4}{6}$

7.  $\frac{2}{3} - \frac{1}{6}$

8.  $\frac{3}{7} - \frac{1}{3}$

9.  $\frac{6}{7} - \frac{1}{9}$

10.  $\frac{2}{4} - \frac{3}{8}$

ภาคผนวก ก (3)

แบบทดสอบ เรื่อง การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน

**คำชี้แจง** ให้นักเรียนพิจารณาข้อคำถามต่อไปนี้ เลือกข้อคำถามที่นักเรียนสามารถหาคำตอบได้จำนวน 1 ข้อ อ่านข้อคำถามตั้ง ๆ ชัด ๆ คิดวิธีหาคำตอบ และดำเนินการหาคำตอบ พร้อมทั้งพูดสิ่งที่คิดออกมาให้มากที่สุด

1.  $1\frac{1}{2} + 2\frac{1}{2}$

2.  $4\frac{3}{5} - 3\frac{1}{5}$

3.  $2\frac{3}{10} + 2\frac{5}{10}$

4.  $6\frac{3}{4} + 4\frac{1}{4}$

5.  $8\frac{5}{8} - 3\frac{1}{8}$

6.  $3\frac{2}{9} - 1\frac{1}{9}$

7.  $9\frac{3}{5} - 7\frac{2}{5}$

8.  $8\frac{7}{10} - 3\frac{1}{10}$

9.  $3\frac{7}{9} + 2\frac{8}{9}$

10.  $2\frac{3}{8} - 1\frac{1}{8}$

ภาคผนวก ก (4)

แบบทดสอบ เรื่อง การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน

คำชี้แจง ให้นักเรียนพิจารณาข้อคำถามต่อไปนี้ เลือกข้อคำถามที่นักเรียนสามารถหาคำตอบได้จำนวน 1 ข้อ อ่านข้อคำถามตั้ง ๆ ชัด ๆ คิดวิธีหาคำตอบ และดำเนินการหาคำตอบ พร้อมทั้งพูดสิ่งที่คิดออกมาให้มากที่สุด

1.  $1\frac{1}{2} + 2\frac{2}{3}$

2.  $3\frac{2}{3} - 1\frac{1}{2}$

3.  $2\frac{4}{7} - 1\frac{1}{3}$

4.  $6\frac{1}{2} + 7\frac{1}{5}$

5.  $2\frac{4}{5} + 4\frac{2}{3}$

6.  $8\frac{3}{4} - 3\frac{1}{2}$

7.  $3\frac{1}{2} + 2\frac{3}{5}$

8.  $4\frac{2}{3} - 1\frac{1}{6}$

9.  $3\frac{2}{3} - 1\frac{3}{4}$

10.  $3\frac{1}{3} + 2\frac{7}{8}$



ภาคผนวก ข  
 แบบการคิดที่ผิดที่ใช้ในการบวกเศษส่วน  
 ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 -6)

แบบการคิดที่ผิดสำหรับคุณลักษณะที่ 1

คุณลักษณะที่	แบบการคิดที่ผิด	ตัวอย่างการคำนวณ
1.การบวก เศษส่วน	1.ไม่เข้าใจวิธีการบวกเศษส่วนแท้ที่มีตัว ส่วนเท่ากัน	$\frac{1}{5} + \frac{2}{5} = \frac{3}{5}$ (คำตอบถูก)
แท้ที่มี ตัวส่วน เท่ากัน	1.1 เศษบวกเศษ ตัวส่วนบวกตัวส่วน	$\frac{1}{5} + \frac{2}{5} = \frac{1+2}{5+5} = \frac{3}{10}$
	1.2 เศษส่วนตัวตั้งบวกกลับกันเป็นเศษ เศษส่วน ตัวบวกกลับ บวกกลับกันเป็นตัวส่วน	$\frac{1}{5} + \frac{2}{5} = \frac{1+5}{2+5} = \frac{6}{7}$
	2. บวกกลับจำนวนนับผิด	
	2.1 บวกกลับเลขผิด	
	2.2 ดูเครื่องหมายสลับบวกเป็นลบ ลบเป็นบวก	

## แบบการคิดที่ผิดสำหรับคุณลักษณะที่ 2

คุณลักษณะที่	แบบการคิดที่ผิด	ตัวอย่างการคำนวณ
2.การบวก ลบเศษส่วน แท้ที่มีตัว ส่วนไม่ เท่ากัน	<b>1.ไม่เข้าใจวิธีการทำตัวส่วนให้เท่ากัน</b>	$\frac{1}{3} + \frac{2}{5} = \frac{(1 \times 5) + (2 \times 3)}{3 \times 5}$ $= \frac{5 + 6}{15} = \frac{11}{15}$ <p style="text-align: center; color: red;">(คำตอบถูก)</p>
	1.1 เอาตัวส่วนที่มากกว่าเป็นตัวส่วน เศษบวกลบเศษ	$\frac{1}{3} + \frac{2}{5} = \frac{1+2}{5} = \frac{3}{5}$
	1.2 เอาตัวส่วนที่น้อยกว่าเป็นตัวส่วน เศษบวกลบเศษ	$\frac{1}{3} + \frac{2}{5} = \frac{1+2}{3} = \frac{3}{3} = 1$
	1.3 เอาตัวส่วนคูณกันเป็นตัวส่วน เศษบวกลบเศษ หรือ หา ค.ร.น. เป็นตัวส่วน เศษบวกลบเศษ	$\frac{1}{3} + \frac{2}{5} = \frac{1+2}{15} = \frac{3}{15} = \frac{1}{5}$
	1.4 หา ค.ร.น. เป็นตัวส่วน เอา ค.ร.น. คูณเศษ เศษบวกลบเศษ	$\frac{1}{3} + \frac{2}{5} = \frac{(1 \times 15) + (2 \times 15)}{15}$ $= \frac{15 + 30}{15} = \frac{45}{15} = 3$
	<b>2.ไม่เข้าใจวิธีการบวกลบเศษส่วนแท้ที่มีตัว ส่วนไม่เท่ากัน</b>	$\frac{1}{3} + \frac{2}{5} = \frac{1+2}{3+5} = \frac{3}{8}$
	2.1 เศษบวกลบเศษ ตัวส่วนบวกลบตัวส่วน	
	2.2 เศษส่วนตัวตั้งบวกลบกันเป็นเศษ เศษส่วน ตัวบวกลบ บวกลบกันเป็นตัวส่วน	$\frac{1}{3} + \frac{2}{5} = \frac{1+3}{2+5} = \frac{4}{7}$
<b>3. บวกลบจำนวนนับผิด</b>		
3.1 บวกลบเลขผิด		
3.2 ดูเครื่องหมายสลับบวกเป็นลบ ลบเป็นบวก		

แบบการคิดที่ผิดสำหรับคุณลักษณะที่ 3 (1)

คุณลักษณะที่	แบบการคิดที่ผิด	ตัวอย่างการคำนวณ
3.การบวก ลบจำนวน คละที่มีตัว ส่วนเท่ากัน	<b>1. ไม่เข้าใจวิธีทำจำนวนคละให้เป็น เศษเกิน</b>	$1\frac{1}{5} + 2\frac{3}{5} = (1+2)\frac{1+3}{5} = 3\frac{4}{5}$ <p>หรือ</p> $1\frac{1}{5} + 2\frac{3}{5} = \frac{6}{5} + \frac{13}{5} = \frac{6+13}{5}$ $= \frac{19}{5} = 3\frac{4}{5}$ <p style="text-align: center;"><b>(คำตอบถูก)</b></p>
	1.1 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เศษบวกลบเศษ	$1\frac{1}{5} + 2\frac{3}{5} = \frac{11}{5} + \frac{23}{5} = \frac{11+23}{5}$ $= \frac{34}{5} = 6\frac{4}{5}$
	1.2 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เศษบวกลบเศษ	$1\frac{1}{5} + 2\frac{3}{5} = \frac{7}{5} + \frac{10}{5} = \frac{7+10}{5}$ $= \frac{17}{5} = 3\frac{2}{5}$
	<b>2. ไม่เข้าใจวิธีทำจำนวนคละให้เป็น เศษเกิน และคำตอบไม่สมบูรณ์</b>	
	2.1 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เศษบวกลบเศษ และไม่ทำเศษเกินให้เป็น จำนวนคละ หรือ ไม่ทำให้เป็นเศษส่วนอย่างต่ำ	$1\frac{1}{5} + 2\frac{3}{5} = \frac{11}{5} + \frac{23}{5} = \frac{11+23}{5}$ $= \frac{34}{5}$
2.2 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เศษบวกลบเศษ และไม่ทำเศษเกินให้เป็น จำนวนคละ หรือ ไม่ทำให้เป็นเศษส่วนอย่างต่ำ	$1\frac{1}{5} + 2\frac{3}{5} = \frac{7}{5} + \frac{10}{5} = \frac{7+10}{5}$ $= \frac{17}{5}$	

แบบการคิดที่ผิดสำหรับคุณลักษณะที่ 3 (2)

คุณลักษณะที่	แบบการคิดที่ผิด	ตัวอย่างการคำนวณ
	<p><b>3. ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน และไม่เข้าใจวิธีการบวกเศษส่วนที่มีส่วนเท่ากัน</b></p> <p>3.1 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เศษบวกเศษ ตัวส่วนบวกเศษตัวส่วน</p>	$1\frac{1}{5} + 2\frac{3}{5} = \frac{11}{5} + \frac{23}{5} = \frac{11+23}{5+5}$ $= \frac{34}{10} = 3\frac{4}{10}$
	<p>3.2 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เศษส่วนตัวตั้งบวกกลับกันเป็นเศษ เศษส่วนตัวบวกกลับ บวกกลับกันเป็นตัวส่วน</p>	$1\frac{1}{5} + 2\frac{3}{5} = \frac{11}{5} + \frac{23}{5} = \frac{11+5}{23+5}$ $= \frac{16}{28} = \frac{8}{14} = \frac{4}{7}$
	<p>3.3 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เศษบวกเศษ ตัวส่วนบวกกลับตัวส่วน</p>	$1\frac{1}{5} + 2\frac{3}{5} = \frac{7}{5} + \frac{10}{5} = \frac{7+10}{5+5}$ $= \frac{17}{10} = 1\frac{7}{10}$
	<p>3.4 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เศษส่วนตัวตั้งบวกกลับกันเป็นเศษ เศษส่วนตัวบวกกลับ บวกกลับกันเป็นตัวส่วน</p>	$1\frac{1}{5} + 2\frac{3}{5} = \frac{7}{5} + \frac{10}{5} = \frac{7+5}{10+5}$ $= \frac{12}{15}$
	<p><b>4. ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน ไม่เข้าใจวิธีการบวกเศษส่วนที่มีส่วนเท่ากัน และคำตอบไม่สมบูรณ์</b></p> <p>4.1 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เศษบวกเศษ ตัวส่วนบวกกลับตัวส่วน และ ไม่ทำเศษเกินให้เป็นจำนวนคละ หรือ ไม่ทำให้เป็นเศษส่วนอย่างต่ำ</p>	$1\frac{1}{5} + 2\frac{3}{5} = \frac{11}{5} + \frac{23}{5} = \frac{11+23}{5+5}$ $= \frac{34}{10}$
	<p>4.2 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เศษส่วนตัวตั้งบวกกลับกันเป็นเศษ เศษส่วนตัวบวกกลับ บวกกลับกันเป็นตัวส่วน และไม่ทำ เศษเกินให้เป็นจำนวนคละ หรือ ไม่ทำให้เป็นเศษส่วนอย่างต่ำ</p>	$1\frac{1}{5} + 2\frac{3}{5} = \frac{11}{5} + \frac{23}{5} = \frac{11+5}{23+5}$ $= \frac{16}{28}$

แบบการคิดที่ผิดสำหรับคุณลักษณะที่ 3 (3)

คุณลักษณะที่	แบบการคิดที่ผิด	ตัวอย่างการคำนวณ
	4.3 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เศษบวกเศษ ตัวส่วนบวกตัวส่วน และ ไม่ทำเศษเกินให้เป็นจำนวนคละ หรือ ไม่ทำให้ เป็นเศษส่วนอย่างต่ำ	$1\frac{1}{5} + 2\frac{3}{5} = \frac{7}{5} + \frac{10}{5} = \frac{7+10}{5+5}$ $= \frac{17}{10}$
	4.4 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เศษส่วนตัวตั้งบวกกลับกันเป็นเศษ เศษส่วนตัว บวกกลับ บวกกลับกันเป็นตัวส่วน และไม่ทำ เศษเกินให้เป็นจำนวนคละ หรือ ไม่ทำให้เป็น เศษส่วนอย่างต่ำ	$1\frac{1}{5} + 2\frac{3}{5} = \frac{7}{5} + \frac{10}{5} = \frac{7+5}{10+5}$ $= \frac{12}{15}$
	<p><b>5. บวกกลับจำนวนนับผิด</b></p> <p>5.1 บวกกลับเลขผิด</p> <p>5.2 ดูเครื่องหมายสลับบวกเป็นลบ ลบเป็นบวก</p>	

แบบการคิดที่ผิดสำหรับคุณลักษณะที่ 4 (1)

คุณลักษณะที่	แบบการคิดที่ผิด	ตัวอย่างการคำนวณ
4.การบวก ลบจำนวน คละที่มีตัว ส่วนไม่ เท่ากัน	1. ไม่เข้าใจวิธีทำจำนวนคละให้เป็น เศษเกิน และไม่เข้าใจวิธีการบวกลบ เศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน	$1\frac{1}{3} + 2\frac{2}{5} = (1+2)\frac{1}{3} + \frac{2}{5}$ $= 3\frac{(5+6)}{5} = 3\frac{11}{5}$ <p>หรือ</p> $1\frac{1}{3} + 2\frac{2}{5} = \frac{4}{3} + \frac{12}{5}$ $= \frac{(4 \times 5)}{3} + \frac{(12 \times 3)}{5}$ $= \frac{20+36}{15} = \frac{56}{15} = 3\frac{11}{15}$ <p>(คำตอบถูก)</p>
	1.1 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เศษบวกลบเศษ ตัวส่วนบวกลบตัวส่วน	$1\frac{1}{3} + 2\frac{2}{5} = \frac{11}{3} + \frac{22}{5}$ $= \frac{11+22}{3+5} = \frac{33}{8} = 4\frac{1}{8}$
	1.2 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เศษส่วนตัวตั้งบวกลบกันเป็นเศษ เศษส่วนตัว บวกลบ บวกลบกันเป็นตัวส่วน	$1\frac{1}{3} + 2\frac{2}{5} = \frac{11}{3} + \frac{22}{5}$ $= \frac{11+3}{22+5} = \frac{14}{27}$
	1.3 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เศษบวกลบเศษ ตัวส่วนบวกลบตัวส่วน	$1\frac{1}{3} + 2\frac{2}{5} = \frac{5}{3} + \frac{9}{5}$ $= \frac{5+9}{3+5} = \frac{14}{8} = 1\frac{6}{8} = 1\frac{3}{4}$
	1.4 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เศษส่วนตัวตั้งบวกลบกันเป็นเศษ เศษส่วนตัว บวกลบ บวกลบกันเป็นตัวส่วน	$1\frac{1}{3} + 2\frac{2}{5} = \frac{5}{3} + \frac{9}{5} = \frac{5+3}{9+5}$ $= \frac{8}{14} = \frac{4}{7}$

แบบการคิดที่ผิดสำหรับคุณลักษณะที่ 4 (2)

คุณลักษณะที่	แบบการคิดที่ผิด	ตัวอย่างการคำนวณ
	<p><b>2. ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน และไม่เข้าใจวิธีการทำตัวส่วนให้เท่ากัน</b></p> <p>2.1 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เอาตัวส่วนที่มากกว่าเป็นตัวส่วน เศษบวกลบเศษ</p>	$1\frac{1}{3} + 2\frac{2}{5} = \frac{11}{3} + \frac{22}{5}$ $= \frac{11+22}{5} = \frac{33}{5} = 6\frac{3}{5}$
	<p>2.2 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เอาตัวส่วนที่น้อยกว่าเป็นส่วน เศษบวกลบเศษ</p>	$1\frac{1}{3} + 2\frac{2}{5} = \frac{11}{3} + \frac{22}{5}$ $= \frac{11+22}{3} = \frac{33}{3} = 11$
	<p>2.3 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เอาตัวส่วนคูณกันเป็นตัวส่วน เศษบวกลบเศษ หรือหา ค.ร.น. เป็นตัวส่วน เศษบวกลบเศษ</p>	$1\frac{1}{3} + 2\frac{2}{5} = \frac{11}{3} + \frac{22}{5}$ $= \frac{11+22}{15} = \frac{33}{15} = 2\frac{3}{15} = 2\frac{1}{5}$
	<p>2.4 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย หา ค.ร.น. เป็นตัวส่วน เอา ค.ร.น. คูณเศษ เศษบวกลบเศษ</p>	$1\frac{1}{3} + 2\frac{2}{5} = \frac{11}{3} + \frac{22}{5}$ $= \frac{(11 \times 5) + (22 \times 3)}{15}$ $= \frac{165 + 330}{15} = \frac{495}{15} = 33$
	<p>2.5 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เอาตัวส่วนที่มากกว่าเป็นตัวส่วน เศษบวกลบเศษ</p>	$1\frac{1}{3} + 2\frac{2}{5} = \frac{5}{3} + \frac{9}{5}$ $= \frac{5+9}{5} = \frac{14}{5} = 2\frac{4}{5}$
	<p>2.6 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เอาตัวส่วนที่น้อยกว่าเป็นส่วน เศษบวกลบเศษ</p>	$1\frac{1}{3} + 2\frac{2}{5} = \frac{5}{3} + \frac{9}{5}$ $= \frac{5+9}{3} = \frac{14}{3} = 4\frac{2}{3}$
	<p>2.7 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เอาตัวส่วนคูณกันเป็นตัวส่วน เศษบวกลบเศษ หรือหา ค.ร.น. เป็นตัวส่วน เศษบวกลบเศษ</p>	$1\frac{1}{3} + 2\frac{2}{5} = \frac{5}{3} + \frac{9}{5}$ $= \frac{5+9}{15} = \frac{14}{15}$

แบบการคิดที่ผิดสำหรับคุณลักษณะที่ 4 (3)

คุณลักษณะที่	แบบการคิดที่ผิด	ตัวอย่างการคำนวณ
	<p>2.8 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ หา ค.ร.น เป็นตัวส่วน เอา ค.ร.น คูณเศษ เศษ บวกลบเศษ</p>	$1\frac{1}{3} + 2\frac{2}{5} = \frac{5}{3} + \frac{9}{5}$ $= \frac{(5 \times 15) + (9 \times 15)}{15}$ $= \frac{75 + 135}{15} = \frac{210}{15} = 14$
	<p><b>3. ไม่เข้าใจวิธีทำจำนวนคละให้เป็น เศษเกิน ไม่เข้าใจวิธีการทำตัวส่วนให้ เท่ากัน และคำตอบไม่สมบูรณ์</b></p> <p>3.1 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เอาตัวส่วนที่มากกว่าเป็นตัวส่วน เศษบวกลบ เศษ ไม่ทำเศษเกินให้เป็นจำนวนคละ หรือไม่ ทำให้เป็นเศษส่วนอย่างต่ำ</p> <p>3.2 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เอาตัวส่วนที่น้อยกว่าเป็นส่วน เศษบวกลบเศษ ไม่ทำเศษเกินให้เป็นจำนวนคละ หรือไม่ทำให้ เป็นเศษส่วนอย่างต่ำ</p> <p>3.3 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เอา ตัวส่วนคูณกันเป็นตัวส่วน เศษบวกลบเศษ หรือหา ค.ร.น เป็นตัวส่วน เศษบวกลบเศษ ไม่ทำเศษเกินให้ เป็นจำนวนคละ หรือไม่ทำให้เป็นเศษส่วนอย่างต่ำ</p> <p>3.4 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย หา ค.ร.น เป็นตัวส่วน เอา ค.ร.น คูณเศษ เศษ บวกลบเศษ ไม่ทำเศษเกินให้เป็นจำนวนคละ หรือไม่ทำให้เป็นเศษส่วนอย่างต่ำ</p> <p>3.5 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เอาตัวส่วนที่มากกว่าเป็นตัวส่วน เศษบวกลบ เศษ ไม่ทำเศษเกินให้เป็นจำนวนคละ หรือไม่ ทำให้เป็นเศษส่วนอย่างต่ำ</p>	$1\frac{1}{3} + 2\frac{2}{5} = \frac{11}{3} + \frac{22}{5}$ $= \frac{11 + 22}{5} = \frac{33}{5}$ $1\frac{1}{3} + 2\frac{2}{5} = \frac{11}{3} + \frac{22}{5}$ $= \frac{11 + 22}{3} = \frac{33}{3}$ $1\frac{1}{3} + 2\frac{2}{5} = \frac{11}{3} + \frac{22}{5}$ $= \frac{11 + 22}{15} = \frac{33}{15}$ $1\frac{1}{3} + 2\frac{2}{5} = \frac{11}{3} + \frac{22}{5}$ $= \frac{(11 \times 15) + (22 \times 15)}{15}$ $= \frac{165 + 330}{15} = \frac{495}{15}$ $1\frac{1}{3} + 2\frac{2}{5} = \frac{5}{3} + \frac{9}{5}$ $= \frac{5 + 9}{5} = \frac{14}{5}$



แบบการคิดที่ผิดสำหรับคุณลักษณะที่ 4 (4)

คุณลักษณะที่	แบบการคิดที่ผิด	ตัวอย่างการคำนวณ
	3.6 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เอาตัวส่วนที่น้อยกว่าเป็นส่วน เศษบวกลบเศษ ไม่ทำเศษเกินให้เป็นจำนวนคละ หรือไม่ทำให้เป็นเศษส่วนอย่างต่ำ	$1\frac{1}{3} + 2\frac{2}{5} = \frac{5}{3} + \frac{9}{5}$ $= \frac{5+9}{3} = \frac{14}{3}$
	3.7 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เอาตัวส่วนคูณกันเป็นตัวส่วน เศษบวกลบเศษ หรือหา ค.ร.น เป็นตัวส่วน เศษบวกลบเศษ ไม่ทำเศษเกินให้เป็นจำนวนคละ หรือไม่ทำให้เป็นเศษส่วนอย่างต่ำ	$1\frac{1}{3} + 2\frac{2}{5} = \frac{5}{3} + \frac{9}{5}$ $= \frac{5+9}{15} = \frac{14}{15}$
	3.8 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ หา ค.ร.น เป็นตัวส่วน เอา ค.ร.น คูณเศษ เศษบวกลบเศษ ไม่ทำเศษเกินให้เป็นจำนวนคละ หรือไม่ทำให้เป็นเศษส่วนอย่างต่ำ	$1\frac{1}{3} + 2\frac{2}{5} = \frac{5}{3} + \frac{9}{5}$ $= \frac{(5 \times 5) + (9 \times 3)}{15}$ $= \frac{25 + 27}{15} = \frac{52}{15}$
	<p><b>4. ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน ไม่เข้าใจวิธีการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน และคำตอบไม่สมบูรณ์</b></p> <p>4.1 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เศษบวกลบเศษ ตัวส่วนบวกลบตัวส่วน ไม่ทำเศษเกินให้เป็นจำนวนคละ หรือไม่ทำให้เป็นเศษส่วนอย่างต่ำ</p>	$1\frac{1}{3} + 2\frac{2}{5} = \frac{11}{3} + \frac{22}{5}$ $= \frac{11+22}{3+5} = \frac{33}{8}$
	4.2 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เศษส่วนตัวตั้งบวกลบกันเป็นเศษ เศษส่วนตัวบวกลบ บวกลบกันเป็นตัวส่วน ไม่ทำเศษเกินให้เป็นจำนวนคละ หรือไม่ทำให้เป็นเศษส่วนอย่างต่ำ	$1\frac{1}{3} + 2\frac{2}{5} = \frac{11}{3} + \frac{22}{5}$ $= \frac{11+3}{22+5} = \frac{14}{27}$
	4.3 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เศษบวกลบเศษ ตัวส่วนบวกลบตัวส่วน ไม่ทำเศษเกินให้เป็นจำนวนคละ หรือไม่ทำให้เป็นเศษส่วนอย่างต่ำ	$1\frac{1}{3} + 2\frac{2}{5} = \frac{5}{3} + \frac{9}{5}$ $= \frac{5+9}{3+5} = \frac{14}{8}$

แบบการคิดที่ผิดสำหรับคุณลักษณะที่ 4 (5)

คุณลักษณะที่	แบบการคิดที่ผิด	ตัวอย่างการคำนวณ
	4.4 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เศษส่วนตัวตั้งบวกกลับกันเป็นเศษ เศษส่วนตัวบวก ลบ บวกกลับกันเป็นตัวส่วน ไม่ทำเศษเกินให้เป็น จำนวนคละ หรือไม่ทำให้เป็นเศษส่วนอย่างต่ำ	$1\frac{1}{3} + 2\frac{2}{5} = \frac{5}{3} + \frac{9}{5}$ $= \frac{5+3}{9+5} = \frac{8}{14} = \frac{4}{7}$
	<b>5. บวกลบจำนวนนับผิด</b> 5.1 บวกลบเลขผิด 5.2 ดูเครื่องหมายสลับบวกเป็นลบ ลบเป็นบวก	

ภาคผนวก ค.

แบบบันทึกกระบวนการคิดแก้ปัญหา  
ในการประเมินเชิงวิจจัยสถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน  
ระดับการศึกษาขั้นพื้นฐาน (ชั้นประถมศึกษาปีที่ 1 – 6)

ตอนที่ 1 ข้อมูลทั่วไปของผู้บันทึก

คำชี้แจง โปรดทำเครื่องหมาย ✓ ลงใน  หน้าข้อความหรือเติมข้อความ  
ที่เป็นข้อมูลเกี่ยวกับตัวท่านตามสภาพที่เป็นจริง

1. เพศ  หญิง  ชาย
2. อายุ  ต่ำกว่า 25 ปี  25 – 30 ปี  
 31 – 35 ปี  36 – 40 ปี  
 41 – 45 ปี  46 ปีขึ้นไป
3. การศึกษาสูงสุด  ต่ำกว่าปริญญาตรี สาขาวิชา/วิชาเอก.....  
 ปริญญาตรี สาขาวิชา/วิชาเอก.....  
 ปริญญาโท สาขาวิชา/วิชาเอก.....  
 ปริญญาเอก สาขาวิชา/วิชาเอก.....
4. สาขาวิชา/วิชาเอกที่สำเร็จการศึกษา.....
5. ประสบการณ์ในการสอนวิชาคณิตศาสตร์ ระดับประถมศึกษา จำนวน.....ปี

ตอนที่ 2 ข้อมูลทั่วไปของการประเมินและผู้ได้รับการประเมิน

คำชี้แจง โปรดทำเครื่องหมาย ✓ ลงใน  หน้าข้อความหรือเติมข้อความ  
ที่เป็นข้อมูลเกี่ยวกับการประเมินและผู้ได้รับการประเมินตามสภาพที่เป็นจริง

1. เพศของนักเรียน  หญิง  ชาย
2. อายุของนักเรียน.....ปี
3. ระดับชั้นที่นักเรียนกำลังศึกษา.....
4. วันที่ประเมิน วันที่.....เดือน.....พ.ศ. ....
5. เริ่มประเมิน เวลา.....น. การประเมินแล้วเสร็จ เวลา.....น.









## ภาคผนวก จ

คู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์  
ในการประเมินเชิงวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วน  
ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6)



# คู่มือการใช้โปรแกรม

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์  
ในการประเมินเชิงวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วน  
ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6)



จัดทำโดย

นางอมรรัตน์ สร้อยสังวาลย์

คู่มือนี้เป็นส่วนหนึ่งของการศึกษาระดับปริญญาโท  
สาขาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา  
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย พ.ศ.2551

## คำนำ

คู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ในการประเมินเชิงวินิจฉัย สถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) จัดทำขึ้นเพื่อสร้างความรู้ความเข้าใจเกี่ยวกับโปรแกรมและการใช้โปรแกรม ซึ่งเป็นส่วนหนึ่งของการดำเนินการจัดทำวิทยานิพนธ์ระดับปริญญาเอก สาขาการวัดและประเมินผล การศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เรื่อง การพัฒนาวิธีการประเมินเชิงวินิจฉัยโดยประยุกต์ใช้โมเดลลำดับขั้นของคุณลักษณะและการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

เนื้อหาที่นำเสนอในคู่มือจำแนกเป็น 3 ส่วน ส่วนที่ 1 บทนำ จะกล่าวถึงข้อมูลทั่วไปเกี่ยวกับโปรแกรม ระบบคอมพิวเตอร์ที่เหมาะสม การติดตั้งโปรแกรมและการเรียกใช้โปรแกรม ส่วนที่ 2 โปรแกรมย่อยการพิมพ์ข้อสอบ และส่วนที่ 3 โปรแกรมย่อยการทดสอบ สำหรับเนื้อหาในส่วนที่ 2 และส่วนที่ 3 จะอธิบายถึงองค์ประกอบที่สำคัญของหน้าจอและขั้นตอนการใช้โปรแกรม

ผู้จัดทำหวังเป็นอย่างยิ่งว่าคู่มือฉบับนี้จะช่วยสร้างความรู้ความเข้าใจเกี่ยวกับโปรแกรมและการใช้โปรแกรมให้กับผู้ที่เกี่ยวข้องและช่วยให้การใช้โปรแกรมดำเนินการไปอย่างมีประสิทธิภาพและเกิดประโยชน์สูงสุด

นางอมรรัตน์ สร้อยสังวาลย์

ผู้จัดทำ

กรกฎาคม 2551

## สารบัญ

	หน้า
บทนำ.....	1
ข้อมูลทั่วไปเกี่ยวกับโปรแกรม.....	3
ระบบคอมพิวเตอร์ที่เหมาะสม.....	4
การติดตั้งโปรแกรม.....	4
การเรียกใช้โปรแกรม.....	4
โปรแกรมย่อยการจัดพิมพ์ข้อสอบ.....	5
องค์ประกอบที่สำคัญของหน้าจอ.....	5
ขั้นตอนการใช้โปรแกรม.....	9
โปรแกรมย่อยการทดสอบ.....	21
องค์ประกอบที่สำคัญของหน้าจอ.....	21
ขั้นตอนการใช้โปรแกรม.....	23

## บทนำ

### ข้อมูลทั่วไปเกี่ยวกับโปรแกรม

โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ (computer adaptive testing program) พัฒนาขึ้นโดยมีวัตถุประสงค์เพื่อเป็นเครื่องมือในการประเมินเชิงวินิจฉัยสถานะความรู้และแบบการคิดที่ผิดในการบวกลบเศษส่วน ระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4 – 6) ใช้แนวคิดจากโมเดลลำดับขั้นของคุณลักษณะ (The Attribute Hierarchy Model) ในการพัฒนาค้างข้อสอบและการกำหนดลำดับขั้นในการเลือกข้อสอบจากคลังข้อสอบ ลักษณะของโปรแกรมและลำดับขั้นการเลือกข้อสอบจากคลังข้อสอบมีรายละเอียด ดังนี้

#### 1. ลักษณะของโปรแกรม

โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะประกอบด้วยโปรแกรมการพิมพ์ข้อสอบและโปรแกรมการทดสอบ ซึ่งจำแนกเป็นโปรแกรมย่อยจำนวน 6 โปรแกรม แต่ละโปรแกรมใช้วิธีการคัดเลือกข้อสอบข้อแรก วิธีการคัดเลือกข้อสอบจากคลัง วิธีการประมาณค่าความสามารถของผู้สอบและการรายงานผลการทดสอบเหมือนกัน แต่ต่างกันที่จำนวนข้อสอบที่กำหนดให้ทำในแต่ละลำดับขั้น เกณฑ์การเลื่อนลำดับขั้นของคุณลักษณะเพื่อเลือกข้อสอบข้อต่อไปและเกณฑ์การหยุดสอบ ดังนี้

ตารางที่ 1 รายละเอียดโปรแกรมย่อยที่ใช้ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

โปรแกรมย่อยที่	จำนวนข้อสอบที่กำหนดให้ทำในแต่ละชั้น อย่างน้อยจำนวน	เกณฑ์การเลื่อนลำดับชั้นของคุณลักษณะเพื่อเลือกข้อสอบข้อต่อไป	เกณฑ์การหยุดสอบ
1	3	ตอบข้อสอบถูกจำนวน 2 ใน 3 ข้อ หรือ ร้อยละ 67	ใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 2 ใน 3 ข้อ หรือ ร้อยละ 67
2	4	ตอบข้อสอบถูกจำนวน 3 ใน 4 ข้อ หรือ ร้อยละ 75	ใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 3 ใน 4 ข้อ หรือ ร้อยละ 75
3	5	ตอบข้อสอบถูกจำนวน 4 ใน 5 ข้อ หรือ ร้อยละ 80	ใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 4 ใน 5 ข้อ หรือ ร้อยละ 80
4	3	ตอบข้อสอบถูกจำนวน 2 ใน 3 ข้อ หรือ ร้อยละ 67 และมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3	ใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 2 ใน 3 ข้อ หรือ ร้อยละ 67 และมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3
5	4	ตอบข้อสอบถูกจำนวน 3 ใน 4 ข้อ หรือ ร้อยละ 75 และมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3	ใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 3 ใน 4 ข้อ หรือ ร้อยละ 75 และมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3
6	5	ตอบข้อสอบถูกจำนวน 4 ใน 5 ข้อ หรือ ร้อยละ 80 และมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3	ใช้แบบการคิดที่ผิดแบบเดียวกัน จำนวน 4 ใน 5 ข้อ หรือ ร้อยละ 80 และมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำกว่า 0.3

## 2. ลำดับขั้นการเลือกข้อสอบจากคลังข้อสอบ

ข้อสอบในคลังข้อสอบมีจำนวน 438 ข้อ จำแนกเป็นข้อสอบในแต่ละชุดของข้อสอบ ดังนี้

คุณลักษณะที่ 1 การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน จำนวน 105 ข้อ

คุณลักษณะที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน จำนวน 115 ข้อ

คุณลักษณะที่ 3 การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน จำนวน 106 ข้อ

คุณลักษณะที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน จำนวน 112 ข้อ

ลำดับขั้นการเลือกข้อสอบจากคลังข้อสอบกำหนดจากลำดับขั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วนระดับการศึกษาขั้นพื้นฐาน (ชั้นประถมศึกษาปีที่ 4 – 6) ซึ่งจำแนกลักษณะการเลือกข้อสอบจากคลังข้อสอบในแต่ละลำดับขั้นได้ดังนี้

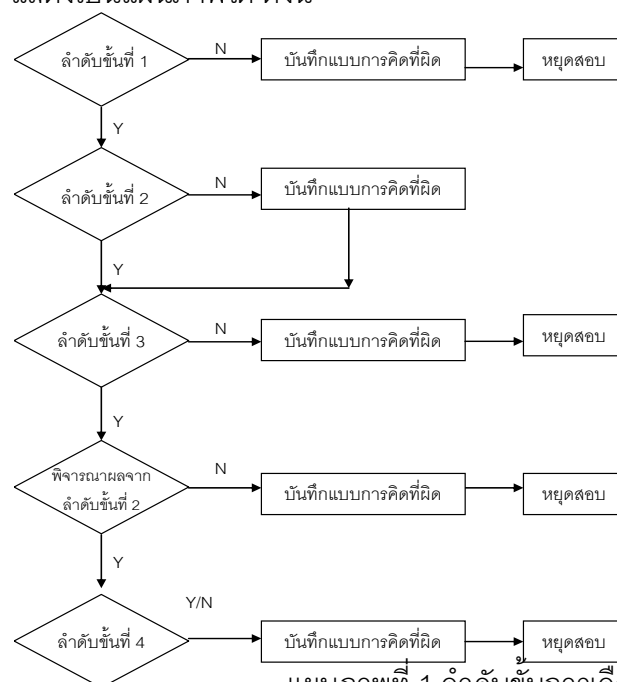
กรณีที่ 1 เลือกข้อสอบจากคุณลักษณะที่ 1 ผ่านคุณลักษณะที่ 1 เลือกข้อสอบจากคุณลักษณะที่ 2 ผ่านคุณลักษณะที่ 2 เลือกข้อสอบจากคุณลักษณะที่ 3 ผ่านคุณลักษณะที่ 3 เลือกข้อสอบจากคุณลักษณะที่ 4 หยุดสอบ

กรณีที่ 2 เลือกข้อสอบจากคุณลักษณะที่ 1 ผ่านคุณลักษณะที่ 1 เลือกข้อสอบจากคุณลักษณะที่ 2 ไม่ผ่านคุณลักษณะที่ 2 เลือกข้อสอบจากคุณลักษณะที่ 3 ผ่านคุณลักษณะที่ 3 หยุดสอบ

กรณีที่ 3 เลือกข้อสอบจากคุณลักษณะที่ 1 ผ่านคุณลักษณะที่ 1 เลือกข้อสอบจากคุณลักษณะที่ 2 ผ่านคุณลักษณะที่ 2 เลือกข้อสอบจากคุณลักษณะที่ 3 ไม่ผ่านคุณลักษณะที่ 3 หยุดสอบ

กรณีที่ 4 เลือกข้อสอบจากคุณลักษณะที่ 1 ไม่ผ่านคุณลักษณะที่ 1 หยุดสอบ

แสดงเป็นแผนภาพได้ ดังนี้



แผนภาพที่ 1 ลำดับขั้นการเลือกข้อสอบจากคลังข้อสอบ

## ระบบคอมพิวเตอร์ที่เหมาะสม

โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะนี้พัฒนาขึ้นเพื่อให้สามารถใช้ได้กับเครื่องไมโครคอมพิวเตอร์ โดยกำหนดลักษณะพื้นฐานของเครื่องไมโครคอมพิวเตอร์ ดังนี้

- Pentium III หรือสูงกว่า
- มีหน่วยความจำ (RAM) อย่างน้อย 512 เมกะไบต์
- ใช้กับระบบปฏิบัติการ Windows XP หรือสูงกว่า
- Hard Drive ที่ว่างสำหรับติดตั้งโปรแกรมประมาณ 15 MB และมีที่ว่างสำหรับการทำงาน 10% ของความจุทั้งหมด

## การติดตั้งโปรแกรม

การติดตั้งโปรแกรมดำเนินการโดยสร้างโฟลเดอร์และคัดลอกไฟล์ทั้งหมดลงไปในโฟลเดอร์นั้น

## การเรียกใช้โปรแกรม

การเรียกใช้โปรแกรมดำเนินการโดยการเปิดโฟลเดอร์และไฟล์การพิมพ์ข้อสอบหรือการทดสอบที่ต้องการ โดยดับเบิลคลิกที่ไฟล์นั้น ๆ

## โปรแกรมการจัดพิมพ์ข้อสอบ

โปรแกรมการจัดพิมพ์ข้อสอบเป็นโปรแกรมที่ใช้ในการเพิ่ม ลด หรือปรับปรุงแก้ไขข้อสอบ  
จำแนกการนำเสนอเป็น 2 ส่วน ส่วนแรกเป็นองค์ประกอบที่สำคัญของหน้าจอและคำสั่งในการ  
จัดพิมพ์ข้อสอบ ส่วนที่ 2 เป็นขั้นตอนการใช้โปรแกรมการจัดพิมพ์ข้อสอบ มีรายละเอียด ดังนี้

### องค์ประกอบที่สำคัญของหน้าจอและคำสั่งในการจัดพิมพ์ข้อสอบ

#### 1. องค์ประกอบที่สำคัญของหน้าจอการจัดพิมพ์ข้อสอบ

องค์ประกอบที่สำคัญของหน้าจอ ดังแสดงในภาพประกอบที่ 1 มีดังนี้

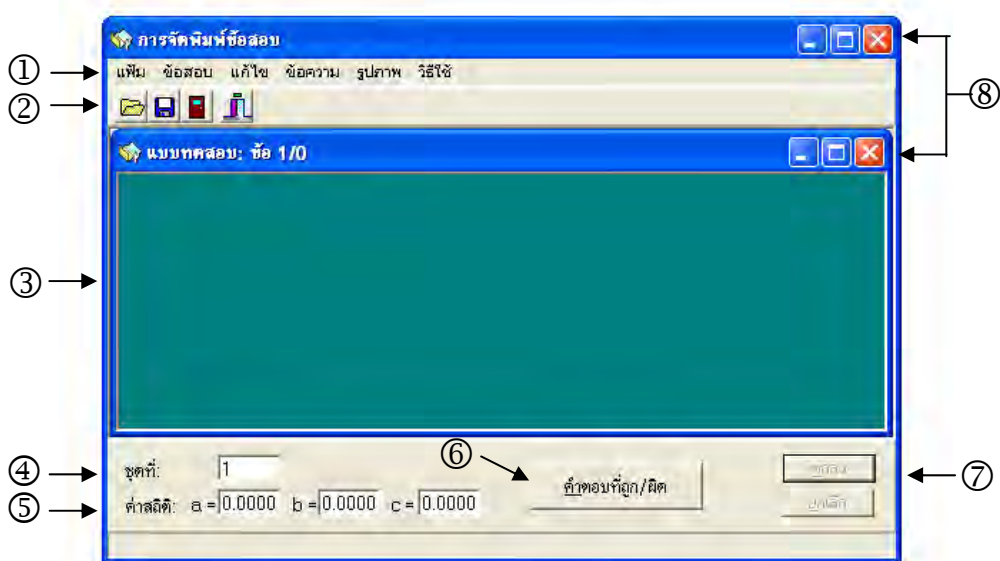
- ① กลุ่มคำสั่ง ประกอบด้วย เพิ่ม ข้อสอบ แก้ไข ข้อความ รูปภาพ วิธีใช้
- ② กลุ่มไอคอน ซึ่งเป็นภาพแทนคำสั่ง ประกอบด้วย รูปภาพแทนคำสั่ง เปิด

บันทึก ปิด เลิกการทำงาน

- ③ พื้นที่ว่างเพื่อใช้ในการพิมพ์ข้อสอบ
- ④ พื้นที่ระบุชุดของข้อสอบ
- ⑤ พื้นที่ระบุค่าสถิติของข้อสอบ ประกอบด้วย ค่าอำนาจจำแนก (a) ค่าความยาก (b)

และค่าการเดา (c)

- ⑥ ปุ่มเชื่อมโยงเพื่อระบุคำตอบถูก/ผิด
- ⑦ ปุ่มคำสั่ง ประกอบด้วย ตกลงและยกเลิก
- ⑧ ปุ่มคำสั่ง ได้แก่ ย่อจอภาพ ขยายจอภาพและปิดจอภาพ

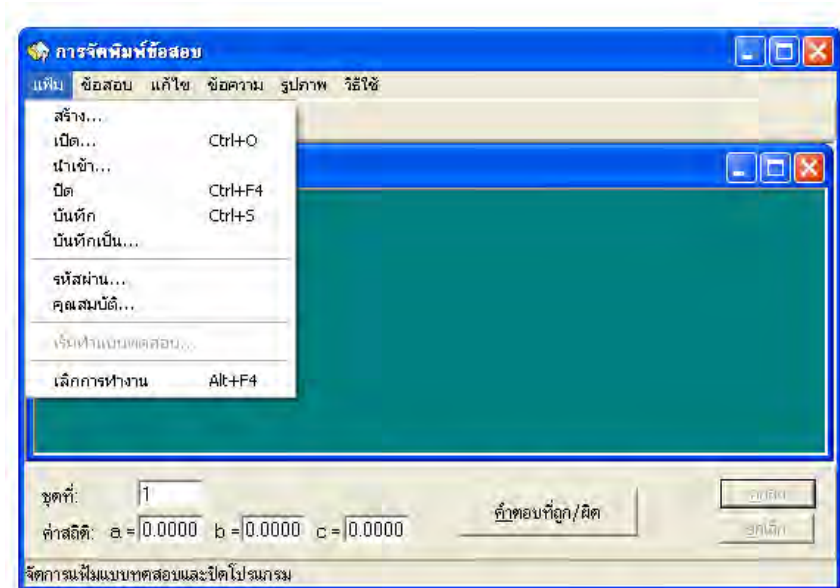


ภาพประกอบที่ 1 หน้าจอการจัดพิมพ์ข้อสอบ



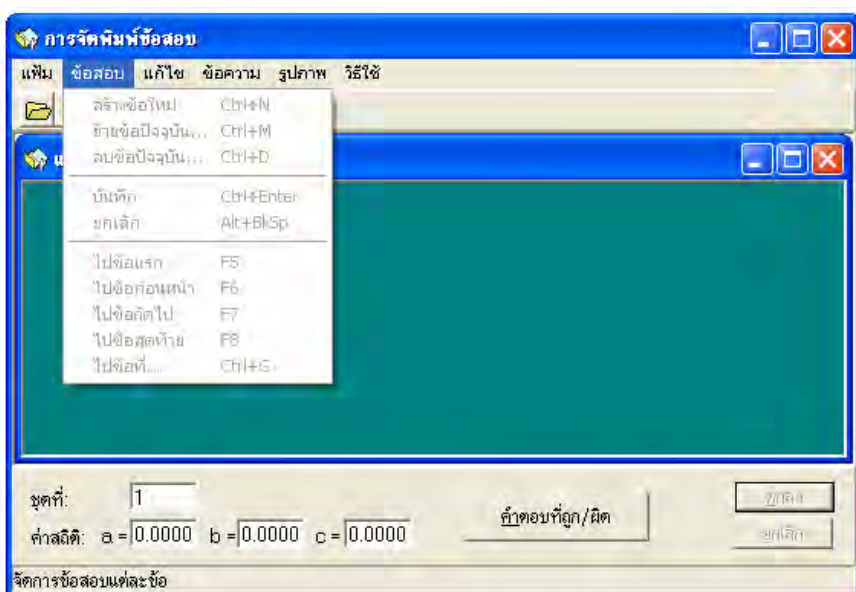
## 2. คำสั่งในการจัดพิมพ์ข้อสอบ

2.1 คำสั่งแฟ้ม ประกอบด้วย สร้าง เปิด นำเข้า ปิด บันทึก บันทึกเป็น รหัสผ่าน คุณสมบัติ เริ่มทำแบบทดสอบและเลิกการทำงาน ดังแสดงในภาพประกอบที่ 2



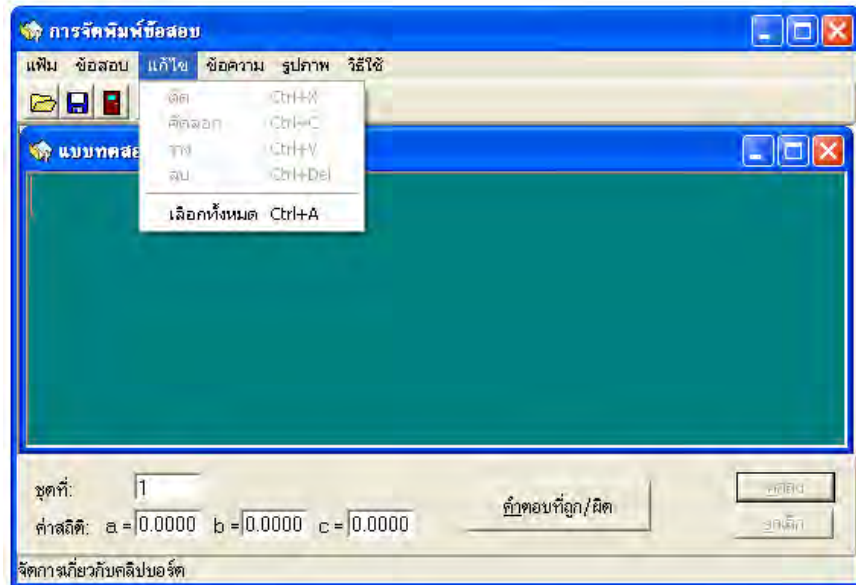
ภาพประกอบที่ 2 หน้าจอการจัดการพิมพ์ข้อสอบเมื่อเลื่อนเมาส์ไปที่คำสั่งแฟ้ม

2.2 คำสั่งข้อสอบ ประกอบด้วย สร้างข้อสอบใหม่ ย้ายข้อปัจจุบัน ลบข้อปัจจุบัน บันทึก ยกเลิก ไปข้อแรก ไปข้อก่อนหน้า ไปข้อถัดไป ไปข้อสุดท้าย และไปข้อที่ ดังแสดงในภาพประกอบที่ 3



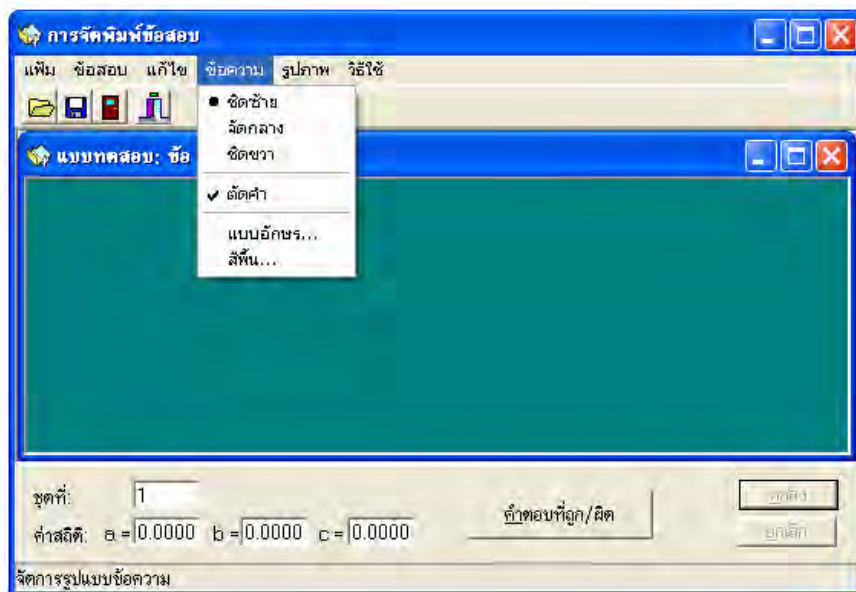
ภาพประกอบที่ 3 หน้าจอการจัดพิมพ์ข้อสอบเมื่อเลื่อนเมาส์ไปที่คำสั่งข้อสอบ

2.3 คำสั่งแก้ไข ประกอบด้วย ตัด คัดลอก วาง ลบ และเลือกทั้งหมด ดังแสดง  
ในภาพประกอบที่ 4



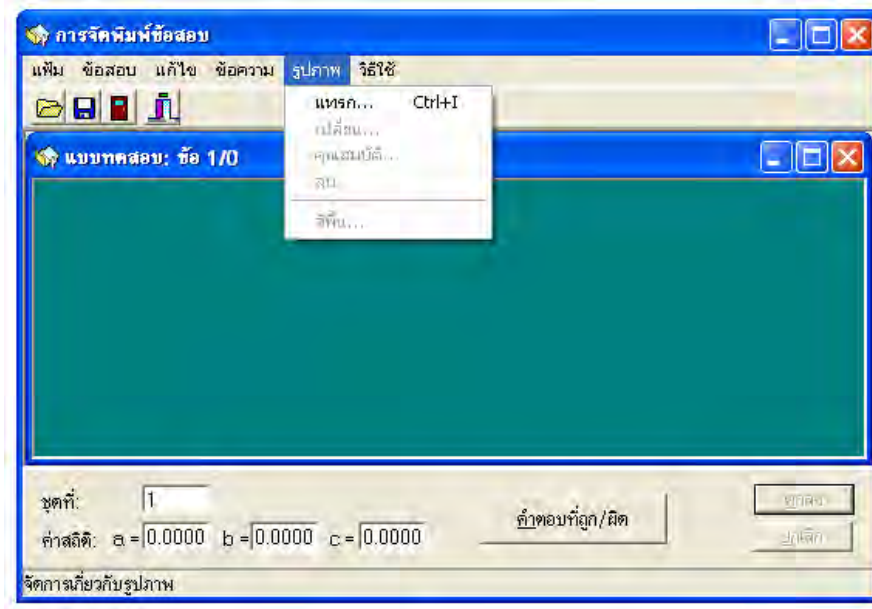
ภาพประกอบที่ 4 หน้าจอการจัดพิมพ์ข้อสอบเมื่อเลื่อนเมาส์ไปที่คำสั่งแก้ไข

2.4 คำสั่งข้อความ ประกอบด้วย ชิดซ้าย จัดกลาง ชิดขวา ตัดคำ แบบอักษร และสีพื้น  
ดังแสดงในภาพประกอบที่ 5



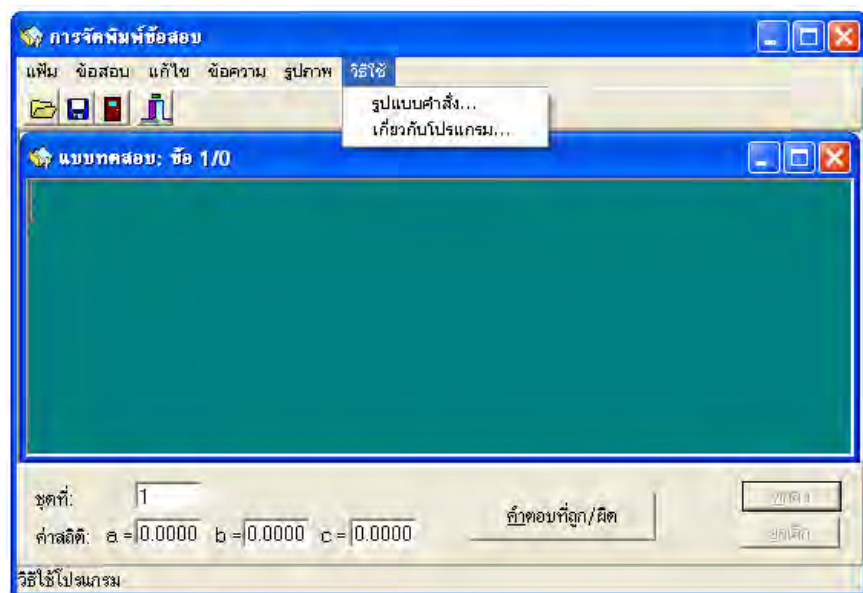
ภาพประกอบที่ 5 หน้าจอการจัดพิมพ์ข้อสอบเมื่อเลื่อนเมาส์ไปที่คำสั่งข้อความ

2.5 คำสั่งรูปภาพ ประกอบด้วย แทรก เปลี่ยน คุณสมบัติ ลบ และสีพื้น ดังแสดงในภาพประกอบที่ 6



ภาพประกอบที่ 6 หน้าจอการจัดพิมพ์ข้อสอบเมื่อเลื่อนเมาส์ไปที่คำสั่งรูปภาพ

2.6 คำสั่งวิธีใช้ ประกอบด้วย รูปแบบคำสั่งและเกี่ยวกับโปรแกรม ดังแสดงในภาพประกอบที่ 7

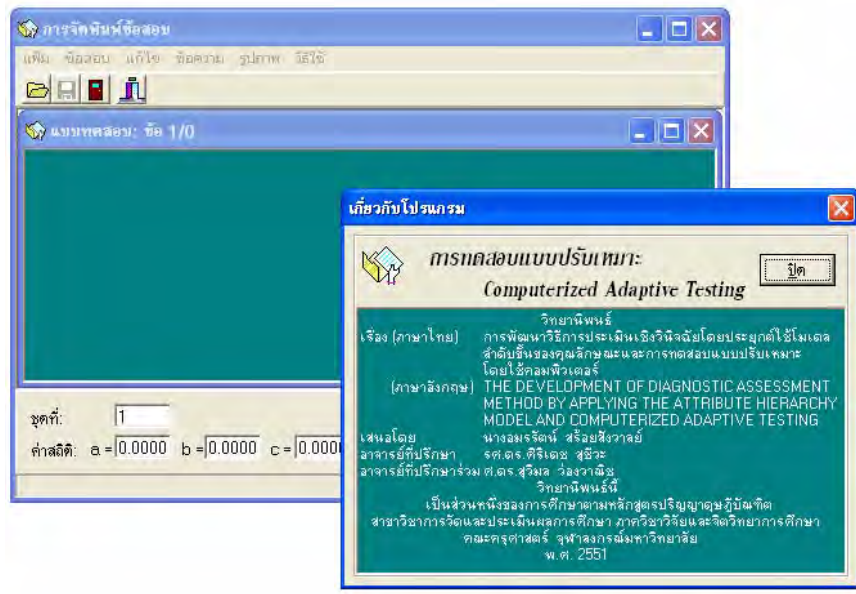


ภาพประกอบที่ 7 หน้าจอการจัดพิมพ์ข้อสอบเมื่อเลื่อนเมาส์ไปที่คำสั่งวิธีใช้

## ขั้นตอนการใช้โปรแกรมการจัดพิมพ์ข้อสอบ

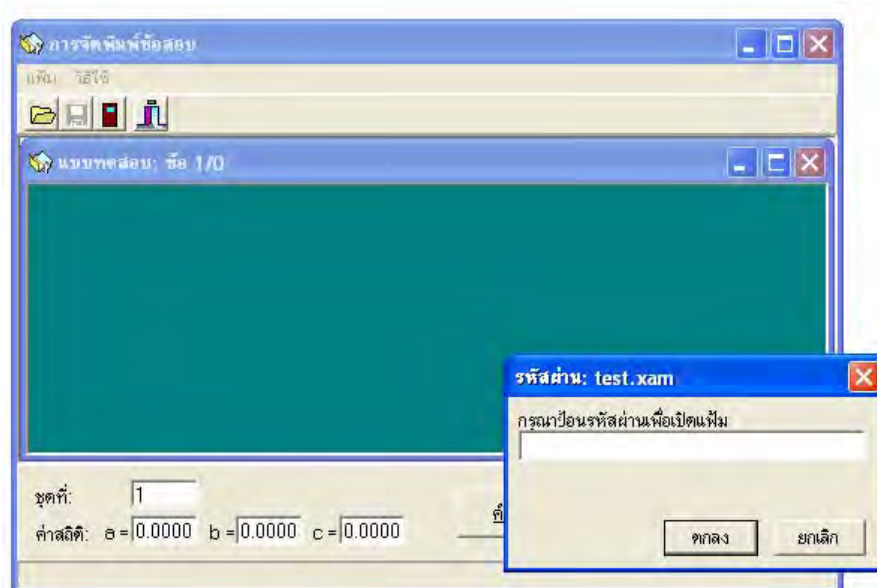
ขั้นตอนการใช้โปรแกรมนำเสนอโดยใช้ภาพประกอบคำบรรยาย ดังนี้

เมื่อดับเบิลคลิกที่ edit test ในระบบปฏิบัติการวินโดวส์ หรือ พิมพ์ test/d กดแป้น enter ในระบบปฏิบัติการ Dos จะปรากฏหน้าจอ ดังภาพประกอบที่ 8



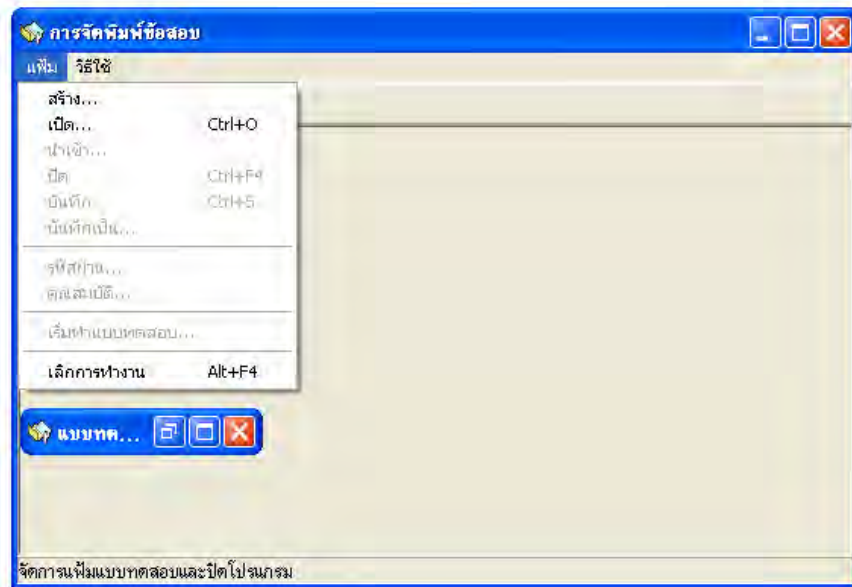
ภาพประกอบที่ 8 โลโก้หน้าแรกแสดงโปรแกรมการจัดพิมพ์ข้อสอบ

เมื่อคลิกปุ่มเปิด หรือกด Enter บนแป้นพิมพ์ โปรแกรมจะแสดงหน้าจอถัดไปให้ป้อนรหัสผ่านเพื่อเข้าสู่โปรแกรมการจัดพิมพ์ข้อสอบ ดังภาพประกอบที่ 9



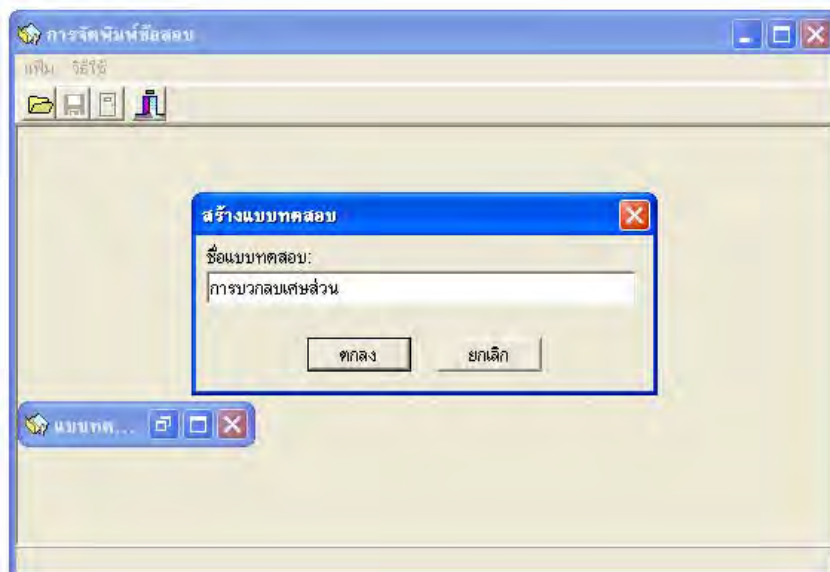
ภาพประกอบที่ 9 หน้าจอของโปรแกรมเพื่อป้อนรหัสผ่าน

ผู้ใช้โปรแกรมสามารถสร้างแบบทดสอบใหม่ได้ โดยคลิกที่แป้นตกลง โปรแกรมจะแสดงหน้าจอจัดการเพิ่มแบบทดสอบและเปิดโปรแกรม เมื่อคลิกที่คำสั่งเพิ่ม จะปรากฏคำสั่งย่อยที่พร้อมการใช้งาน 3 คำสั่ง ได้แก่ สร้าง เปิด และเลิกการทำงาน ดังภาพประกอบที่ 10



ภาพประกอบที่ 10 หน้าจอของโปรแกรมเมื่อเลื่อนเมาส์ไปที่คำสั่งเพิ่ม

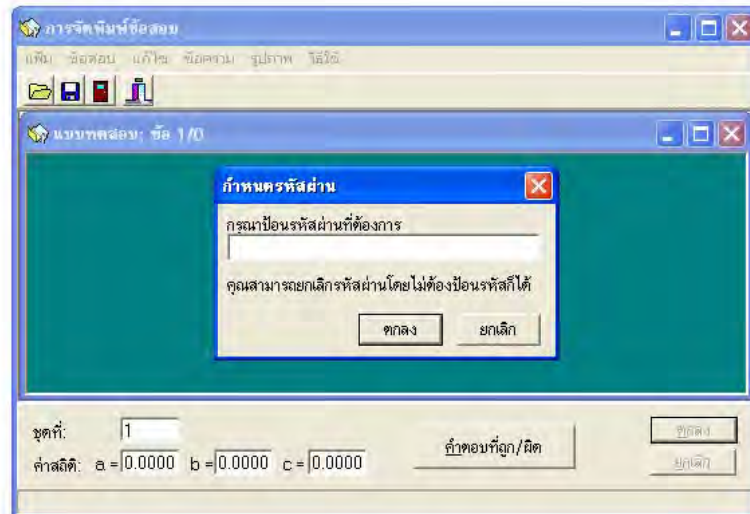
กรณีต้องการสร้างแบบทดสอบจากแบบทดสอบที่มีอยู่แล้วให้คลิกที่คำสั่งเปิดเพื่อเลือกเปิดแบบทดสอบที่ต้องการ กรณีที่ต้องการสร้างแบบทดสอบใหม่ให้คลิกที่คำสั่งสร้าง โปรแกรมจะปรากฏหน้าจอให้ระบุชื่อแบบทดสอบ ดังภาพประกอบที่ 11





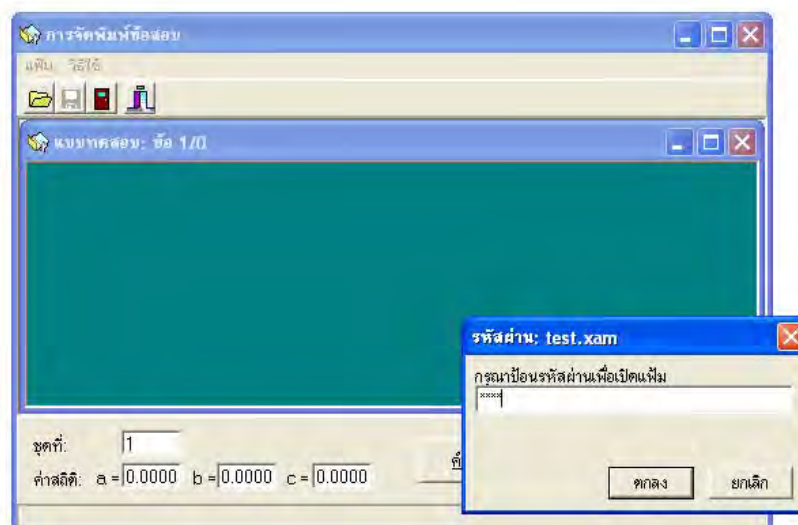
ภาพประกอบที่ 11 หน้าจอของโปรแกรมเพื่อระบุชื่อแบบทดสอบ

เมื่อคลิกที่ปุ่มตกลง โปรแกรมจะแสดงหน้าจอเพื่อให้สร้างข้อสอบข้อที่ 1 กรณีที่ต้องการกำหนดรหัสผ่านให้เลื่อนเมาส์ไปที่คำสั่งแฟ้ม และคลิกที่คำสั่ง รหัสผ่าน โปรแกรมจะปรากฏหน้าจอให้กำหนดรหัสผ่าน ดังภาพประกอบที่ 12



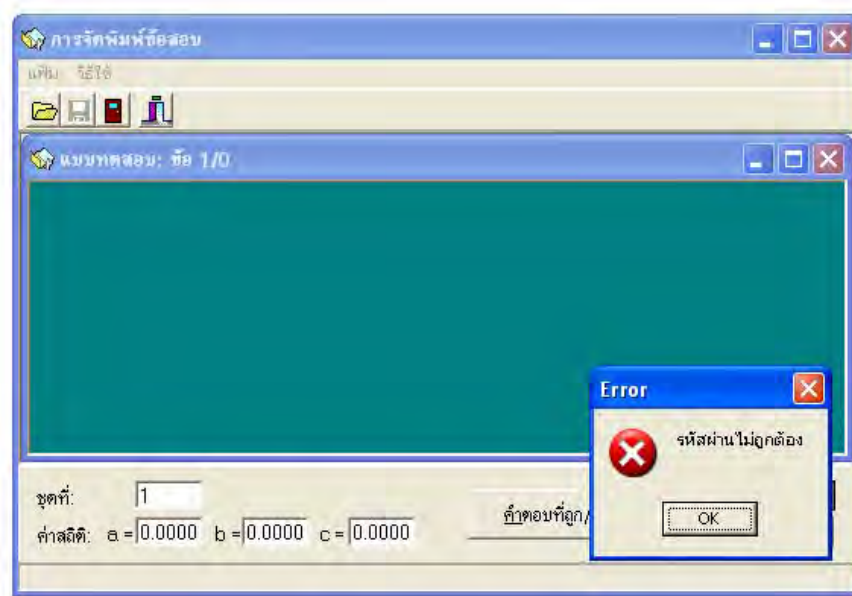
ภาพประกอบที่ 12 หน้าจอของโปรแกรมเพื่อกำหนดรหัสผ่าน

เมื่อพิมพ์รหัสผ่านและคลิกที่ปุ่มตกลง โปรแกรมจะกลับไปหน้าจอเพื่อให้สร้างข้อสอบข้อที่ 1 จากภาพประกอบที่ 9 กรณีที่ต้องการเพิ่ม ลด หรือแก้ไขปรับปรุงข้อสอบจากแบบทดสอบที่มีอยู่แล้ว ให้พิมพ์รหัสผ่านเพื่อเปิดแฟ้ม ดังภาพประกอบที่ 13



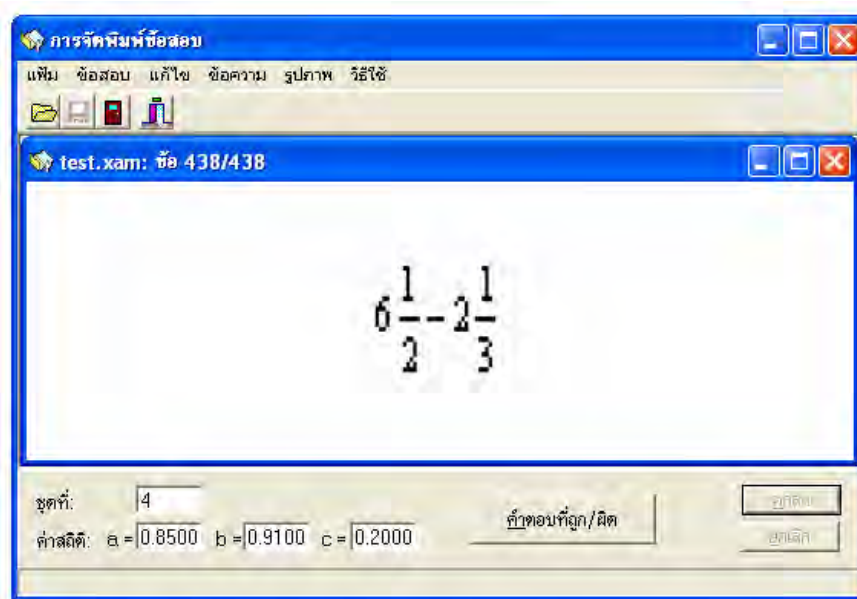
ภาพประกอบที่ 13 หน้าจอของโปรแกรมเพื่อป้อนรหัสผ่าน

เมื่อป้อนรหัสผ่านเรียบร้อยแล้ว ให้กดปุ่มตกลง กรณีที่ป้อนรหัสผ่านไม่ถูกต้อง โปรแกรมจะปรากฏหน้าจอคำสั่ง Error เพื่อเตือนผู้ใช้โปรแกรมให้แก้ไขรหัสผ่านให้ถูกต้อง ดังภาพประกอบที่ 14



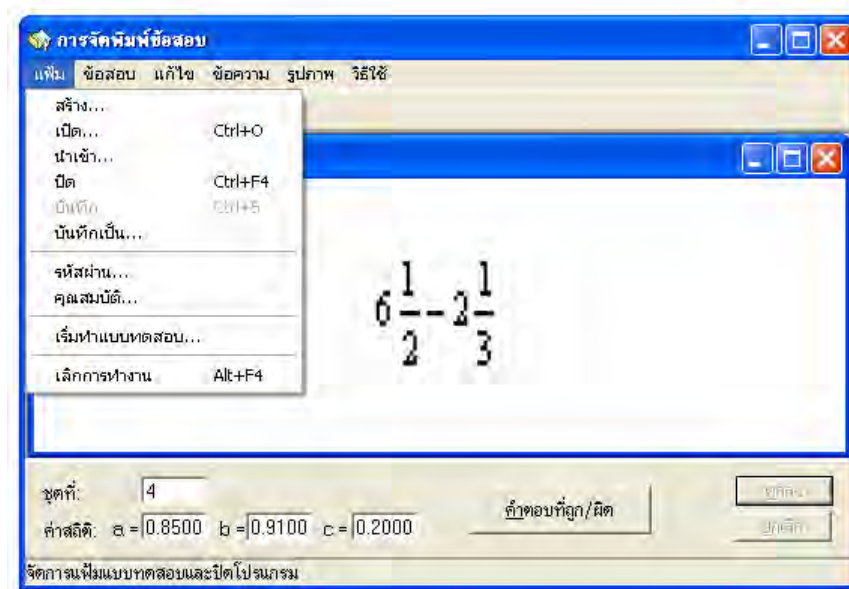
ภาพประกอบที่ 14 หน้าจอของโปรแกรมเพื่อเตือนการระบุรหัสผ่านไม่ถูกต้อง

เมื่อป้อนรหัสผ่านได้ถูกต้องและคลิกเมาส์ที่ปุ่มตกลง โปรแกรมจะให้ผ่านเข้าสู่โปรแกรมการจัดการพิมพ์ข้อสอบ ดังภาพประกอบที่ 15



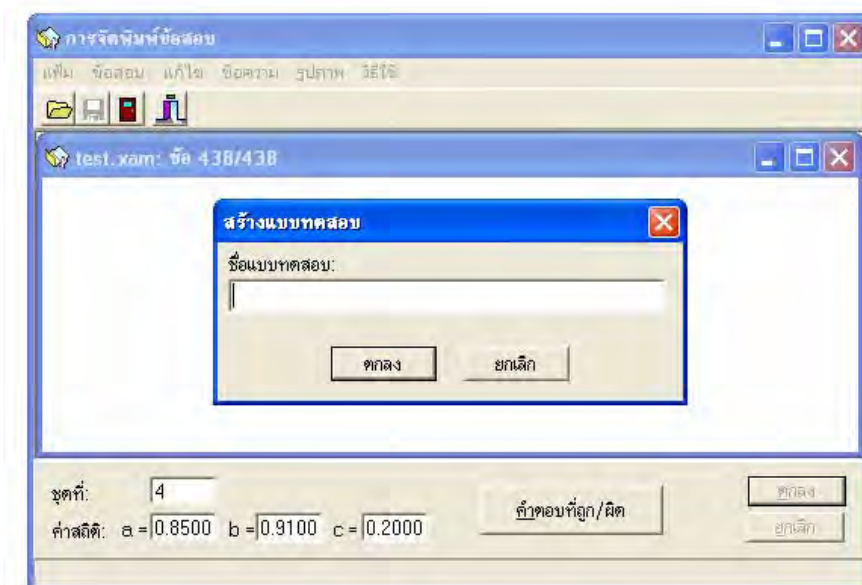
ภาพประกอบที่ 15 หน้าจอแสดงโปรแกรมการจัดการพิมพ์ข้อสอบ

เมื่อคลิกที่คำสั่งแฟ้ม โปรแกรมจะปรากฏหน้าจอแสดงคำสั่งที่พร้อมใช้งาน 9 คำสั่ง ดังภาพประกอบที่ 16



ภาพประกอบที่ 16 หน้าจอโปรแกรมแสดงคำสั่งแฟ้ม-สร้าง

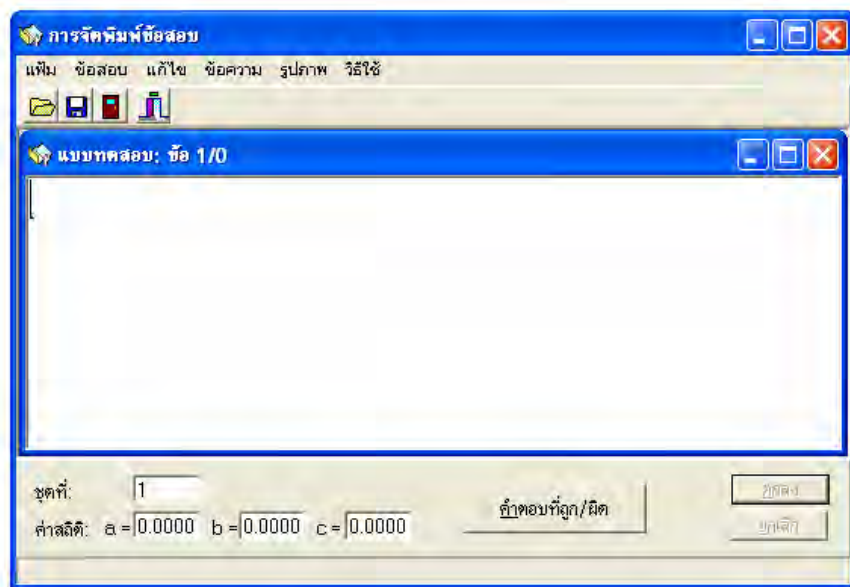
เมื่อคลิกที่คำสั่งสร้าง โปรแกรมจะกำหนดให้ระบุชื่อแบบทดสอบ ดังภาพประกอบที่ 17



ภาพประกอบที่ 17 หน้าจอโปรแกรมเพื่อระบุชื่อแบบทดสอบ (2)

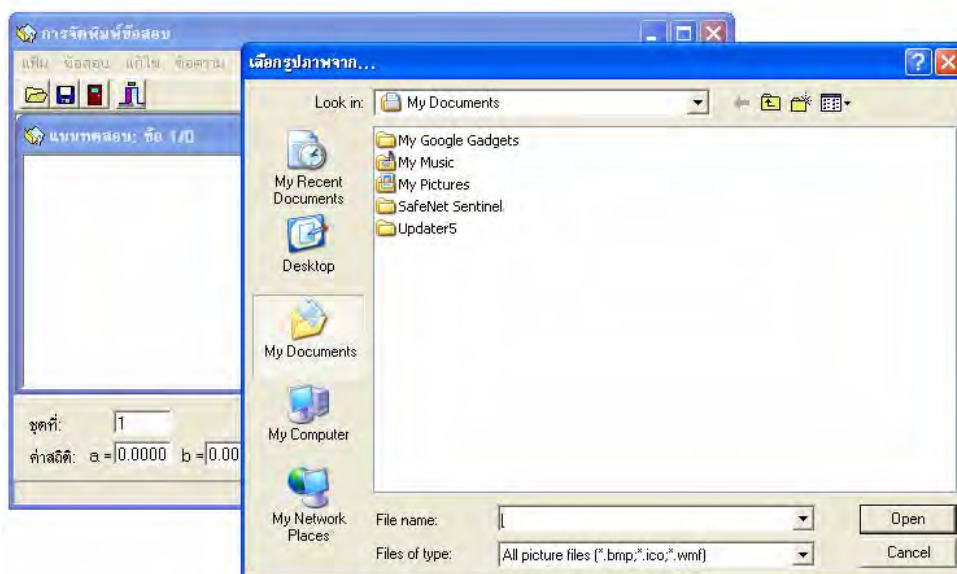


เมื่อระบุชื่อแบบทดสอบเรียบร้อยแล้วให้คลิกที่ปุ่มตกลง โปรแกรมจะปรากฏหน้าจอเพื่อให้สร้างข้อสอบข้อที่ 1 ดังภาพประกอบที่ 18



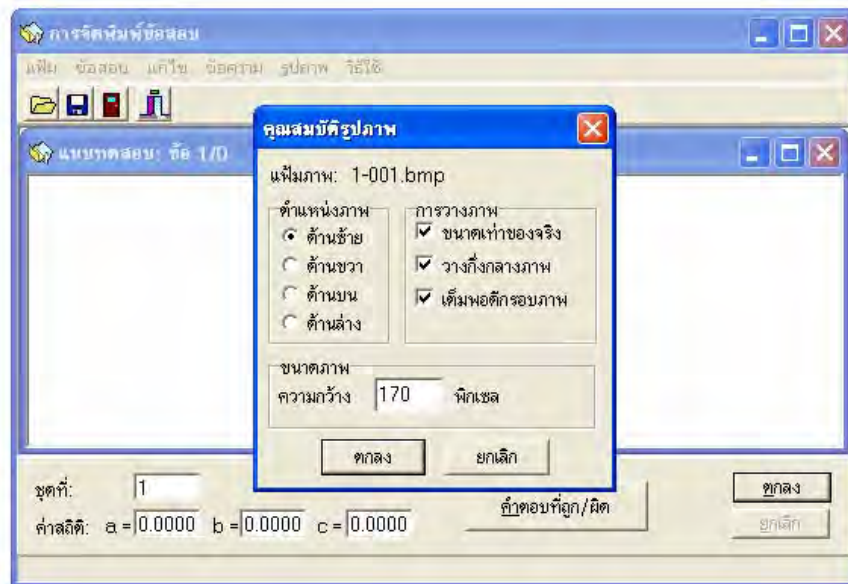
ภาพประกอบที่ 18 หน้าจอโปรแกรมเพื่อจัดพิมพ์ข้อสอบ

เมื่อต้องการสร้างข้อสอบ ให้คลิกที่คำสั่งรูปภาพ เลือกคำสั่งแทรก โปรแกรมจะปรากฏหน้าจอให้เลือกแฟ้มที่เก็บรูปภาพข้อสอบ ดังภาพประกอบที่ 19



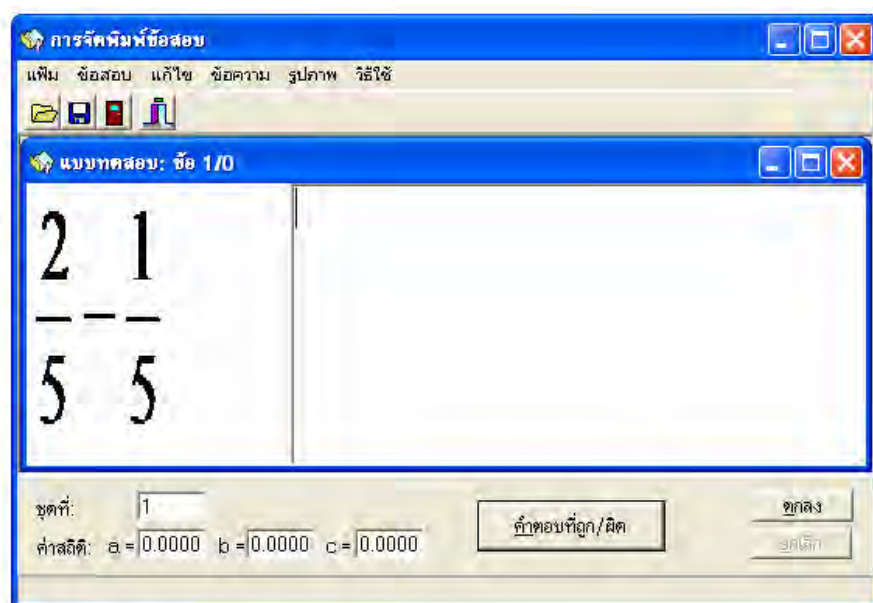
ภาพประกอบที่ 19 หน้าจอโปรแกรมเพื่อเลือกแฟ้มที่เก็บรูปภาพข้อสอบ

เมื่อเลือกรูปภาพข้อสอบและคลิกที่ปุ่ม Open โปรแกรมจะแสดงหน้าจอให้กำหนดคุณสมบัติรูปภาพ ดังภาพประกอบที่ 20



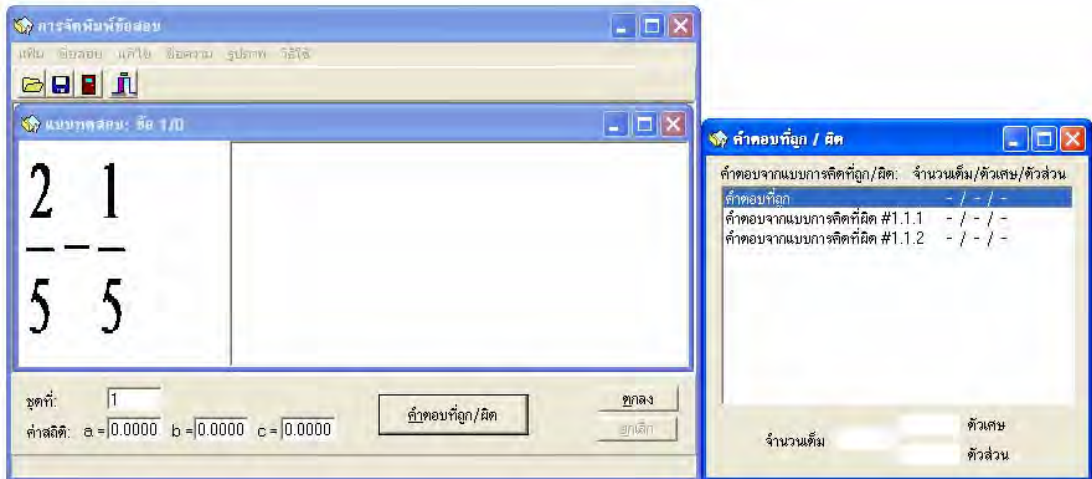
ภาพประกอบที่ 20 หน้าจอโปรแกรมเพื่อกำหนดคุณสมบัติรูปภาพ

เมื่อกำหนดคุณสมบัติรูปภาพและคลิกที่ปุ่มตกลง โปรแกรมจะปรากฏหน้าจอแสดงรูปภาพข้อสอบที่เลือกตามคุณสมบัติที่กำหนด ดังภาพประกอบที่ 21



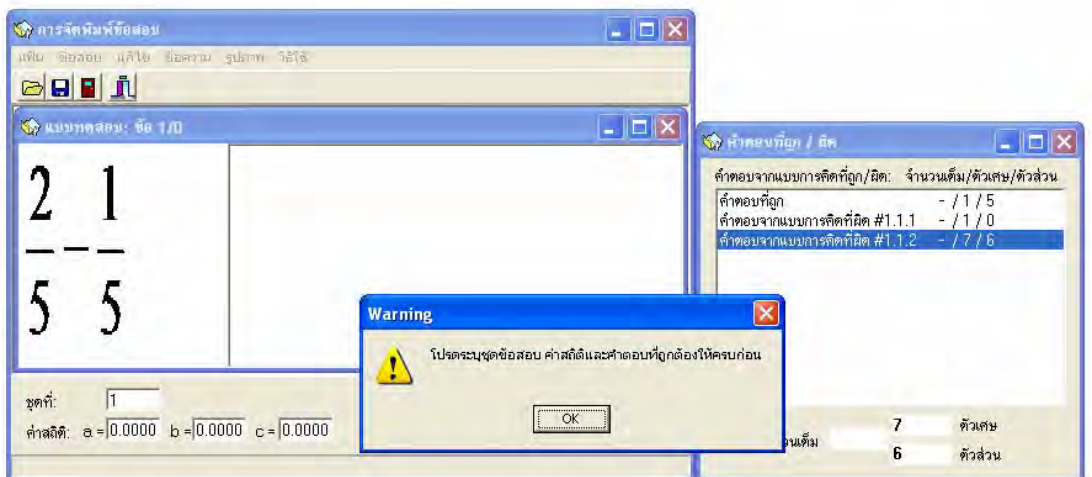
ภาพประกอบที่ 21 หน้าจอโปรแกรมแสดงรูปภาพข้อสอบที่เลือกตามคุณสมบัติ

เมื่อคลิกที่ปุ่ม คำตอบที่ถูก/ผิด โปรแกรมจะปรากฏหน้าจอให้บันทึกคำตอบที่ถูกและคำตอบที่ผิด ดังภาพประกอบที่ 22



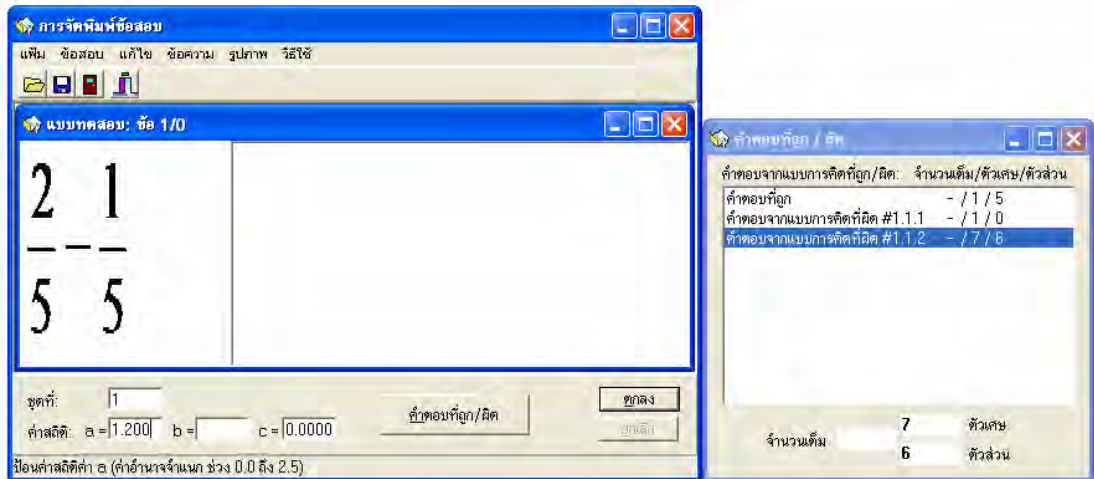
ภาพประกอบที่ 22 หน้าจอโปรแกรมเพื่อบันทึกคำตอบที่ถูกและคำตอบที่ผิด

เมื่อบันทึกคำตอบที่ถูก/ผิด เรียบร้อยแล้ว แต่ยังไม่ได้รับชุดของข้อสอบและค่าสถิติ โปรแกรมจะเตือนให้ระบุชุดของข้อสอบ ค่าสถิติและคำตอบที่ถูกต้องให้ครบถ้วน ดังภาพประกอบที่ 23

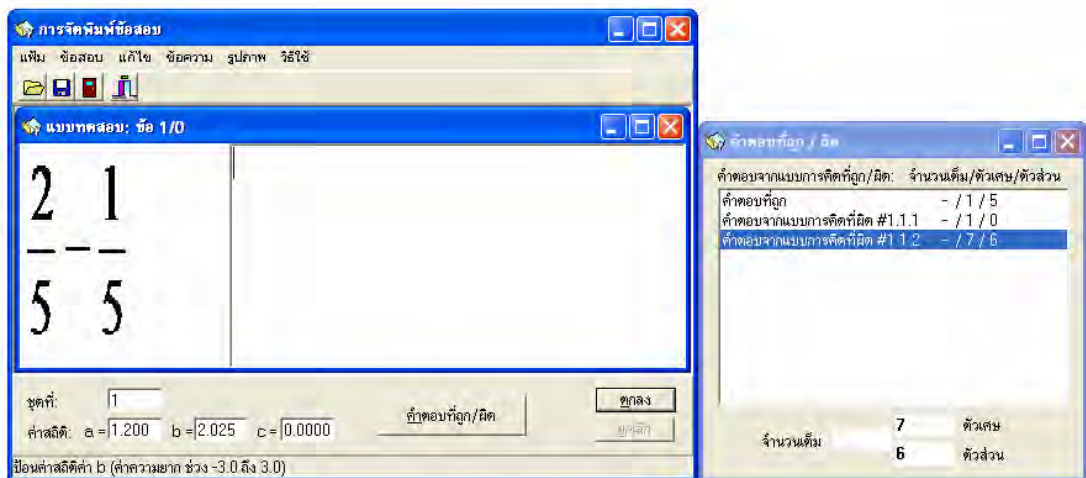


ภาพประกอบที่ 23 หน้าจอโปรแกรมเพื่อเตือนให้ระบุชุดของข้อสอบ ค่าสถิติและคำตอบที่ถูกต้องให้ครบถ้วน

ในขณะที่บันทึกค่าสถิติแต่ละค่า จะมีขอบเขตของค่าสถิติที่กำหนดไว้ แสดงบนแถบหน้าจอด้านล่าง เพื่อเตือนไม่ให้นักศึกษาผิดพลาด ดังแสดงในภาพประกอบที่ 24,25 และ 26

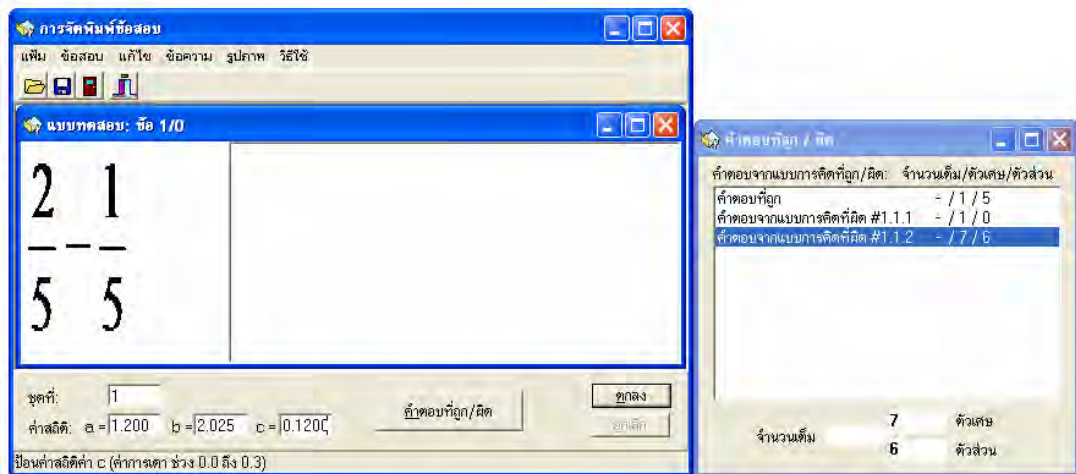


ภาพประกอบที่ 24 หน้าจอโปรแกรมเพื่อบันทึกค่าอำนาจจำแนก



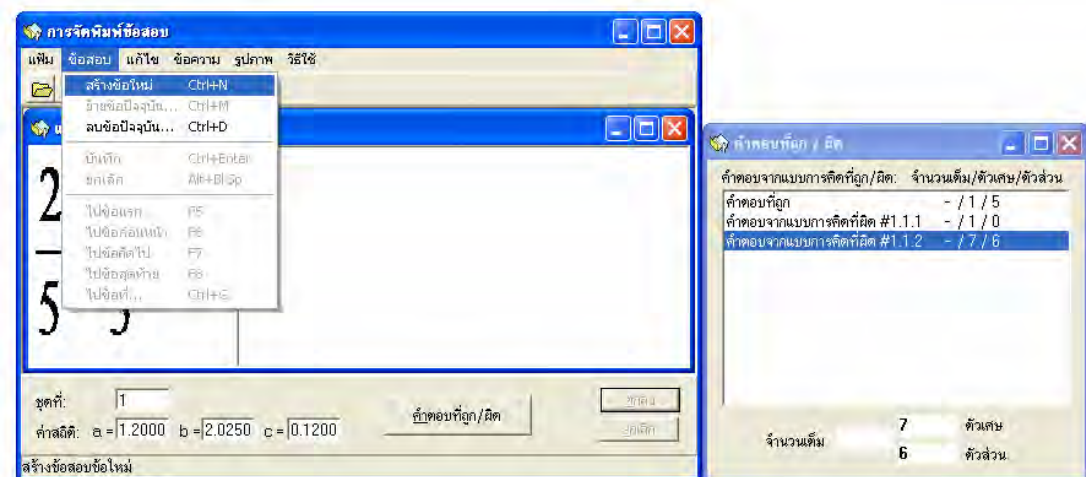
ภาพประกอบที่ 25 หน้าจอโปรแกรมเพื่อบันทึกค่าความยาก



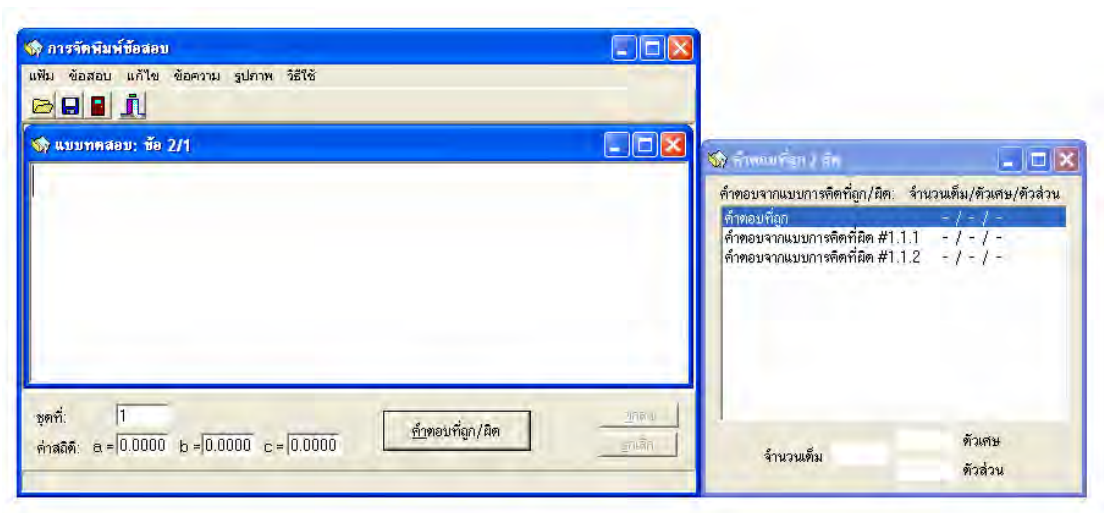


ภาพประกอบที่ 26 หน้าจอโปรแกรมเพื่อบันทึกค่าการเดา

เมื่อบันทึกค่าที่กำหนดครบถ้วนแล้ว ให้เลื่อนเมาส์ไปที่คำสั่งข้อสอบ และคลิกที่คำสั่งสร้างข้อใหม่ ดังแสดงในภาพประกอบที่ โปรแกรมจะแสดงหน้าจอให้สร้างข้อสอบข้อที่ 2 ดังภาพประกอบที่ 27 และ 28

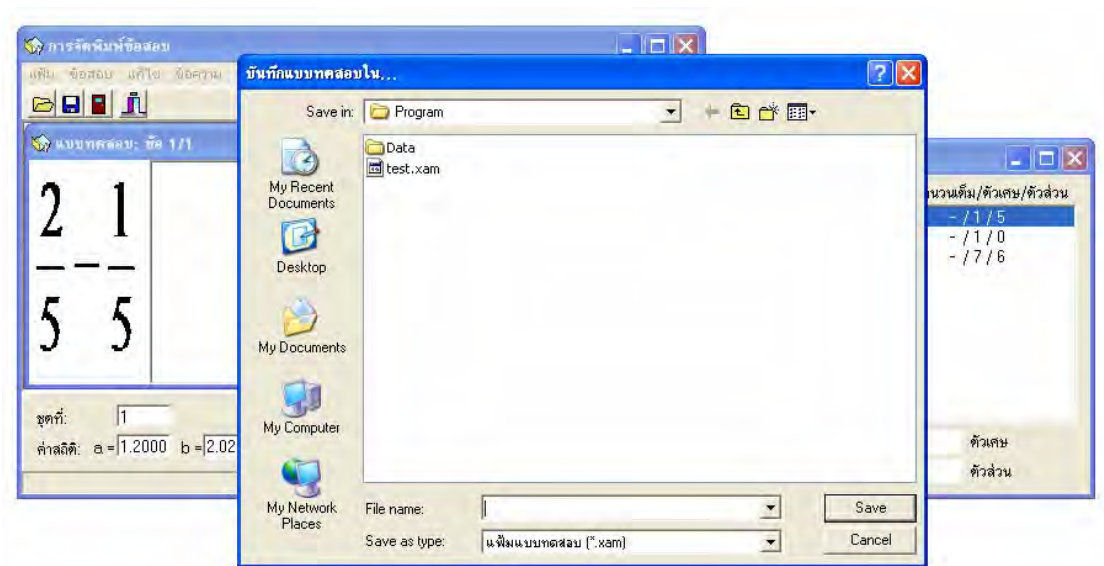


ภาพประกอบที่ 27 หน้าจอโปรแกรมเมื่อเลื่อนเมาส์ไปที่คำสั่งข้อสอบ



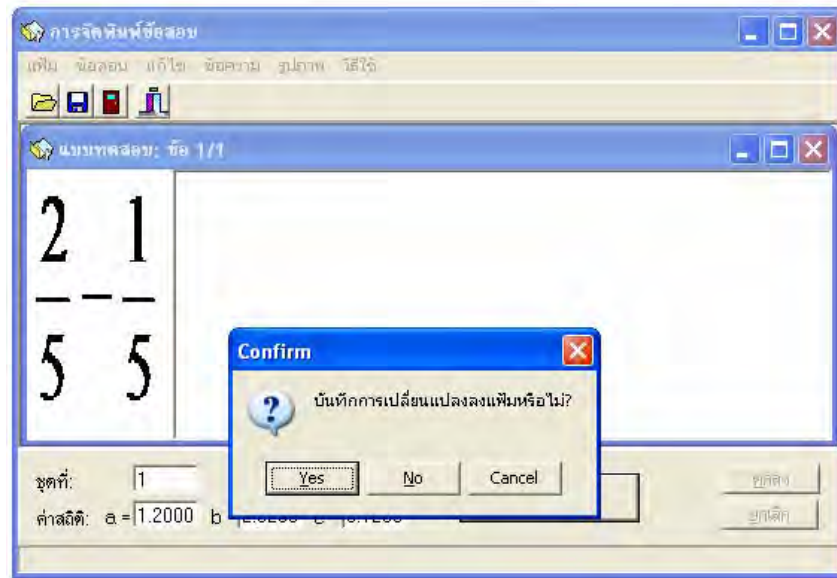
ภาพประกอบที่ 28 หน้าจอโปรแกรมเพื่อให้สร้างข้อสอบข้อที่ 2

เมื่อต้องการเลิกการทำงาน ให้เลือกคำสั่งเพิ่ม และคลิกที่คำสั่งบันทึก โปรแกรมจะปรากฏหน้าจอให้เลือกตำแหน่งที่จะบันทึกเพิ่ม ดังแสดงในภาพประกอบที่ 29



ภาพประกอบที่ 29 หน้าจอโปรแกรมเพื่อเลือกตำแหน่งที่จะบันทึกเพิ่ม

กรณีที่คลิกที่ภาพกากบาทเพื่อปิดโปรแกรม โดยที่ยังไม่มีการบันทึก โปรแกรมจะปรากฏ หน้าจอเตือน ให้บันทึกแฟ้มก่อนปิด ดังแสดงในภาพประกอบที่ 30



ภาพประกอบที่ 30 หน้าจอโปรแกรมเพื่อให้เลือกบันทึกแฟ้มก่อนปิด

## โปรแกรมการทดสอบ

โปรแกรมการทดสอบเป็นโปรแกรมที่ใช้ในการทดสอบแบบปรับเหมาะ จำแนกการนำเสนอเป็น 2 ส่วน ส่วนแรกเป็นองค์ประกอบที่สำคัญของหน้าจอและคำสั่งในการทดสอบ ส่วนที่ 2 เป็นขั้นตอนการใช้โปรแกรมการทดสอบ มีรายละเอียด ดังนี้

### องค์ประกอบที่สำคัญของหน้าจอและคำสั่งในการทดสอบ

#### 1. องค์ประกอบที่สำคัญของหน้าจอการทดสอบ

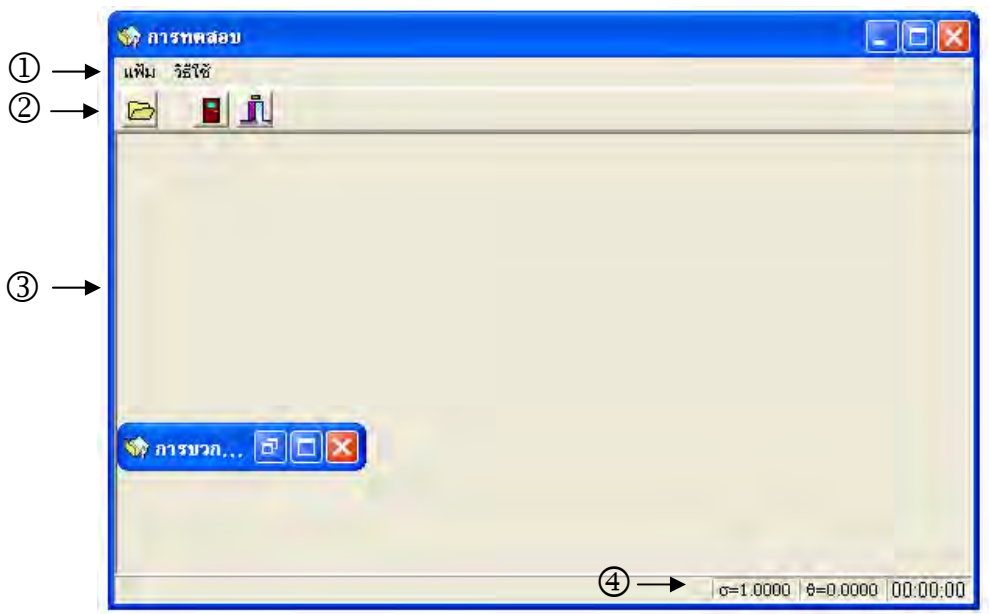
องค์ประกอบที่สำคัญของหน้าจอการทดสอบ ดังแสดงในภาพประกอบที่ 31 มีดังนี้

- ① กลุ่มคำสั่ง ประกอบด้วย แฟ้ม และวิธีใช้
- ② กลุ่มไอคอน ซึ่งเป็นภาพแทนคำสั่ง ประกอบด้วย รูปภาพแทนคำสั่ง เปิด ปิด และเลิกการทำงาน

③ พื้นที่ว่างเพื่อแสดงข้อสอบ

④ กลุ่มค่าสถิติ ประกอบด้วย ค่าความคลาดเคลื่อนมาตรฐาน ค่าระดับ

ความสามารถของผู้สอบและเวลาที่ใช้ในการสอบ



ภาพประกอบที่ 31 หน้าจอการทดสอบ



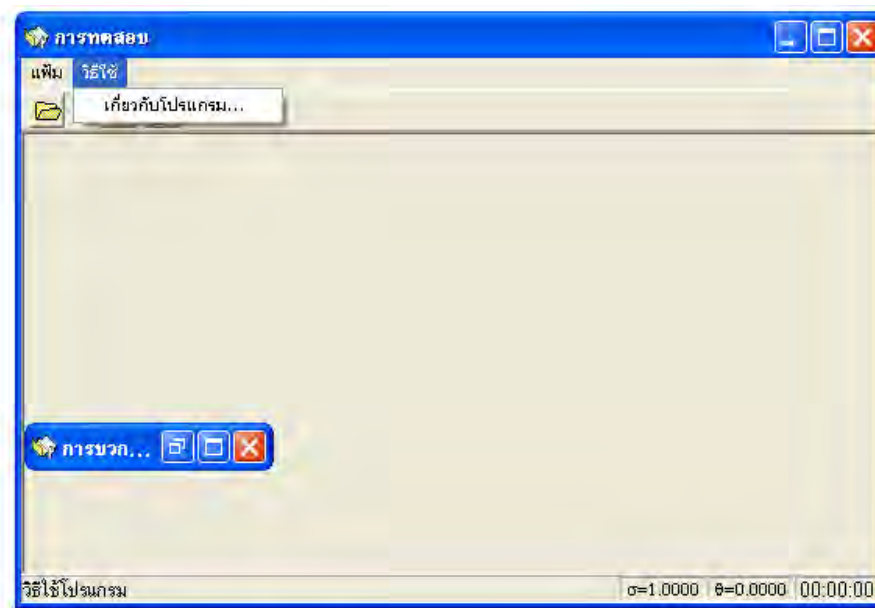
## 2. คำสั่งในการทดสอบ

2.1 คำสั่งเพิ่ม ประกอบด้วย สร้าง เปิด ปิด และเลิกการทำงาน ดังแสดงในภาพประกอบที่ 32



ภาพประกอบที่ 32 หน้าจอการทดสอบเมื่อเลื่อนเมาส์ไปที่คำสั่งเพิ่ม

2.1 คำสั่งวิธีใช้ ประกอบด้วย เกี่ยวกับโปรแกรม ดังแสดงในภาพประกอบที่ 33

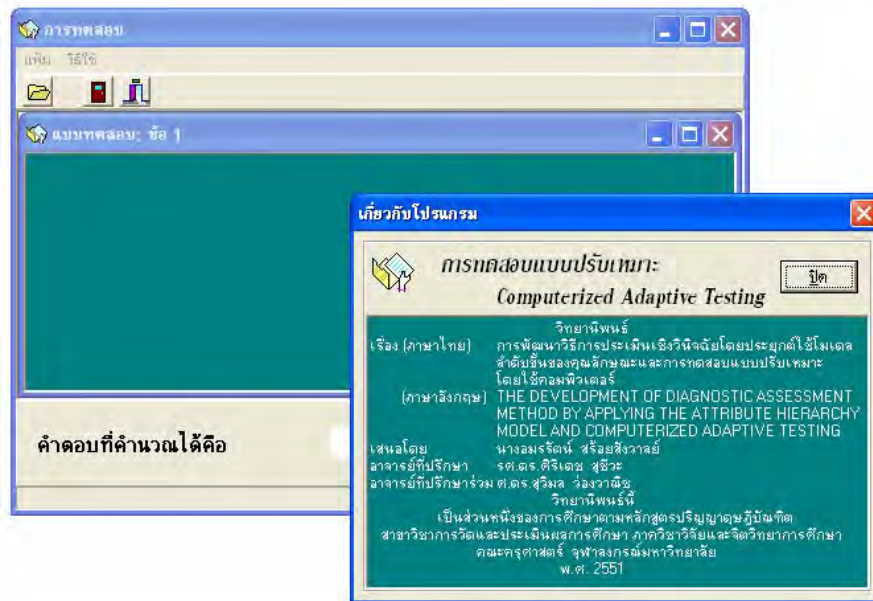


ภาพประกอบที่ 33 หน้าจอการทดสอบเมื่อเลื่อนเมาส์ไปที่คำสั่งวิธีใช้

## ขั้นตอนการใช้โปรแกรมการทดสอบ

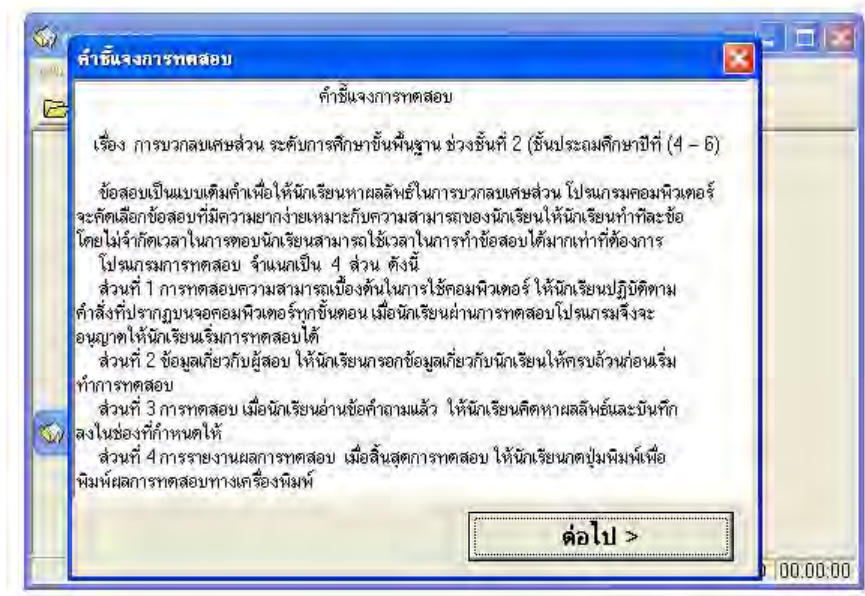
ขั้นตอนการใช้โปรแกรมการทดสอบนำเสนอโดยใช้ภาพประกอบคำบรรยาย ดังนี้

เมื่อดับเบิลคลิกที่ do test 1 ถึง do test 6 ในระบบปฏิบัติการวินโดว หรือ พิมพ์ test 1 ถึง test 6 กดแป้น enter ในระบบปฏิบัติการ Dos จะปรากฏหน้าจอ ดังภาพประกอบที่ 34



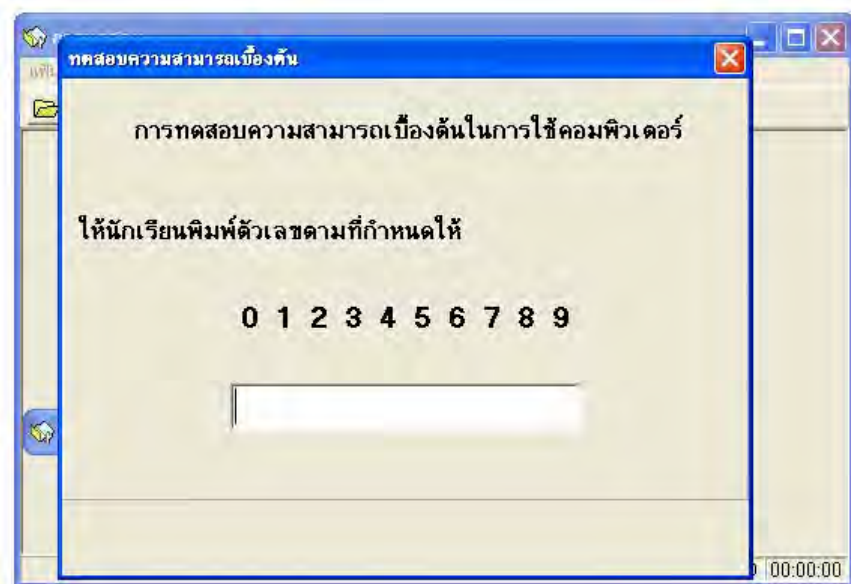
ภาพประกอบที่ 34 โลโก้หน้าแรกแสดงโปรแกรมการทดสอบ

เมื่อคลิกที่ปุ่มปิด โปรแกรมจะปรากฏหน้าจอคำชี้แจงการทดสอบ ดังภาพประกอบที่ 35

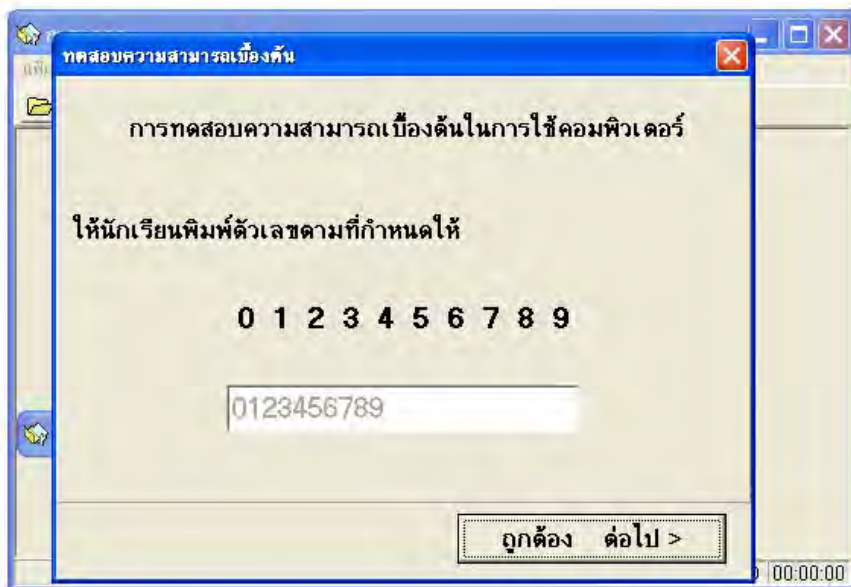


ภาพประกอบที่ 35 หน้าจอแสดงคำชี้แจงการทดสอบ

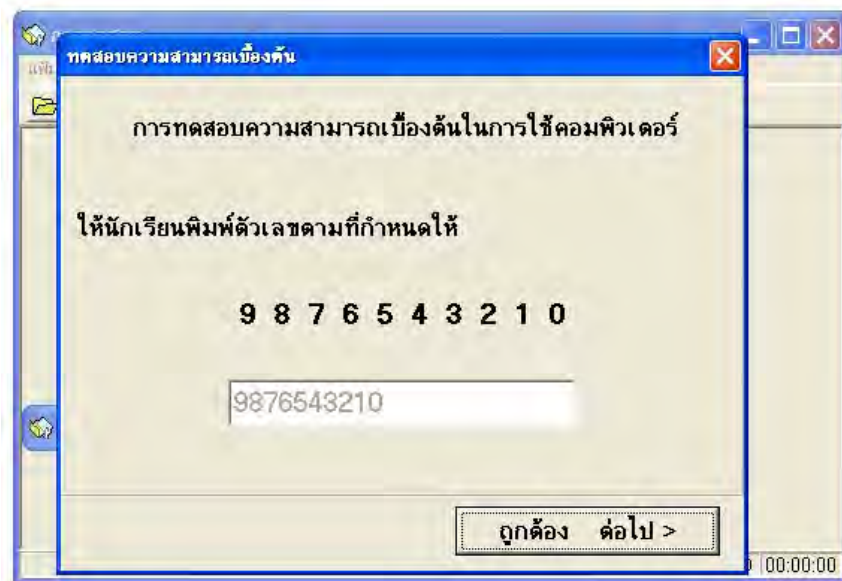
เมื่อคลิกที่ปุ่มต่อไป โปรแกรมจะปรากฏหน้าจอเพื่อทดสอบความสามารถในการใช้คอมพิวเตอร์ ดังแสดงในภาพประกอบที่ 36 ถึง 42



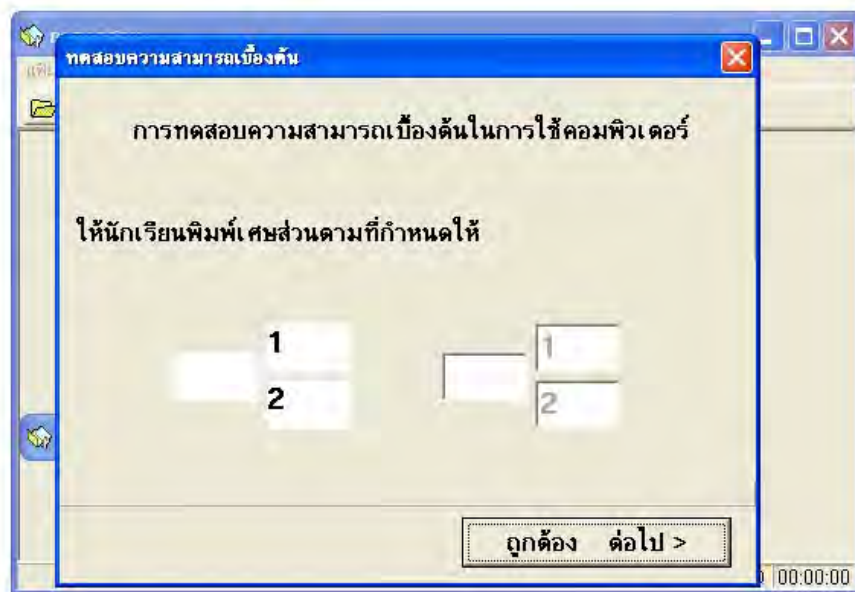
ภาพประกอบที่ 36 หน้าจอเพื่อทดสอบความสามารถในการใช้คอมพิวเตอร์ ภาพที่ 1



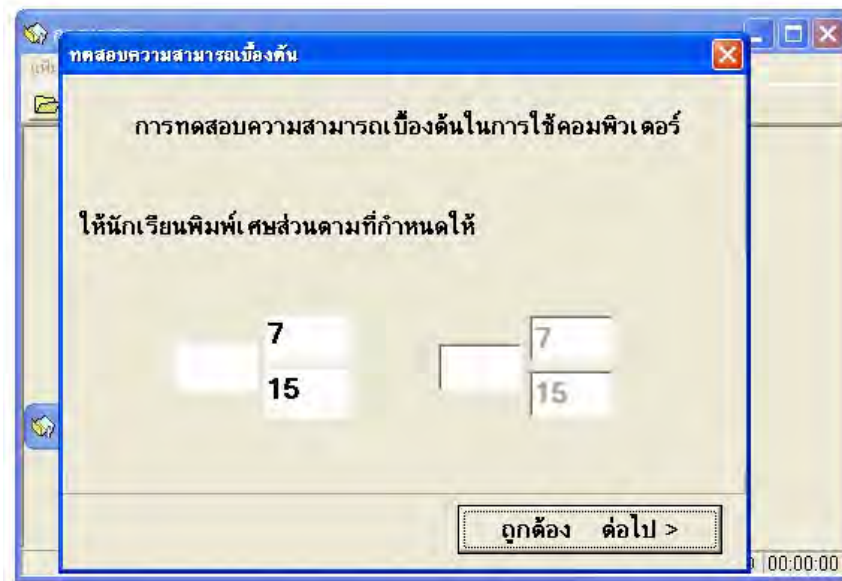
ภาพประกอบที่ 37 หน้าจอเพื่อทดสอบความสามารถในการใช้คอมพิวเตอร์ ภาพที่ 2



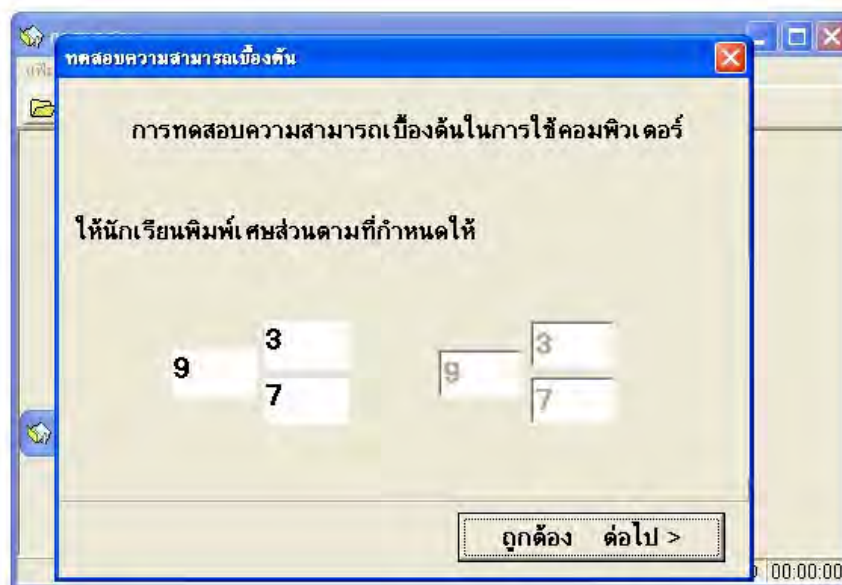
ภาพประกอบที่ 38 หน้าจอเพื่อทดสอบความสามารถในการใช้คอมพิวเตอร์ ภาพที่ 3



ภาพประกอบที่ 39 หน้าจอเพื่อทดสอบความสามารถในการใช้คอมพิวเตอร์ ภาพที่ 4

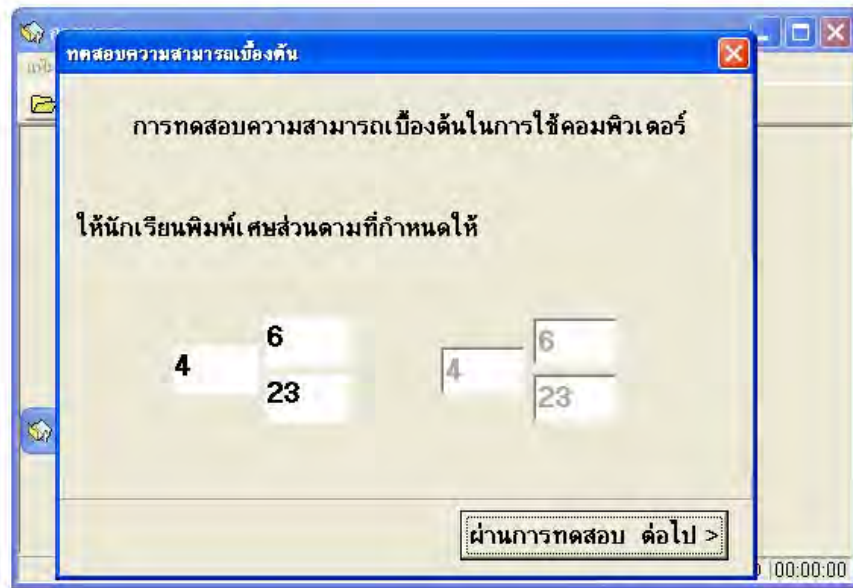


ภาพประกอบที่ 40 หน้าจอเพื่อทดสอบความสามารถในการใช้คอมพิวเตอร์ ภาพที่ 5



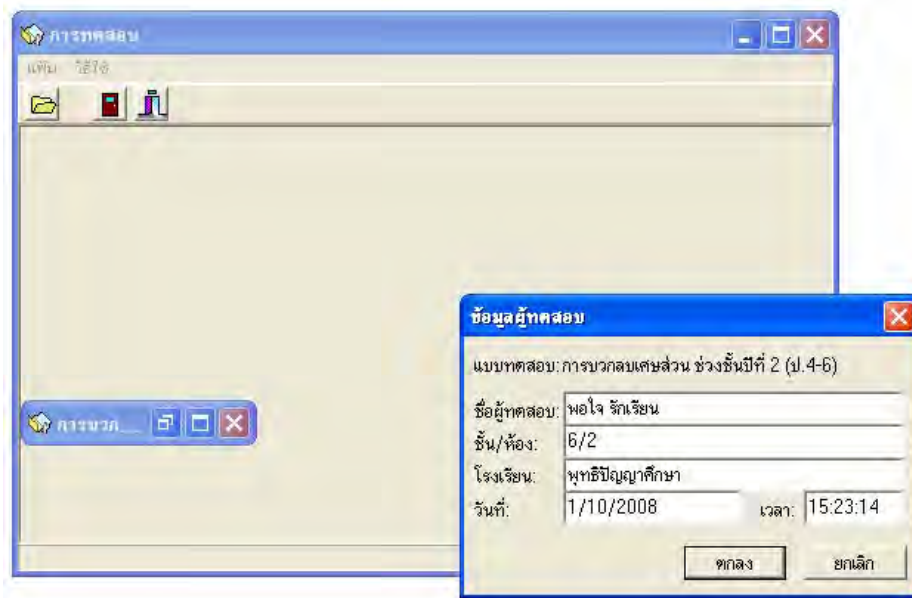
ภาพประกอบที่ 41 หน้าจอเพื่อทดสอบความสามารถในการใช้คอมพิวเตอร์ ภาพที่ 6





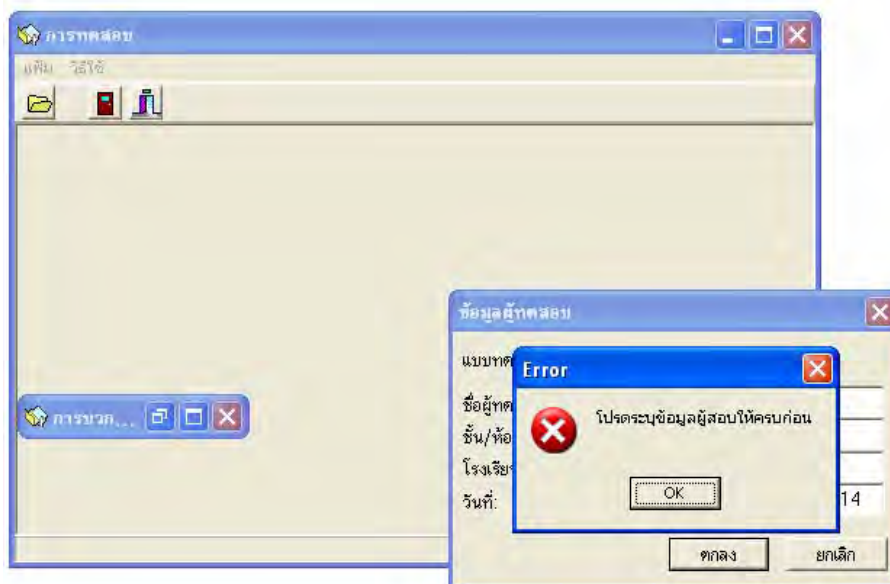
ภาพประกอบที่ 42 หน้าจอเพื่อทดสอบความสามารถในการใช้คอมพิวเตอร์ ภาพที่ 7

เมื่อผ่านการทดสอบความสามารถในการใช้คอมพิวเตอร์หน้าจอสุดท้าย ให้คลิกที่ปุ่ม ผ่านการทดสอบ ต่อไป > โปรแกรมจะปรากฏหน้าจอเพื่อให้ผู้สอบระบุข้อมูลผู้ทดสอบ ดังแสดงในภาพประกอบที่ 43



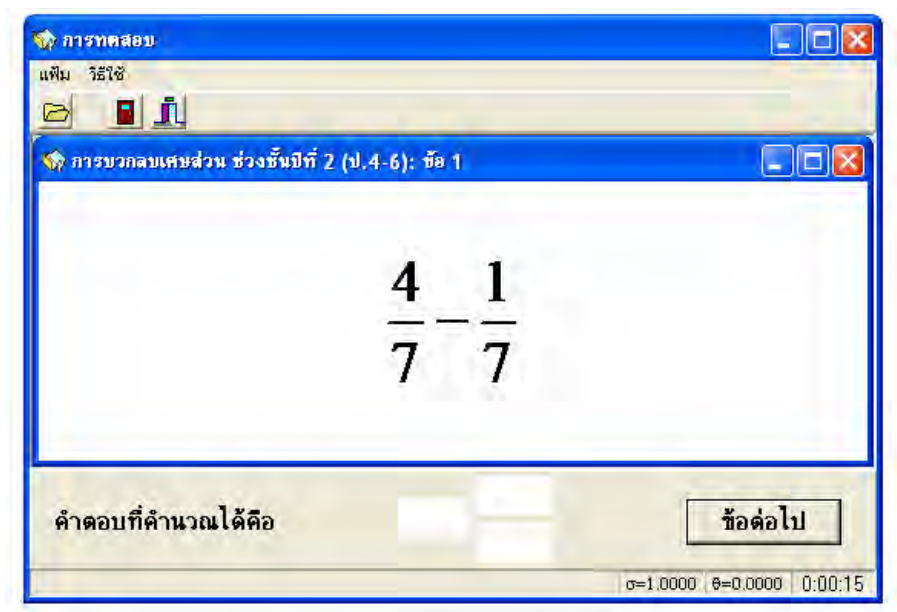
ภาพประกอบที่ 43 หน้าจอเพื่อให้ผู้สอบบันทึกข้อมูลผู้ทดสอบ

กรณีที่ผู้สอบระบุข้อมูลไม่ครบ โปรแกรมจะปรากฏหน้าจอคำสั่ง Error เพื่อเตือนผู้สอบให้ระบุข้อมูลผู้สอบให้ครบ ดังแสดงในภาพประกอบที่ 44



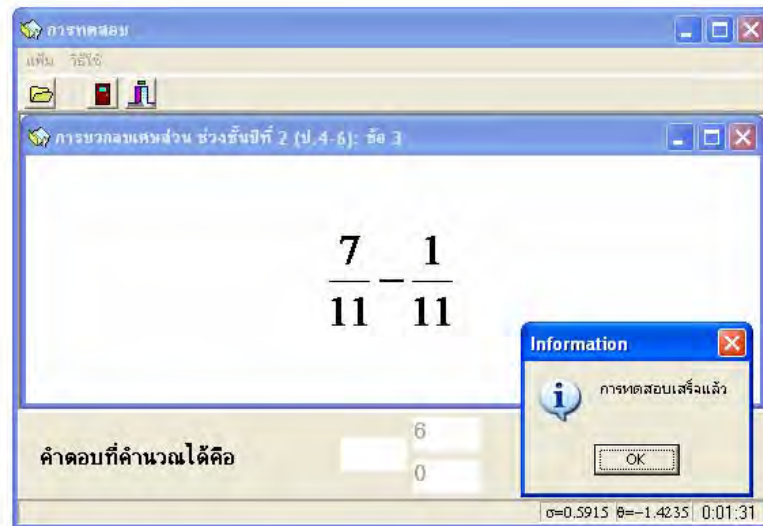
ภาพประกอบที่ 44 หน้าจอเพื่อให้ผู้สอบระบุข้อมูลผู้สอบให้ครบ

เมื่อระบุข้อมูลผู้ทดสอบครบแล้ว คลิกที่ปุ่มตกลง โปรแกรมจะปรากฏหน้าจอเพื่อให้ผู้สอบเริ่มทำข้อสอบข้อที่ 1 ดังแสดงในภาพประกอบที่

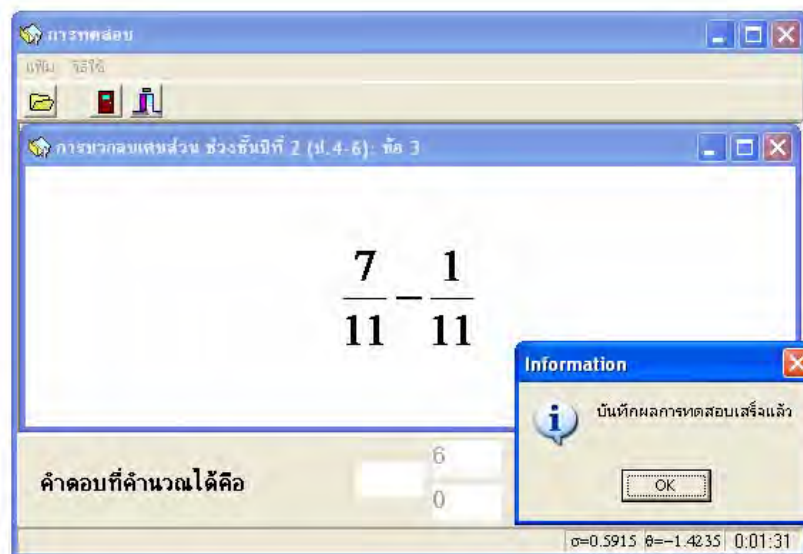


ภาพประกอบที่ 45 หน้าจอเพื่อให้ผู้สอบทำข้อสอบข้อที่ 1

จากภาพประกอบที่ 15 ผู้สอบดำเนินการทดสอบโดยพิมพ์คำตอบลงในช่องคำตอบที่คำนวณได้ และกดปุ่มข้อต่อไป โปรแกรมจะคัดเลือกข้อสอบให้ทำงานกระทั่งถึงเกณฑ์การหยุดสอบ เมื่อถึงเกณฑ์การหยุดสอบ โปรแกรมจะปรากฏหน้าจอเพื่อแสดงว่าการทดสอบเสร็จแล้ว ดังแสดงในภาพประกอบที่ 46



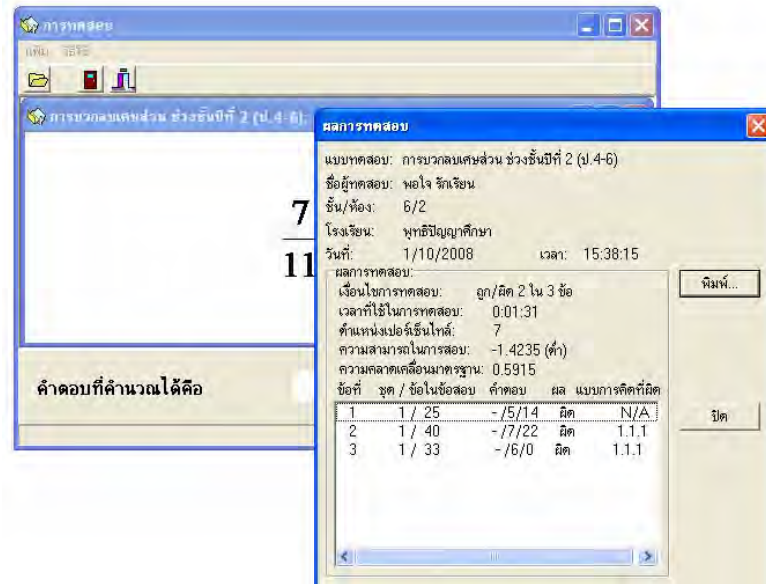
ภาพประกอบที่ 46 หน้าจอเพื่อแสดงว่าการทดสอบเสร็จแล้ว เมื่อคลิกที่ปุ่ม OK โปรแกรมจะปรากฏหน้าจอเพื่อแสดงว่าบันทึกผลการทดสอบเสร็จแล้ว ดังแสดงในภาพประกอบที่ 47



ภาพประกอบที่ 47 หน้าจอเพื่อแสดงว่าบันทึกผลการทดสอบเสร็จแล้ว



เมื่อคลิกที่ปุ่ม OK โปรแกรมจะปรากฏหน้าจอแสดงผลการทดสอบ ดังแสดงในภาพประกอบที่ 48 ผู้สอบสามารถสั่งพิมพ์ผลการทดสอบได้ โดยคลิกที่ปุ่มพิมพ์ หรือปิดหน้าจอการทดสอบโดยคลิกที่ปุ่มปิด ผลการทดสอบจะถูกบันทึกไว้ใน folder data



ภาพประกอบที่ 48 หน้าจอแสดงผลการทดสอบ

#### ภาคผนวก จ

- สภาพทั่วไปของกลุ่มตัวอย่างในการสำรวจและจัดเรียงลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน
- สภาพทั่วไปของกลุ่มตัวอย่างในการประเมินความถูกต้องของคุณลักษณะที่ใช้ในการบวกลบเศษส่วนและประเมินแผนภาพลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน
- สภาพทั่วไปของผู้ตอบแบบประเมินความครบถ้วนและความครอบคลุมของแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน

## สภาพทั่วไปของกลุ่มตัวอย่างในการสำรวจและจัดเรียงลำดับชั้นของคุณลักษณะ ที่ใช้ในการบวกลบเศษส่วน

กลุ่มตัวอย่างในการสำรวจและจัดเรียงลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ได้แก่ ครูต้นแบบ วิชาคณิตศาสตร์ ระดับประถมศึกษาของสำนักงานคณะกรรมการการศึกษาแห่งชาติ (สกศ.) ตั้งแต่ปี 2541 - 2544 จำนวน 57 คน ซึ่งตอบแบบสำรวจและแบบเปรียบเทียบลำดับชั้นของคุณลักษณะกลับและมีข้อมูลสมบูรณ์สามารถนำมาวิเคราะห์ได้ จำนวน 26 คน ส่วนใหญ่เป็นเพศหญิง (ร้อยละ 84.62) มีอายุ 41 ปี ขึ้นไปมากที่สุด (ร้อยละ 92.31) รองลงมา มีอายุระหว่าง 31 ถึง 35 ปี (ร้อยละ 7.69) ด้านวุฒิการศึกษาสูงสุด ส่วนใหญ่จบการศึกษาในระดับปริญญาโท (ร้อยละ 57.69) รองลงมา คือ ปริญญาตรี (ร้อยละ 42.31) โดยสำเร็จการศึกษาในสาขาวิชา/วิชาเอกคณิตศาสตร์มากที่สุด (ร้อยละ 61.54) รองลงมา ได้แก่ ประถมศึกษา (ร้อยละ 19.23) สังคมศึกษา (ร้อยละ 7.69) พลศึกษา บริหารการศึกษา และวิทยาศาสตร์ (ร้อยละ 3.85) ส่วนใหญ่มีประสบการณ์ในการสอนวิชาคณิตศาสตร์ ระหว่าง 21 ถึง 30 ปี (ร้อยละ 53.85) รองลงมา ต่ำกว่า 10 ปี 11-20 ปี (ร้อยละ 15.38) 31-40 ปี และ 41 ปีขึ้นไป (ร้อยละ 7.69) ดังแสดงในตารางที่ 1

ตารางที่ 1 สภาพทั่วไปของกลุ่มตัวอย่างในการสำรวจและจัดเรียงลำดับชั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน

สภาพทั่วไป	จำนวน	ร้อยละ
<b>เพศ</b>		
หญิง	22	84.62
ชาย	4	15.38
<b>อายุ</b>		
31 – 35 ปี	2	7.69
36 – 40 ปี	0	0.00
41 – 45 ปี	12	46.15
46 ปี ขึ้นไป	12	46.15
<b>วุฒิการศึกษาสูงสุด</b>		
ปริญญาตรี	11	42.31
ปริญญาโท	15	57.69
<b>สาขาวิชา/วิชาเอกที่สำเร็จการศึกษา</b>		
คณิตศาสตร์	16	61.54
ประถมศึกษา	5	19.23
พลศึกษา	1	3.85
บริหารการศึกษา	1	3.85
วิทยาศาสตร์	1	3.85
สังคมศึกษา	2	7.69
<b>ประสบการณ์ในการสอนวิชาคณิตศาสตร์ ระดับประถมศึกษา</b>		
ต่ำกว่า 10 ปี	4	15.38
11-20 ปี	4	15.38
21-30 ปี	14	53.85
31-40 ปี	2	7.69
41 ปีขึ้นไป	2	7.69

**สภาพทั่วไปของของกลุ่มตัวอย่างในการประเมินความถูกต้องของคุณลักษณะที่ใช้ในการบวกลบ  
เศษส่วนและประเมินแผนภาพลำดับขั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน**

กลุ่มตัวอย่างในการประเมินความถูกต้องของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน และกลุ่มตัวอย่างในการประเมินแผนภาพลำดับขั้นของคุณลักษณะที่ใช้ในการบวกลบเศษส่วน ประกอบด้วย ผู้เชี่ยวชาญด้านหลักสูตรคณิตศาสตร์หรือการเรียนการสอนคณิตศาสตร์ เรื่อง การบวกลบเศษส่วน จำนวน 3 คน เป็นนักวิชาการจากสำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน จำนวน 2 คน และครูผู้สอนคณิตศาสตร์ ระดับประถมศึกษา สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน ตำแหน่ง ครู คศ.3 จำนวน 1 คน นักวิชาการ จำแนกเป็น นักวิชาการสอบ ระดับ 8ว. ซึ่งดูแลงานด้านการวัดและประเมินผล ระดับการศึกษา ขั้นพื้นฐาน จำนวน 1 คน และนักวิชาการศึกษา ระดับ 7ว.ดูแลงานด้านหลักสูตรและการสอน คณิตศาสตร์ ระดับการศึกษาขั้นพื้นฐาน จำนวน 1 คน กลุ่มตัวอย่างทั้งหมดจบการศึกษาใน ระดับปริญญาโท โดยศึกษาในสาขาวิชา/วิชาเอกคณิตศาสตร์ในระดับปริญญาตรีทั้ง 3 คน ส่วนในระดับปริญญาโท ศึกษาในสาขาวิชา/วิชาเอกการสอนคณิตศาสตร์ จำนวน 1 คน การวัด และประเมินผลการศึกษา จำนวน 1 คน และบริหารการศึกษา จำนวน 1 คน ด้านประสบการณ์ ทำงานเกี่ยวกับการเรียนการสอนคณิตศาสตร์ พบว่า ทั้งหมดมีประสบการณ์ในการทำงาน ดังกล่าวมากกว่า 25 ปี

**สภาพทั่วไปของผู้ตอบแบบประเมินความครบถ้วนและความครอบคลุมของแบบการคิด  
ที่ผิดที่ใช้ในการบวกลบเศษส่วน**

ตารางที่ 2 สภาพทั่วไปของผู้ตอบแบบประเมินความครบถ้วนและความครอบคลุมของแบบการคิด  
ที่ผิดที่ใช้ในการบวกลบเศษส่วน

สภาพทั่วไป	จำนวน	ร้อยละ
<b>เพศ</b>		
หญิง	5	71.43
ชาย	2	28.57
<b>อายุ</b>		
41 – 45 ปี	2	28.57
46 ปี ขึ้นไป	5	71.43
<b>วุฒิการศึกษาสูงสุด</b>		
ปริญญาตรี	7	100
<b>สาขาวิชา/วิชาเอกที่สำเร็จการศึกษา</b>		
คณิตศาสตร์	1	14.28
ประถมศึกษา	3	42.86
พลศึกษา	1	14.28
บรรณารักษศาสตร์	1	14.28
วิทยาศาสตร์ทั่วไป	1	14.28
<b>ประสบการณ์ในการสอนวิชาคณิตศาสตร์ ระดับประถมศึกษา</b>		
ต่ำกว่า 20 ปี	4	57.14
20 ปีขึ้นไป	3	42.86

จากตารางที่ 2 แสดงให้เห็นว่าผู้ตอบแบบประเมินความครบถ้วนและความครอบคลุมของแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วนส่วนใหญ่เป็นเพศหญิง (ร้อยละ 71.43) มีอายุ 46 ปีขึ้นไปมากที่สุด (71.43) รองลงมาคืออายุระหว่าง 41-45 ปี (ร้อยละ 28.57) และไม่มีผู้ที่มีอายุต่ำกว่า 45 ปี ด้านวุฒิการศึกษาสูงสุด พบว่า ทั้งหมดจบการศึกษาในระดับปริญญาตรี โดยสำเร็จการศึกษาสาขา/วิชาเอกประถมศึกษามากที่สุด (ร้อยละ 42.86) รองลงมา ได้แก่ คณิตศาสตร์ พลศึกษา บรรณารักษศาสตร์และวิทยาศาสตร์ทั่วไป (ร้อยละ 14.28 ทั้ง 4 สาขา/วิชาเอก) ส่วนประสบการณ์ในการสอนวิชาคณิตศาสตร์ระดับประถมศึกษาพบว่า เป็นผู้ที่มีประสบการณ์ต่ำกว่า 20 ปี (ร้อยละ 57.14) มากกว่าเป็นผู้ที่มีประสบการณ์ 20 ปีขึ้นไป (ร้อยละ 42.86)



## ภาคผนวก ช

ผลการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน  
จากการทดลองใช้โปรแกรมคอมพิวเตอร์



## ผลการประเมินเชิงวิจักษ์สถานะความรู้และแบบการคิดที่ผิดที่ใช้ในการบวกลบเศษส่วน จากการทดลองใช้โปรแกรมคอมพิวเตอร์

### 1. ข้อมูลเบื้องต้นเกี่ยวกับการทดสอบจากการทดลองใช้โปรแกรมคอมพิวเตอร์

จากการทดลองใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะแก่นักเรียนระดับการศึกษาขั้นพื้นฐาน ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) จำนวน 54 คน พบว่า นักเรียนใช้เวลาในการสอบเฉลี่ย จำนวน 15.92 นาที โดยใช้เวลาดำสุด จำนวน 3 นาที และสูงสุด จำนวน 42 นาที ส่วนข้อสอบที่ทำทั้งหมด มีจำนวนเฉลี่ย 26.80 ข้อ โดยมีจำนวนข้อต่ำสุด 4 ข้อ และสูงสุด 69 ข้อ เมื่อพิจารณาตำแหน่งเปอร์เซ็นต์ไทล์ของผู้สอบ พบว่า มีค่าเฉลี่ยอยู่ตำแหน่งเปอร์เซ็นต์ไทล์ที่ 15.76 โดยมีตำแหน่งเปอร์เซ็นต์ไทล์ต่ำสุดที่ 0 และตำแหน่งเปอร์เซ็นต์ไทล์สูงสุดที่ 69 ส่วนระดับความสามารถของผู้สอบ พบว่า มีระดับความสามารถเฉลี่ย -1.32 โดยมีระดับความสามารถต่ำสุด -3.54 และระดับความสามารถสูงสุด 0.52 ด้านความคลาดเคลื่อนมาตรฐานในการประมาณค่า พบว่ามีความคลาดเคลื่อนมาตรฐานในการประมาณค่าเฉลี่ย 0.30 โดยมีความคลาดเคลื่อนมาตรฐานในการประมาณค่าต่ำสุด 0.17 และความคลาดเคลื่อนมาตรฐานในการประมาณค่าสูงสุด 0.63 ดังแสดงในตารางที่ 1

ตารางที่ 1 ข้อมูลเบื้องต้นเกี่ยวกับการทดสอบจากการทดลองใช้โปรแกรมคอมพิวเตอร์

ข้อมูลเบื้องต้นเกี่ยวกับการทดสอบ	จำนวนผู้สอบ	$\bar{x}$	S.D.	ต่ำสุด	สูงสุด
1. เวลาที่ใช้ในการทดสอบ	54	15.92	8.99	3	42
2. จำนวนข้อสอบที่ทำทั้งหมด	54	26.80	14.29	4	69
3. ตำแหน่งเปอร์เซ็นต์ไทล์ของผู้สอบ	54	15.76	17.97	0	69
4. ระดับความสามารถของผู้สอบ	54	-1.32	1.01	-3.54	0.52
5. ความคลาดเคลื่อนมาตรฐานในการประมาณค่า	54	0.30	0.09	0.17	0.63

### 2. ผลการประเมินเชิงวิจักษ์จากการทดลองใช้โปรแกรมคอมพิวเตอร์

#### 2.1 ผลการประเมินเชิงวิจักษ์สถานะความรู้

การประเมินเชิงวิจักษ์สถานะความรู้จากการทดลองใช้โปรแกรมคอมพิวเตอร์ในการทดสอบแบบปรับเหมาะ พบว่า นักเรียนส่วนใหญ่มีสถานะความรู้ ทักษะ หรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากันและจำนวนคละที่มีตัวส่วนเท่ากัน โดยมีแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 2 และ 4 มากที่สุด (ร้อยละ 29.6) รองลงมา ได้แก่ มีความรู้ ทักษะหรือ

กระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน เศษส่วนแท้ที่มีตัวส่วนไม่เท่ากันและจำนวนคละที่มีตัวส่วนเท่ากัน โดยมีแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 5 ร้อยละ 16.7 มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากันและเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน โดยแบบแผนการตอบข้อสอบที่คาดหวังที่ แบบแผนที่ 3 ร้อยละ 13.0 ไม่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วน โดยมีแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 1 ร้อยละ 9.3 และมีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน เศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน จำนวนคละที่มีตัวส่วนเท่ากันและจำนวนคละที่มีตัวส่วนไม่เท่ากัน โดยมีแบบแผนการตอบข้อสอบที่คาดหวังแบบแผนที่ 6 ร้อยละ 1.9 ดังแสดงในตารางที่ 2 ตารางที่ 2 ผลการประเมินเชิงวินิจฉัยสถานะความรู้จากการทดลองใช้โปรแกรมคอมพิวเตอร์

ที่	แบบแผนการตอบข้อสอบที่คาดหวัง	สถานะความรู้	จำนวน	ร้อยละ
1	0000	ไม่มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วน	5	9.3
2	1000	มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน	16	29.6
3	1100	มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากันและการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน	7	13.0
4	1010	มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากันและการบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน	16	29.3
5	1110	มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากันและการบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน	9	16.7
6	1111	มีความรู้ ทักษะหรือกระบวนการที่ใช้ในการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน การบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน การบวกลบจำนวนคละที่มีตัวส่วนเท่ากันและการบวกลบจำนวนคละที่มีตัวส่วนไม่เท่ากัน	1	1.9

## 2.2 ผลการประเมินเชิงวิจักษ์แบบการคิดที่ผิดที่ใช้

การประเมินเชิงวิจักษ์แบบการคิดที่ผิดที่ใช้ สำหรับคุณลักษณะที่ 1 การบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน พบว่า นักเรียนส่วนใหญ่ไม่ได้ใช้แบบการคิดที่ผิด (ร้อยละ 90.7) รองลงมา คือ วิจักษ์ไม่ได้ (ร้อยละ 5.6) และไม่เข้าใจวิธีการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน โดยใช้วิธีเศษ บวกลบเศษส่วนบวกลบส่วน (ร้อยละ 3.7) ส่วนคุณลักษณะที่ 2 การบวกลบเศษส่วนแท้ที่มีตัวส่วน ไม่เท่ากัน พบว่า นักเรียนส่วนใหญ่ไม่เข้าใจวิธีการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน โดยใช้วิธีเศษ บวกลบเศษ ตัวส่วนบวกลบตัวส่วน (ร้อยละ 38.8) รองลงมา ได้แก่ ไม่ได้ใช้แบบการคิดที่ผิด (ร้อยละ 31.5) วิจักษ์ไม่ได้และไม่ได้ทดสอบในลำดับขั้นนี้ (ร้อยละ 9.3) ไม่เข้าใจวิธีการทำตัวส่วนให้เท่ากัน โดย ใช้วิธีเอาตัวส่วนที่มากกว่าเป็นตัวส่วน เศษบวกลบเศษ (ร้อยละ 7.4) และใช้วิธีเอาตัวส่วนคูณกันเป็น ตัวส่วน เศษบวกลบเศษ หรือหา ค.ร.น. เป็นตัวส่วน เศษบวกลบเศษ (ร้อยละ 3.7)

เมื่อพิจารณาคคุณลักษณะที่ 3 การบวกลบจำนวนคละที่มีตัวส่วนเท่ากัน พบว่า นักเรียนส่วนใหญ่ไม่ได้ใช้แบบการคิดที่ผิด (ร้อยละ 46.3) รองลงมา ได้แก่ วิจักษ์ไม่ได้ (ร้อยละ 33.1) ไม่ได้ทดสอบในลำดับขั้นนี้ (ร้อยละ 9.3) ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน และ คำตอบไม่สมบูรณ์ โดยใช้วิธีจำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เศษบวกลบเศษ และไม่ ทำเศษเกินให้เป็นจำนวนคละ หรือ ไม่ทำให้เป็นเศษส่วนอย่างต่ำ (ร้อยละ 5.6) ไม่เข้าใจวิธีทำ จำนวนคละให้เป็นเศษเกิน โดยใช้วิธี ตัวส่วนบวกรจำนวนนับ บวกเศษ เป็นเศษ เศษบวกลบเศษ และ ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน และคำตอบไม่สมบูรณ์ โดยใช้วิธีตัวส่วนบวกรจำนวนนับ บวกเศษ เป็นเศษ เศษบวกลบเศษ และไม่ทำเศษเกินให้เป็นจำนวนคละ หรือ ไม่ทำให้เป็น เศษส่วนอย่างต่ำ (ร้อยละ 1.9) สำหรับคุณลักษณะที่ 4 การบวกลบจำนวนคละที่มีตัวส่วนไม่ เท่ากัน พบว่า นักเรียนส่วนใหญ่ไม่ได้ทดสอบในลำดับขั้นนี้ (ร้อยละ 83.3) รองลงมา ได้แก่ วิจักษ์ ไม่ได้ (ร้อยละ 11.1) ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน ไม่เข้าใจวิธีการทำตัวส่วนให้เท่ากัน และคำตอบไม่สมบูรณ์ โดยใช้วิธีตัวส่วนบวกรจำนวนนับ บวกเศษ เป็นเศษ เอาตัวส่วนคูณกันเป็นตัว ส่วน เศษบวกลบเศษ หรือหา ค.ร.น. เป็นตัวส่วน เศษบวกลบเศษ ไม่ทำเศษเกินให้เป็นจำนวน คละ หรือไม่ทำให้เป็นเศษส่วนอย่างต่ำ (ร้อยละ 3.7) และไม่ได้ใช้แบบการคิดที่ผิด (ร้อยละ 1.9) ดังแสดงในตารางที่ 3

ตารางที่ 3 ผลการประเมินเชิงวิจนิจัยแบบการคิดที่ผิดที่ใช้จากการทดลองใช้โปรแกรมคอมพิวเตอร์

คุณลักษณะที่	แบบการคิดผิดที่ใช้	จำนวน	ร้อยละ
1.การบวกลบ	1. ไม่ได้ใช้แบบการคิดที่ผิด	49	90.7
เศษส่วนแท้ที่มี	2. วิจนิจัยไม่ได้	3	5.6
ตัวส่วนเท่ากัน	3. ไม่เข้าใจวิธีการบวกลบเศษส่วนแท้ที่มีตัวส่วนเท่ากัน โดยใช้วิธี เศษบวกลบเศษส่วนบวกลบส่วน	2	3.7
2.การบวกลบ	1. ไม่ได้ใช้แบบการคิดที่ผิด	17	31.5
เศษส่วนแท้ที่มี	2. วิจนิจัยไม่ได้	5	9.3
ตัวส่วนไม่เท่ากัน	3. ไม่ได้ทดสอบในลำดับขั้นนี้	5	9.3
	4. ไม่เข้าใจวิธีการทำตัวส่วนให้เท่ากัน โดยใช้วิธี		
	4.1 เอาตัวส่วนที่มากกว่าเป็นตัวส่วน เศษบวกลบเศษ	4	7.4
	4.2 เอาตัวส่วนคูณกันเป็นตัวส่วน เศษบวกลบเศษ หรือ หา ค.ร.น เป็นตัวส่วน เศษบวกลบเศษ	2	3.7
	5. ไม่เข้าใจวิธีการบวกลบเศษส่วนแท้ที่มีตัวส่วนไม่เท่ากัน โดยใช้วิธี เศษบวกลบเศษ ตัวส่วนบวกลบตัวส่วน	21	38.8
3.การบวกลบ	1. ไม่ได้ใช้แบบการคิดที่ผิด	25	46.3
จำนวนคละที่มี	2. วิจนิจัยไม่ได้	19	33.1
ตัวส่วนเท่ากัน	3. ไม่ได้ทดสอบในลำดับขั้นนี้	5	9.3
	4. ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน โดยใช้วิธี	1	1.9
	ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เศษบวกลบเศษ		
	5. ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกินและคำตอบไม่สมบูรณ์ โดยใช้วิธี		
	5.1 จำนวนเต็มเป็นหลักสิบ เศษเป็นหลักหน่วย เศษบวกลบเศษ และ ไม่ทำเศษเกินให้เป็นจำนวนคละ หรือ ไม่ทำให้เป็นเศษส่วนอย่างต่ำ	3	5.6
	5.2 ตัวส่วนบวกจำนวนนับ บวกเศษ เป็นเศษ เศษบวกลบเศษ และไม่ ทำเศษเกินให้เป็นจำนวนคละ หรือ ไม่ทำให้เป็นเศษส่วนอย่างต่ำ	1	1.9
4.การบวกลบ	1. ไม่ได้ใช้แบบการคิดที่ผิด	1	1.9
จำนวนคละที่มี	2. วิจนิจัยไม่ได้	6	11.1
ตัวส่วนไม่เท่ากัน	3. ไม่ได้ทดสอบในลำดับขั้นนี้	45	83.3
	4. ไม่เข้าใจวิธีทำจำนวนคละให้เป็นเศษเกิน ไม่เข้าใจวิธีการทำตัว ส่วนให้เท่ากัน และคำตอบไม่สมบูรณ์ โดยใช้วิธีตัวส่วนบวกจำนวน นับ บวกเศษ เป็นเศษ เอาตัวส่วนคูณกันเป็นตัวส่วน เศษบวกลบ เศษ หรือหา ค.ร.น เป็นตัวส่วน เศษบวกลบเศษ ไม่ทำเศษเกินให้ เป็นจำนวนคละ หรือ ไม่ทำให้เป็นเศษส่วนอย่างต่ำ	2	3.7



## ประวัติผู้เขียนวิทยานิพนธ์

นางอมรรัตน์ สร้อยสังวาลย์ เกิดวันที่ 26 พฤศจิกายน พ.ศ. 2514 สำเร็จ การศึกษาครุศาสตรบัณฑิต เกียรตินิยมอันดับ 1 วิชาเอกการศึกษาปฐมวัย ในโครงการคุรุทายาท คณะครุศาสตร์ วิทยาลัยครุนครราชสีมา ปีการศึกษา 2535 ศึกษาศาสตรบัณฑิต สาขาการวัด และประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช ปีการศึกษา 2540 และครุศาสตรมหาบัณฑิต สาขาวิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2539 เข้าศึกษาต่อระดับดุษฎีบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อปีการศึกษา 2547 ปัจจุบันรับราชการตำแหน่งนักวิชาการศึกษา ระดับ 6 สำนักติดตามและประเมินผลการจัด การศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ

