

**การพัฒนาแบบทดสอบวัดความสามารถด้านการคิดวิเคราะห์
สำหรับนักเรียนชั้นมัธยมศึกษาปีที่สาม
A Development Analytical Ability Test
for Mattayomkuksa III Students**

สุทธจิวรรณ พิศัคดิ์โสภณ¹

บทคัดย่อ

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อ 1) สร้างและพัฒนาแบบทดสอบวัดความสามารถด้านการคิดวิเคราะห์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่สาม โดยวิเคราะห์หาคุณภาพข้อสอบรายข้อ ได้แก่ ค่าความยาก ค่าอำนาจจำแนก หาคุณภาพทั้งฉบับ ได้แก่ ความเชื่อมั่นของแบบทดสอบ และ 2) ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ กลุ่มตัวอย่างที่ใช้ในการวิจัย ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่สาม ตั้งแต่ปีการศึกษา 2552–2553 จำนวน 803 ถึง 935 คน ของโรงเรียนรัฐบาลสังกัดสำนักงานการศึกษาขั้นพื้นฐานทั้งในส่วนกลางและภูมิภาค เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบวัดความสามารถด้านการคิดวิเคราะห์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่สามจำนวนหกฉบับ ผลการวิจัยพบว่า 1) แบบทดสอบวัดความสามารถด้านการคิดวิเคราะห์ทั้งหกฉบับมีค่าพารามิเตอร์ของข้อสอบเมื่อใช้วิธีทฤษฎีการตอบข้อสอบ มีดังนี้ ค่าอำนาจจำแนก (a) มีค่าตั้งแต่ .116 ถึง 3.659 ค่าความยากมีค่าตั้งแต่ -2.626 ถึง 4.957 ค่าการเดามีค่าตั้งแต่ 0.045 ถึง 0.396 ค่าอินฟอร์เมชันสารสนเทศของแบบทดสอบฉบับ ข และ ง มีค่าสูงสุดที่ประมาณ 2 แบบทดสอบฉบับ ฉ มีค่าต่ำสุดประมาณ 0.50 สำหรับค่าความเชื่อมั่นเมื่อใช้สูตรของคูเดอร์ริชาร์ดสัน 20 มีค่าความเชื่อมั่นตั้งแต่ 0.6460 ถึง 0.8146 และ 2) จำนวนข้อสอบที่ทำหน้าที่ต่างกันคิดเป็นร้อยละ 13 และส่วนใหญ่เป็นข้อสอบที่โอกาสเพศชายจะตอบถูกมากกว่าเพศหญิง

คำสำคัญ : แบบวัดความสามารถด้านการคิดวิเคราะห์ การทำหน้าที่ต่างกันของข้อสอบ พารามิเตอร์ของข้อสอบ

¹ สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ

Abstract

The purposes of this study were to construct and develop an analytical thinking ability test for Mattayomsuksa III students, which was then analyzed by item difficulty, discrimination, reliability and determine differential item function. The sample consisted of 803 to 935 Mattayomsuksa III students and between the academic years 2009 and 2011 from government schools of the office of the basic education commission in the central and provincial area. The research instruments were six analytical thinking ability tests. The results are the following : 1) The item parameters of the tests using the item response theory method were the following: the discrimination parameter (a) ranged from .116 to 3.659, difficulty parameter (b) ranged from -2.626 to 4.957, the guessing parameter (c) ranged from 0.045 to 0.396, the information function values for Test B and D were the highest at approximately 2.0 and for the Test F the lowest at 0.50 of Test F. Furthermore, the reliability of the tests using the Kuder-Richarson 20 ranged from 0.6060 to 0.8146. 2) The differential item functioning in the tests was 13 percent of total items, with male students having more opportunity responses than female students.

Keywords : An analytical thinking ability test, Differential item function, Item parameters

บทนำ

การวัดผลและประเมินผลการศึกษาเป็นกระบวนการที่ใช้ตรวจสอบคุณภาพของการเรียนการสอนว่านักเรียนมีสัมฤทธิ์ผลตามที่จุดหมายของหลักสูตรได้กำหนดไว้หรือไม่ โดยเฉพาะเรื่องของการคิดวิเคราะห์เป็นเรื่องที่สำคัญและมีความจำเป็นอย่างยิ่งที่จะต้องพัฒนาผู้เรียนให้มีความสามารถด้านนี้ ดังที่หลักสูตรการศึกษาขั้นพื้นฐานได้กำหนดเกณฑ์มาตรฐานการเรียนรู้ของหลักสูตรฯ ไว้ว่า ผู้เรียนจะจบการศึกษาภาคบังคับและจบการศึกษาขั้นพื้นฐานได้ต้องผ่านการประเมินการอ่าน คิด วิเคราะห์ ให้ได้ตามเกณฑ์ที่สถานศึกษากำหนด (กรมวิชาการ. 2545) ซึ่งการที่ผู้เรียนจะมีความสามารถทางการอ่าน การคิด วิเคราะห์ได้นั้นครูผู้สอนต้องจัดกิจกรรมการเรียนการสอนเพื่อที่จะฝึกให้ผู้เรียนมีสมรรถนะตามที่หลักสูตรต้องการ และนับแต่ที่มีการประกาศใช้หลักสูตรการศึกษาขั้นพื้นฐาน สถานศึกษาต่างๆ ได้พยายามจัดการเรียนการสอนเพื่อที่จะพัฒนาให้ผู้เรียนเป็นไปตามที่หลักสูตรต้องการ แต่เครื่องมือที่ใช้วัดและประเมินความสามารถในด้านดังกล่าวที่ได้มาตรฐานยังมีไม่เพียงพอกับความต้องการของสถานศึกษา จึงได้ทำการสร้างและ

พัฒนาเครื่องมือวัดความสามารถด้านนี้ขึ้นมาโดยใช้แนวคิดของบลูม (ซวาล แพร์ตกุล. มปป.) ซึ่งยังเป็นที่ยอมรับใช้กันอย่างแพร่หลาย โดยจำแนกความสามารถด้านการคิดวิเคราะห์ออกเป็น การวิเคราะห์ความสำคัญ การวิเคราะห์ความสัมพันธ์และการวิเคราะห์หลักการ และวิเคราะห์หาคุณภาพของเครื่องมือโดยใช้ทฤษฎีการทดสอบแบบมาตรฐานเดิมและทฤษฎีการตอบข้อสอบ แต่ทฤษฎีต่างมีจุดเด่นที่ต่างกันนั่นคือ การวิเคราะห์โดยใช้ทฤษฎีการทดสอบแบบมาตรฐานเดิมผลการวิเคราะห์ให้ค่าสถิติรายตัวเลือกของข้อสอบ และการวิเคราะห์โดยทฤษฎีการตอบข้อสอบค่าสถิติที่ได้จะไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่าง นอกจากนี้ยังตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเพื่อไม่ให้เกิดความได้เปรียบเสียเปรียบระหว่างผู้สอบที่มีเพศต่างกัน

วัตถุประสงค์ของการวิจัย

1. เพื่อสร้างและพัฒนาแบบทดสอบวัดความสามารถด้านการอ่านคิดวิเคราะห์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่สาม โดยหาคุณภาพของข้อคำถาม ได้แก่ ค่าความยาก ค่าอำนาจจำแนก และความเชื่อมั่นของแบบทดสอบ
2. ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

วิธีการวิจัย

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ แบ่งเป็น

1. กลุ่มตัวอย่างที่ใช้ทดสอบครั้งที่หนึ่ง เป็นกลุ่มตัวอย่างที่ใช้เพื่อหาคุณภาพของข้อสอบรายข้อ ได้แก่ ค่าความยาก และค่าอำนาจจำแนก เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 ภาคเรียนที่ 1 ปีการศึกษา 2552 ของโรงเรียนรัฐบาลสังกัดสำนักงานการศึกษาขั้นพื้นฐานทั้งในส่วนกลางและภูมิภาค จำนวน 871 คน
2. กลุ่มตัวอย่างที่ใช้ทดสอบครั้งที่สอง เป็นกลุ่มตัวอย่างที่ใช้เพื่อหาคุณภาพของข้อสอบรายข้อ ได้แก่ ค่าความยาก และค่าอำนาจจำแนก และตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เป็นนักเรียนชั้นมัธยมศึกษาปีที่สาม ภาคเรียนที่ 2 ปีการศึกษา 2552 ของโรงเรียนรัฐบาลสังกัดสำนักงานการศึกษาขั้นพื้นฐานทั้งในส่วนกลางและภูมิภาค จำนวน 1,244 คน
3. กลุ่มตัวอย่างที่ใช้ทดสอบครั้งที่สาม เป็นกลุ่มตัวอย่างที่ใช้ทดสอบหาคุณภาพทั้งฉบับ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2553 ของโรงเรียนรัฐบาลสังกัดสำนักงานการศึกษาขั้นพื้นฐานทั้งในส่วนกลางและภูมิภาค การทดสอบครั้งนี้จัดแบบทดสอบเป็น 6 ฉบับ ให้กลุ่มตัวอย่างตอบแบบทดสอบคนละ 2 ฉบับๆ ละ 803 ถึง 935 คน

วิธีดำเนินการวิจัย

การสร้างเครื่องมือที่ใช้ในการวิจัย

ผู้วิจัยศึกษาเอกสารที่เกี่ยวข้องกับแนวคิด ทฤษฎี งานวิจัยที่เกี่ยวข้องกับการวัดความสามารถคิดวิเคราะห์ กำหนดกรอบนิยาม จัดประชุมปฏิบัติการสร้าง และให้ผู้ทรงคุณวุฒิพิจารณาความเที่ยงตรงของเครื่องมือวัด ได้ข้อสอบที่คัดเลือกและปรับปรุงจำนวน 252 ข้อ จัดแบบทดสอบเป็น 5 ฉบับ แล้วนำไปทดสอบกับกลุ่มตัวอย่างครั้งที่ 1 ก่อนดำเนินการทดสอบได้ประชุมชี้แจงการดำเนินการสอบให้แก่ผู้ช่วยเก็บข้อมูลเพื่อให้การดำเนินการสอบเป็นไปในแนวทางเดียวกัน นำผลการทดสอบครั้งที่หนึ่งมาวิเคราะห์หาคุณภาพของข้อสอบเป็นรายข้อ ปรับปรุงข้อคำถามและตัวเลือกบางข้อที่ไม่อยู่ในเกณฑ์ใช้ได้ แล้วคัดเลือกข้อสอบมาจำนวน 170 ข้อ จัดแบบทดสอบเป็น 3 ฉบับ เพื่อนำไปทดสอบกับกลุ่มตัวอย่างครั้งที่ 2 วิเคราะห์หาคุณภาพรายข้อ ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ แล้วนำข้อสอบมาจัดชุดแบบทดสอบออกเป็น 6 ฉบับ

ลักษณะของเครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบวัดความสามารถคิดวิเคราะห์สร้างตามแนวคิดของบลูม จำนวน 170 ข้อ จัดแบบทดสอบเป็น 6 ฉบับๆ ละ 40 ข้อ แต่ละฉบับประกอบด้วยลักษณะของเครื่องมือ 2 ลักษณะๆ ละ 20 ข้อ จากลักษณะต่อไปนี้

ลักษณะ A เป็นการกำหนดข้อความ เรียงราว บทสนทนา หรือสถานการณ์อย่างใดอย่างหนึ่งมาให้ แล้วให้วิเคราะห์หาคำตอบจากสิ่งที่กำหนดมาให้ มี 4 ชุดย่อย (A1 – A4) ชุดย่อยละ 10 ข้อ

ลักษณะ B โจทย์กำหนดสถานการณ์หรือเงื่อนไขเกี่ยวกับเรื่องราวใดเรื่องราวหนึ่ง แล้วให้วิเคราะห์โยงความสัมพันธ์กันระหว่างสิ่งที่กำหนดให้ มี 4 ชุดย่อย (B1 – B4) ชุดย่อยละ 10 ข้อ

ลักษณะ C โจทย์กำหนดคำหรือกลุ่มคำมาให้ ให้พิจารณาว่าคำหรือกลุ่มคำเหล่านั้นอยู่ร่วมกันโดยอาศัยหลักการใด หรือต่างกันโดยอาศัยหลักการใด แบ่งเป็น 2 ชุดย่อย (C1 – C2) ชุดย่อยละ 20 ข้อ

ลักษณะ D โจทย์กำหนดคำหรือกลุ่มคำมาให้ แล้วให้พิจารณาว่ากลุ่มคำเหล่านั้นมีความสัมพันธ์กันดังภาพเวนนไดอะแกรมใด มี 2 ชุดย่อย (D1 – D2) ชุดย่อยละ 20 ข้อ

การเขียนข้อสอบวัดความสามารถด้านคิดวิเคราะห์

การเขียนข้อสอบวัดความสามารถด้านคิดวิเคราะห์ที่ใช้ในการวิจัยครั้งนี้ สร้างตามแนวคิดของบลูม (ซวาล แพร์ตกุล. มปป.) ที่นิยามว่าการวิเคราะห์ เป็นความสามารถในการแยกแยะสิ่งต่างๆ หรือเรื่องราวต่างๆ ออกเป็นส่วนย่อยๆ ตามหลักการ และกฎเกณฑ์ที่กำหนดให้ เพื่อค้นหาความจริงต่างๆ ที่ซ่อนแฝงอยู่ภายในเรื่องราวนั้นแบ่งเป็น 3 ประเภท คือ

1. การวิเคราะห์ความสำคัญ เป็นการค้นหาคุณลักษณะที่เด่นชัดของเรื่องราวในแง่มุมต่างๆ ตามกฎเกณฑ์ที่กำหนดให้ เช่น ให้หาความเด่นด้อยของข้อความ ให้จำแนกชนิดประเภทและให้ค้นหาเลขนัยของคำพูดและการกระทำต่างๆ เป็นต้น

2. การวิเคราะห์ความสัมพันธ์ เป็นการค้นหาความเกี่ยวข้องระหว่างคุณลักษณะสำคัญใดๆ ของเรื่องราวและสิ่งต่างๆ การถามความสัมพันธ์ต้องเกี่ยวข้องกับของสองสิ่ง หรือสองเรื่องใดๆ เป็นอย่างน้อย ซึ่งเป็นสิ่งที่มีความหมายนัยสำคัญ หรือมีอิทธิพลเด่นชัดของเรื่องเหล่านั้นเสมอ

3. การวิเคราะห์หลักการ เป็นการค้นหาโครงสร้างและระบบของวัตถุสิ่งของเรื่องราว และการกระทำต่างๆ ว่าสิ่งเหล่านั้นรวมตัวกันเป็นรูป จนดำรงสภาพอยู่ได้ด้วยอะไร ยึดอะไรเป็นหลัก หรือมีสิ่งใดมาเป็นตัวเชื่อมโยง และการถามเป็นการถามให้วิเคราะห์หาโครงสร้าง เช่น ผู้ประพันธ์ใช้แนวการเขียนแบบใด หรือผู้แต่งยึดคติและทักษะใด การกระทำนี้ยึดหลักการใด

การเขียนข้อสอบลักษณะ A

การเขียนข้อสอบลักษณะนี้เป็นการกำหนดข้อความ เรื่องราว บทสนทนา หรือสถานการณ์อย่างใดอย่างหนึ่งมาให้ ซึ่งสถานการณ์ที่กำหนดมาให้นั้นต้องเป็นสถานการณ์ที่มีความหมาย มีนัย ให้แก่งคิด และมีความถูกต้องตามหลักวิชา แล้วให้ผู้ตอบค้นหาสิ่งที่มีความหมายนัยสำคัญของเรื่องราว หรือสถานการณ์ต่างๆ ที่กำหนดให้ สถานการณ์ที่กำหนดมาให้ต้องมีความสมบูรณ์พอที่จะให้ผู้ตอบสามารถลงสรุปได้ วินิจฉัยได้ ลักษณะคำถามที่จะให้วิเคราะห์อาจให้วิเคราะห์จากเรื่องราวทั้งหมด หรือวิเคราะห์จากตอนใดตอนหนึ่ง ซึ่งการถามอาจถามให้วิเคราะห์ชนิด ความสำคัญของเรื่องราว วิเคราะห์เลขนัย วัตถุประสงค์และผลลัพธ์ ถามเกี่ยวกับสาเหตุและผลของเรื่องราว

การเขียนข้อสอบลักษณะ B

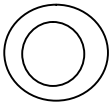
การเขียนคำถามลักษณะนี้เป็นลักษณะของการวิเคราะห์ความสัมพันธ์ โจทย์กำหนดสถานการณ์ต่างๆ ที่เป็นเงื่อนไขมาให้ แล้วหาความสัมพันธ์ระหว่างตัวแปรจากสถานการณ์ที่กำหนด การกำหนดสถานการณ์ต้องมีความรัดกุม เงื่อนไขต่างๆ ต้องมีความสัมพันธ์กัน

การเขียนข้อสอบลักษณะ C

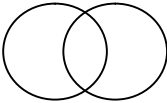
การเขียนข้อสอบประเภทนี้เป็นการวิเคราะห์หลักการตามแนวคิดของบลูม ซึ่งเป็นการให้ค้นหาโครงสร้างและระบบของวัตถุสิ่งของ เรื่องราว และการกระทำต่างๆ ว่าสามารถรวมกันจนเป็นสภาพเช่นนั้นได้ด้วยอะไร มีอะไรเป็นหลักยึด หรือมีอะไรเป็นตัวเชื่อมโยง ข้อสอบในแนวนี้อาจไม่ได้ยึดเนื้อหาวิชาเป็นหลัก การถามจะยกเอาคำมาให้ประมาณ 3 – 4 คำ แล้วให้พิจารณาว่าคำเหล่านั้นมีตัวร่วมกันคืออะไร ซึ่งผู้ตอบต้องวิเคราะห์คำทุกคำ บางครั้งอาจต้องใช้จินตนาการโยงความสัมพันธ์แต่ละสิ่งอย่างดี แล้วนำมาผสมกลมกลืนเป็นสิ่งใหม่ที่สามารถรับรู้และเข้าใจตรงกัน (ล้วน สายยศ และอังคณา สายยศ. 2541 : 132) การเขียนข้อสอบแบบนี้แยกเป็นแบบหาตัวร่วมและหาตัวต่าง

การเขียนข้อสอบลักษณะ D

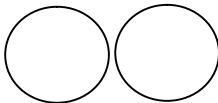
การเขียนข้อสอบลักษณะนี้ให้หาความสัมพันธ์ระหว่างสิ่งของหรือกลุ่มของคำที่กำหนดให้ ว่ามีความสัมพันธ์กันอย่างไร โดยโจทย์กำหนดคำมาให้สามคำแล้วให้วงกลมแทนสิ่งของหรือกลุ่มของคำเหล่านั้น และใช้ความสัมพันธ์หรือความเกี่ยวข้องกันของคำเหล่านั้นด้วยการซ้อนกันของวงกลม โดยใช้ภาพเวนนีไดอะแกรม ขนาดของวงกลมไม่ได้บ่งชี้ถึงขนาดของความสัมพันธ์ ลักษณะของการซ้อนกันของวงกลมมีความหมายดังนี้



แทน ทั้งหมดของสิ่งหนึ่งเป็นประเภทหนึ่งในอีกสิ่งหนึ่ง เช่น ไม่บรรทัด และอุปกรณ์การเรียน



แทน ทั้งสองสิ่งมีบางส่วนเป็นประเภทเดียวกันและมีบางส่วนเป็นคนละประเภทกัน เช่น ของเหลว และน้ำหวาน



แทน ทั้งสองสิ่งเป็นคนละประเภทกัน เช่น แว่นตา และไม่บรรทัด

สรุปผลการวิจัย

การสร้างและพัฒนาแบบทดสอบวัดความสามารถด้านการคิดวิเคราะห์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่สาม สรุปผลการวิจัยได้ดังนี้

ผลการวิเคราะห์หาคุณภาพของเครื่องมือวัดครั้งที่หนึ่ง

การวิเคราะห์หาคุณภาพรายข้อของแบบทดสอบวัดความสามารถด้านการคิดวิเคราะห์ครั้งที่หนึ่ง จำนวนห้าฉบับ ทั้งหมด 252 ข้อ ตัดออกจากการวิเคราะห์ 3 ข้อเนื่องจากพบข้อบกพร่องเกี่ยวกับโจทย์และตัวเลือกที่มีการพิมพ์ผิดพลาด มีค่าความยากตั้งแต่ .13 ถึง .88 ค่าอำนาจตั้งแต่ -.06 ถึง .69 รวมจำนวนข้อที่อยู่ในเกณฑ์ใช้ได้ 73 ข้อ คิดเป็นร้อยละ 70.68

ผลการวิเคราะห์หาคุณภาพของเครื่องมือวัดครั้งที่สอง

ผู้วิจัยได้นำข้อสอบมาปรับปรุงคำถามบางข้อ แล้วคัดเลือกมาจำนวน 170 ข้อ มาจัดเป็นแบบทดสอบฉบับใหม่จำนวนสามฉบับๆ ละ 51 52 และ 67 ข้อตามลำดับ นำไปทดสอบกับกลุ่มตัวอย่างเพื่อหาคุณภาพของเครื่องมือวัดครั้งที่สอง มีค่าความยากตั้งแต่ .11 ถึง .92 ค่าอำนาจจำแนกตั้งแต่ -.11 ถึง .67 รวมจำนวนข้อที่อยู่ในเกณฑ์ใช้ได้ 140 ข้อ คิดเป็นร้อยละ 82.35

นอกจากนี้ ผู้วิจัยได้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) โดยใช้ทฤษฎีการตอบข้อสอบ ได้ผลดังตาราง 1

ตาราง 1 การทำหน้าที่ต่างกันของแบบทดสอบวัดการคิดวิเคราะห์ จำนวนสามฉบับ

แบบทดสอบ	จำนวนข้อ	ข้อที่ DIF	จำนวนข้อ	DIF ตามเพศ
ฉบับที่หนึ่ง	51	6 ^{**} , 29 ^{**} 39 [*]	3	ชาย ชาย
ฉบับที่สอง	52	8 [*] , 11 [*] , 14 [*] , 15 [*] , 16 ^{**} 23 [*] , 48 [*]	7	ชาย หญิง
ฉบับที่สาม	67	1 ^{**} , 22 [*] , 37 ^{**} , 41 ^{**} , 43 ^{**} , 46 ^{**} , 47 ^{**} , 49 ^{**} , 51 ^{**} , 54 [*] , 55 ^{**} 7 [*]	12	ชาย ชาย หญิง
จำนวนข้อที่ DIF ทั้งหมด			22	

** มีนัยสำคัญทางสถิติที่ระดับ .01 , มีนัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 1 แบบทดสอบวัดการคิดวิเคราะห์ทั้งสามฉบับ จำนวน 170 ข้อ มีข้อที่ข้อสอบทำหน้าที่ต่างกัน 22 ข้อคิดเป็นร้อยละ 13 แบบทดสอบฉบับที่หนึ่ง มีข้อสอบจำนวน 3 ข้อที่ทำหน้าที่ต่างกันโดยที่โอกาสในการตอบถูกของนักเรียนชายมีมากกว่านักเรียนหญิง ได้แก่ข้อ 6 29 และ 39 แบบทดสอบฉบับที่สอง มีข้อสอบ 5 ข้อ ที่โอกาสในการตอบถูกของนักเรียนชายมีมากกว่านักเรียนหญิง ได้แก่ข้อ 8 11 14 15 และ 16 มี 2 ข้อที่โอกาสตอบถูกของนักเรียนหญิงมีมากกว่านักเรียนชาย ได้แก่ข้อ 23 และ 48 แบบทดสอบฉบับที่สาม มี 11 ข้อ ที่โอกาสในการตอบถูกของนักเรียนชายมีมากกว่านักเรียนหญิง ได้แก่ข้อ 1 22 37 41 43 46 47 49 51 54 และ 55 และมี 2 ข้อที่โอกาสตอบถูกของนักเรียนหญิงมีมากกว่านักเรียนชาย ได้แก่ข้อ 7 และ 9 จำนวนข้อสอบที่ทำหน้าที่ต่างกันคิดเป็นร้อยละ 23

ผู้วิจัยได้พิจารณาลักษณะของข้อคำถามและตัวเลือกของข้อที่ทำหน้าที่ต่างกันนำมาปรับปรุง แล้วจัดข้อสอบเป็นหกฉบับ แต่ละฉบับมีข้อสอบร่วมกัน 10 หรือ 20 ข้อ แล้วนำไปทดสอบกับกลุ่มตัวอย่างครั้งที่สาม หาค่าสถิติพื้นฐานและค่าความเชื่อมั่น

ผลการวิเคราะห์หาคุณภาพของเครื่องมือวัดครั้งที่สาม

การทดสอบครั้งที่สามเป็นการหาคุณภาพของแบบทดสอบทั้งหกฉบับ ได้แก่ ค่าสถิติพื้นฐาน ค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยาก อำนาจจำแนก ค่าการเดา และค่าฟังก์ชันอินฟอร์เมชันของแบบทดสอบ ตามทฤษฎีการตอบข้อสอบได้ผลดังตาราง 2

ตาราง 2 ค่าสถิติพื้นฐานของคะแนนแบบทดสอบวัดความสามารถในการคิดวิเคราะห์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

แบบทดสอบ	$\bar{X}$	S^2	C.V.	KR-20
ฉบับ ก	21.2913	37.0527	0.2857	0.7796
ฉบับ ข	22.3500	23.2866	0.2158	0.6460
ฉบับ ค	18.3600	28.7488	0.2919	0.7022
ฉบับ ง	20.0025	43.1740	0.3283	0.7882
ฉบับ จ	19.2413	46.0156	0.3523	0.8146
ฉบับ ฉ	19.9788	45.8356	0.3389	0.8139

จากตาราง 2 พบว่าคะแนนเฉลี่ยของแบบทดสอบวัดคิดวิเคราะห์ที่มีค่าเฉลี่ยตั้งแต่ร้อยละ 50 ขึ้นไป ได้แก่ ฉบับ ก ฉบับ ข และฉบับ ง แบบทดสอบฉบับ ค มีคะแนนเฉลี่ยต่ำสุด ค่าสัมประสิทธิ์การกระจายของแบบทดสอบฉบับ จ มีค่าสูงสุด และฉบับ ข มีค่าต่ำสุด สำหรับค่าความเชื่อมั่นของแบบทดสอบฉบับที่มีค่าความเชื่อมั่นสูงสุดคือ ฉบับ จ และฉบับที่มีค่าความเชื่อมั่นต่ำสุดคือ ฉบับ ข

ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบข้อสอบ

จากผลการทดสอบครั้งที่สามผู้วิจัยได้วิเคราะห์หาคุณภาพของแบบทดสอบโดยใช้ทฤษฎีการตอบข้อสอบ ก่อนหาค่าพารามิเตอร์ของข้อสอบแต่ละฉบับ ได้ทำการตรวจสอบความเป็นมิติเดียวของคะแนนจากแบบทดสอบแต่ละฉบับ โดยวิธีการวิเคราะห์องค์ประกอบโดยสะกัดองค์ประกอบหลัก (Principal Component : PC) แล้วหมุนแกนด้วยวิธีแวนแมกซ์ ปรากฏว่าคะแนนของแบบทดสอบแต่ละฉบับได้องค์ประกอบที่มีค่าไอเกนเกิน 1.0 และค่าไอเกนขององค์ประกอบที่ 1 มีค่าไอเกนสูงกว่าองค์ประกอบอื่นๆ เป็นอย่างมาก และค่าไอเกนขององค์ประกอบ 2 ถึงองค์ประกอบถัดมามีค่าใกล้เคียงกัน แสดงว่า แบบทดสอบการคิดวิเคราะห์แต่ละฉบับมีความเป็นมิติเดียว (พอประมาณ) (Lord. 1980 : 19) นั่นคือแบบทดสอบที่สร้างขึ้นนี้สามารถนำไปวิเคราะห์ตามทฤษฎีการตอบข้อสอบได้

ค่าพารามิเตอร์ของข้อสอบได้แก่ ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าการเดา (c) ได้ผลดังตาราง 3

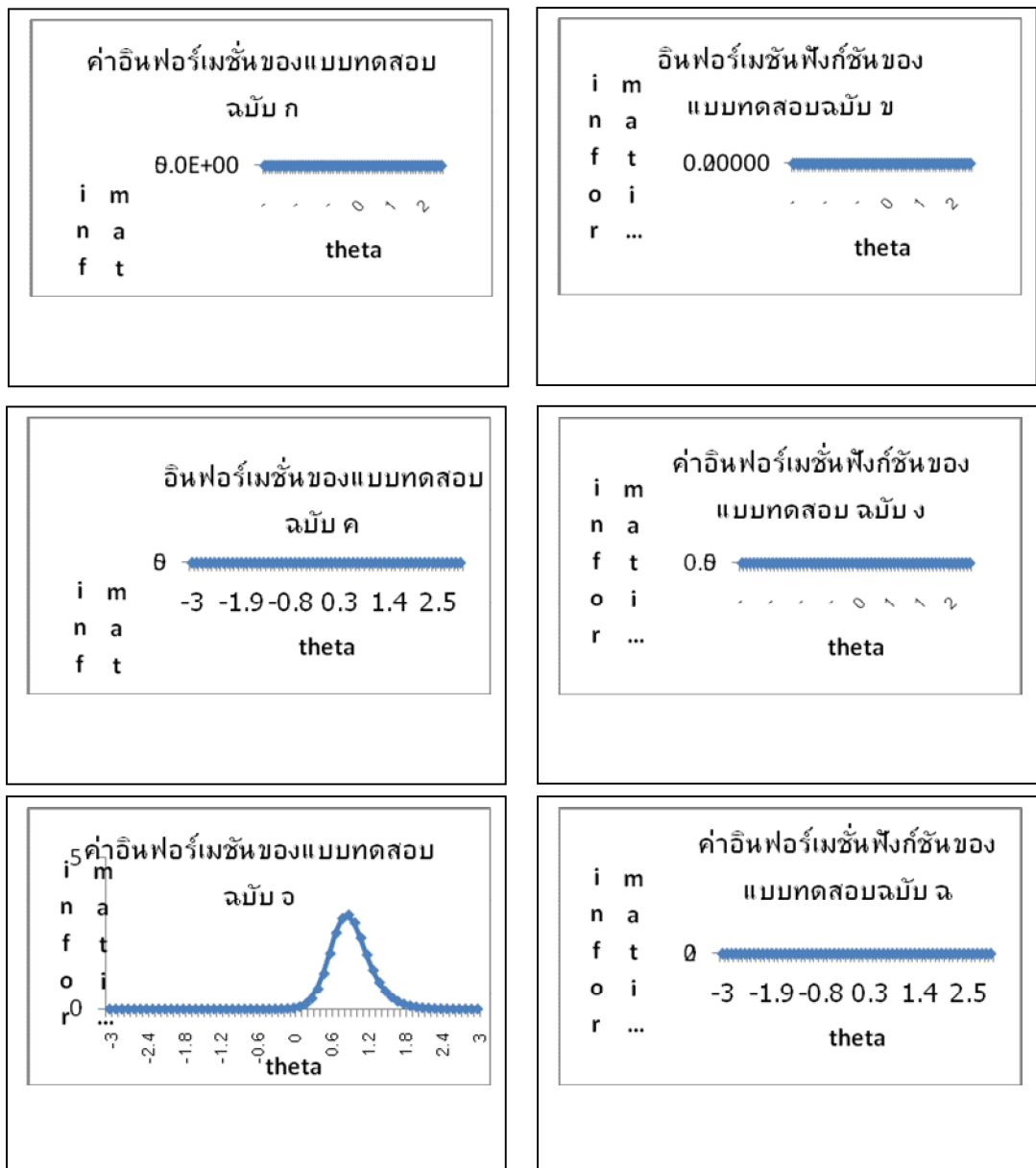
ตาราง 3 ค่าพิสัยของค่า a, b และ c ในแบบทดสอบวัดการคิดวิเคราะห์แต่ละฉบับ

แบบทดสอบ	พิสัยของค่า a	พิสัยของค่า b	พิสัยของค่า c
ฉบับ ก	.116 ถึง 3.659	-2.446 ถึง 4.400	.045 ถึง .315
ฉบับ ข	.160 ถึง 1.432	-2.626 ถึง 4.957	.131 ถึง .368
ฉบับ ค	.124 ถึง 2.096	-1.792 ถึง 4.802	.052 ถึง .246
ฉบับ ง	.236 ถึง 1.244	-2.34 ถึง 3.987	.086 ถึง .396
ฉบับ จ	.357 ถึง 2.667	-.996 ถึง 2.728	.113 ถึง .356
ฉบับ ฉ	.263 ถึง 1.824	-.784 ถึง 2.680	.067 ถึง .500

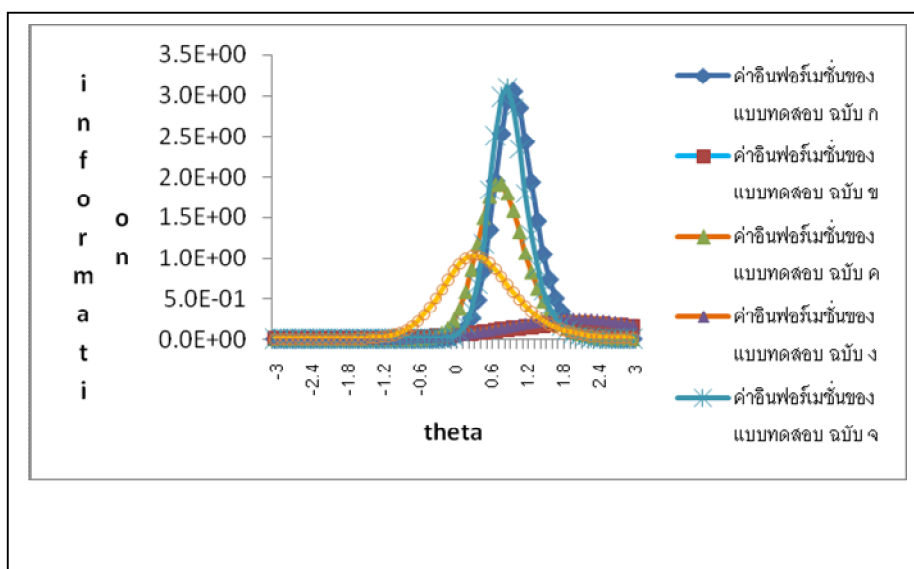
จากตาราง 3 ค่าอำนาจจำแนกต่ำสุดและสูงสุดอยู่ในแบบทดสอบฉบับ ก ข้อที่ง่ายที่สุดและข้อที่ยากที่สุดอยู่ในแบบทดสอบฉบับ ข ค่าการเดาที่น้อยที่สุดอยู่ในแบบทดสอบฉบับ ก และค่าการเดามากที่สุดอยู่ในแบบทดสอบฉบับ ฉ

ค่าฟังก์ชันสารสนเทศของแบบทดสอบ

ผู้วิจัยได้คำนวณค่าฟังก์ชันสารสนเทศของแบบทดสอบวัดการคิดวิเคราะห์แต่ละฉบับจากผลรวมของค่าฟังก์ชันสารสนเทศของข้อสอบทุกข้อในแบบทดสอบฉบับนั้นๆ ที่ระดับความสามารถตั้งแต่ -3 ถึง +3 ดังภาพประกอบ 1 แสดงให้เห็นว่าแบบทดสอบวัดด้านการคิดวิเคราะห์ฉบับ ก ค่าฟังก์ชันสารสนเทศสูงสุดอยู่ที่ระดับความสามารถ 1.0 แบบทดสอบฉบับ ข ค่าฟังก์ชันสารสนเทศสูงสุดอยู่ที่ระดับ 1.9 ถึง 2.3 แบบทดสอบฉบับ ค ค่าฟังก์ชันสารสนเทศสูงสุดอยู่ที่ระดับ 0.7 ถึง 0.8 แบบทดสอบฉบับ ง ค่าฟังก์ชันสารสนเทศสูงสุดอยู่ที่ระดับ 1.9 ถึง 2.1 แบบทดสอบฉบับ จ ค่าฟังก์ชันสารสนเทศสูงสุดอยู่ที่ระดับ 0.9 และแบบทดสอบฉบับ ฉ ค่าฟังก์ชันสารสนเทศสูงสุดอยู่ที่ระดับ 0.2 ถึง 0.5 และจากภาพประกอบ 2 เป็นโค้งฟังก์ชันสารสนเทศของแบบทดสอบวัดความสามารถคิดวิเคราะห์ทั้งหกฉบับ ปรากฏว่า ฉบับ ก และ จ ฉบับ ข และ ง มีรูปร่างของโค้งคล้ายกัน แต่สำหรับแบบทดสอบฉบับ ค และ ฉ พบว่า ค่าฟังก์ชันสารสนเทศของแบบทดสอบฉบับ ค เหมาะกับผู้สอบที่มีระดับความสามารถมากกว่าที่จะสอบในฉบับ ฉ



ภาพประกอบ 1 โค้งฟังก์ชันสารสนเทศของแบบทดสอบฉบับ ก ถึง ฉ



ภาพประกอบ 2 โค้งฟังก์ชันสารสนเทศของแบบทดสอบทั้งหกฉบับ

อภิปรายผล

จากผลการวิเคราะห์แบบทดสอบวัดการคิดวิเคราะห์ครั้งที่หนึ่ง จำนวน 252 ข้อ แต่ต้องตัดออกจากการวิเคราะห์ 3 ข้อเนื่องจากมีข้อบกพร่องเกี่ยวกับโจทย์และตัวเลือกที่มีการพิมพ์ผิดพลาด มีค่าความยากต่ำสุด .14 สูงสุด .86 ค่าอำนาจจำแนกต่ำสุด -.05 และสูงสุด .69 รวมจำนวนข้อที่อยู่ในเกณฑ์ใช้ได้คิดเป็นร้อยละ 70.68 ซึ่งจำนวนข้อที่ใช้ได้อยู่ในเกณฑ์สูงแม้จะเป็นการทดสอบครั้งที่หนึ่งก็ตาม ทั้งนี้เพราะแบบทดสอบเหล่านี้ได้ผ่านการพิจารณาจากผู้ทรงคุณวุฒิที่มีความรู้ความเชี่ยวชาญเกี่ยวกับการสร้างข้อคำถามวัดการคิดวิเคราะห์เป็นอย่างดี อย่างไรก็ตามยังมีข้อสอบบางข้อที่ค่อนข้างง่าย และค่าอำนาจจำแนกติดลบ อาจเป็นเพราะโจทย์บางข้อมีค่าบางค่าที่นักเรียนไม่รู้จัก จึงอาจเดาคำตอบ ดังนั้นผู้วิจัยจึงได้ปรับปรุงข้อคำถามบางข้อและนำไปทดสอบครั้งที่สอง ผลการวิเคราะห์ได้ค่าความยากต่ำสุด .11 ค่าความยากสูงสุด .92 ค่าอำนาจจำแนกต่ำสุด -.11 ค่าอำนาจจำแนกสูงสุด .67 เมื่อคิดเป็นจำนวนข้อที่ผ่านเกณฑ์คิดเป็นร้อยละ 82.35 ซึ่งจำนวนข้อที่ใช้ได้คิดเป็นร้อยละมากกว่าการทดสอบครั้งแรก เพราะข้อสอบได้ผ่านการทดสอบและปรับปรุงมาแล้วเป็นอย่างดี แต่จะอย่างไรก็ตามก็พบว่าเมื่อทดสอบกับกลุ่มตัวอย่างครั้งที่สองข้อสอบจะยากกว่าเดิม และบางข้อง่ายกว่าเดิม ทั้งนี้อาจขึ้นอยู่กับตัวอย่างที่ใช้ทดสอบตามทฤษฎีการวิเคราะห์ข้อสอบแบบมาตรฐานเดิมผลการวิเคราะห์จะแปรเปลี่ยนไปตามกลุ่มตัวอย่างที่สอบ (Hamandbleton, Swaminthan and Rogers.1991 : 3) นอกจากการหาค่าความยากและอำนาจจำแนกแล้ว ผู้วิจัยได้นำผลการตอบไปตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ พบว่า ในจำนวนแบบทดสอบวัดการคิดวิเคราะห์ทั้งสามฉบับ จำนวน 170 ข้อ มีข้อสอบที่ทำหน้าที่ต่างกัน 22 ข้อคิดเป็นร้อยละ 13 ซึ่งมีเป็นจำนวนไม่มากนัก เมื่อพิจารณาในประเด็น

คำถามพบว่าการทำหน้าที่ต่างกันที่โอกาสนักเรียนชายตอบถูกมีมากกว่านักเรียนหญิง ส่วนใหญ่จะเป็นคำถามในแบบทดสอบที่ให้คิดวิเคราะห์หาตัวเหมือนตัวต่างหรือการวิเคราะห์หลักการตามแนวคิดของบลูม (ซวาล แพร์ตกุล. มปป.) และเป็นเรื่องราวที่ค่อนข้างให้โอกาสนักเรียนชายตอบถูกมากกว่านักเรียนหญิง เช่น คำถามเกี่ยวกับหน้าของลูกเต๋า อย่างไรก็ตามผลการวิจัยนี้สอดคล้องกับงานวิจัยของ Katherine และ Meichu (รักชนก ยี่สุนศรี. 2544 : 43 อ้างอิงจาก Katherine and Meichu. 1996) ที่ได้สร้างข้อสอบพบว่าในภาพรวมข้อสอบด้านการประยุกต์ซึ่งต้องอาศัยการคิดวิเคราะห์จะทำหน้าที่ต่างกันโดยเพศชายจะมีโอกาสตอบถูกได้มากกว่าและเป็นข้อสอบที่ยากสำหรับเพศหญิง แม้ว่าจะมีบ้างที่แบบทดสอบการคิดวิเคราะห์นี้มีบางข้อที่ทำหน้าที่ต่างกัน แต่ก็ไม่กระทบต่อคุณภาพของแบบทดสอบ เพราะงานวิจัยของเกษร ห่วงจิตร (2539 :บทคัดย่อ) ได้ตรวจสอบความเที่ยงและความตรงของแบบทดสอบกรณีก่อนและหลังตัดข้อสอบที่ทำหน้าที่ต่างกันออกจากแบบทดสอบจากพบว่าไม่แตกต่างกัน และงานวิจัยของ Rozowski และ Reith (รักชนก ยี่สุนศรี. 2544 : 46 อ้างอิงจาก Rozowski and Reith. 1999) ได้ทำการตรวจสอบคุณภาพในการวัดและความตรงของแบบทดสอบที่ประกอบด้วยข้อสอบที่ DIF และไม่ DIF โดยใช้วิธีแมนเทิล-แฮนเซล และแบ่งกลุ่มตามตัวแปรและเชื้อชาติพบว่าแบบทดสอบที่ประกอบด้วยข้อสอบที่ทำหน้าที่ต่างกันนั้น ไม่ได้ทำให้แบบทดสอบมีคุณภาพในการวัดหรือมีความตรงต่ำลงมากนัก ดังนั้นแบบทดสอบวัดการคิดวิเคราะห์ที่สร้างขึ้นทั้งหกฉบับนี้จึงสามารถนำไปใช้วัดได้แม้จะมีบางข้อที่มีการทำหน้าที่ต่างกันก็ตามซึ่งมีเป็นจำนวนน้อย

เมื่อจัดแบบทดสอบการคิดวิเคราะห์ออกเป็นหกฉบับ พบว่าแบบทดสอบฉบับ ก ฉบับ ข และฉบับ ง ค่าเฉลี่ยมีค่าร้อยละตั้งแต่ 50 ขึ้นไป นอกนั้นต่ำกว่าร้อยละ 50 โดยเฉพาะฉบับ ค มีค่าเฉลี่ยต่ำสุดที่เป็นดังนี้อาจเป็นเพราะรูปแบบของข้อคำถามในฉบับ ก ฉบับ ข เป็นรูปแบบเดียวกัน และฉบับ ง ก็มีบางตอนที่มียูรูปแบบเดียวกับฉบับ ก และ ข ซึ่งเป็นรูปแบบที่ผู้สอบค่อนข้างคุ้นเคยและมักพบในการเรียนการสอนการสอบ ส่วนแบบทดสอบฉบับ ค ค่าเฉลี่ยมีค่าไม่ถึงครึ่งหนึ่งของคะแนนเต็มและมีค่าต่ำสุด ที่เป็นดังนี้อาจเป็นเพราะรูปแบบที่ใช้ในแบบทดสอบฉบับ ค ส่วนหนึ่งเป็นเรื่องของการคิดวิเคราะห์ในหลักการ ซึ่งต้องใช้สมรรถภาพสมองขั้นสูงในการตอบ สำหรับค่าสัมประสิทธิ์การกระจายพบว่า แบบทดสอบฉบับ จ มีค่ามากที่สุด ในขณะที่ฉบับ ข มีค่าต่ำสุด ที่เป็นดังนี้เนื่องจากที่กล่าวไปแล้วว่ารูปแบบในแบบทดสอบฉบับ ข เป็นรูปแบบที่ผู้สอบมักคุ้นเคยแต่ในฉบับ จ จะค่อนข้างยากคะแนนเฉลี่ยไม่ถึงครึ่งหนึ่งของคะแนนเต็มทำให้คะแนนที่ได้แตกต่างกันมากกว่าฉบับอื่นและถ้าแบบทดสอบที่ยากเกินไปหรือง่ายเกินไปการกระจายของคะแนนจะแคบเมื่อพิจารณาค่าความเชื่อมั่นจากการใช้สูตรของคูเดอร์ริชาร์ดสันที่ 20 พบว่า ค่าความเชื่อมั่นของแบบทดสอบฉบับ จ มีค่ามากที่สุดและฉบับ ข มีค่าต่ำสุด ซึ่งเป็นไปตามทฤษฎีที่ว่าถ้าคะแนนของแบบทดสอบฉบับใดก็ตามมีความแตกต่างกันหรือการกระจายของคะแนนมากย่อมทำให้ความเชื่อมั่นที่ได้มีค่าสูง หรือแบบทดสอบที่มีความคลาดเคลื่อนต่ำความเชื่อมั่นของแบบทดสอบก็จะสูงตามไปด้วย (Allen and Yen. 1971 : 73-75)

สำหรับการวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบข้อสอบ พบว่าค่าพารามิเตอร์ของข้อสอบแต่ละฉบับส่วนใหญ่มีค่าพารามิเตอร์ของข้อสอบอยู่ในช่วงพิสัยที่ใช้ในทางปฏิบัติ นั่นคือ ค่า a มีค่าตั้งแต่ 0 ถึง 2 ไม่มีค่าใดติดลบ ค่าความยาก b มีบางข้อของบางฉบับที่มีค่าอยู่นอกพิสัย -3 ถึง +3 นั่นคือบางข้อมีค่ามากกว่า +3 แสดงว่าข้อสอบนั้นค่อนข้างยาก (ศิริชัย กาญจนวาสี. 2545 : 47-48) ที่เป็นดังนี้อาจเป็นเพราะรูปแบบของข้อคำถามในแต่ละฉบับไม่เหมือนกัน ฉบับที่มีข้อคำถามบางข้อยาก ส่วนใหญ่เป็นข้อสอบที่ทำให้คิดวิเคราะห์ความสัมพันธ์ วิเคราะห์หลักการ สำหรับค่าการเดาไม่มากข้อที่มีค่าการเดามากกว่า .30 ซึ่งอาจเป็นเพราะนักเรียนเหนื่อยล้าต่อการสอบ ทำให้ไม่ตั้งใจและเกิดการเดาบ้างในบางข้อ

สำหรับโค้งฟังก์ชันอินฟอร์เมชันของแบบทดสอบวัดความสามารถคิดวิเคราะห์ทั้ง 6 ฉบับปรากฏว่า ฉบับ ก และ จ ฉบับ ข และ ง มีรูปร่างของโค้งคล้ายกัน นั้นแสดงว่าแบบทดสอบดังกล่าวเหมาะกับผู้สอบที่มีระดับความสามารถเหมือนกัน แต่สำหรับแบบทดสอบฉบับ ค และ ฉ พบว่า ค่าฟังก์ชันสารสนเทศของแบบทดสอบฉบับ ค เหมาะกับผู้สอบที่มีระดับความสามารถมากกว่าที่จะสอบในฉบับ ฉ

ข้อเสนอแนะ

ข้อเสนอแนะเพื่อการวิจัย

1. แบบทดสอบวัดความสามารถด้านการคิดวิเคราะห์ที่ได้จากการวิจัยครั้งนี้มีทั้งหมดหกฉบับ แต่ละฉบับมีข้อสอบที่ร่วมกันอยู่ 10 หรือ 20 ข้อ อาจจับคู่แบบทดสอบแล้วนำไปทดสอบเพื่อการเทียบคะแนนโดยวิธีการเทียบต่างๆ

2. ควรศึกษารูปแบบการสร้างแบบทดสอบวัดการคิดวิเคราะห์แบบต่างๆ ด้วยวิธีการที่ต่างกัน แล้วตรวจสอบความเที่ยงตรงของแบบทดสอบ

3. ควรศึกษาวิจัยโดยใช้โมเดลลิสเรลเพื่อตรวจสอบความเป็นเอกพันธ์ของแบบทดสอบความเที่ยงตรงของแบบทดสอบ

4. ควรศึกษาคะแนนเกณฑ์ปกติของแบบทดสอบแต่ละฉบับเพื่อใช้ประเมินความสามารถด้านการคิดวิเคราะห์ของผู้สอบ

ข้อเสนอแนะเพื่อนำไปใช้

1. แบบทดสอบวัดความสามารถด้านคิดวิเคราะห์ที่ได้จากการวิจัยครั้งนี้มีหกฉบับสร้างขึ้นโดยไม่ได้ยึดเนื้อหาวิชาเป็นหลัก และสร้างตามแนวความถนัดทางการเรียน ดังนั้นในการนำไปใช้ควรให้สอดคล้องกับวัตถุประสงค์ของผู้นำไปใช้

2. แบบทดสอบทั้งหกฉบับนี้แต่ละฉบับมีรูปแบบของข้อสอบสองรูปแบบ ฉบับ ก - ฉ ฉบับ ข - จ ฉบับ ค - ง เป็นฉบับที่มีรูปแบบต่างกันและไม่มีข้อสอบร่วมกัน จึงสามารถนำไปทดสอบกับนักเรียนพร้อมกันทั้งสองฉบับได้

หนังสืออ้างอิง

- กรมวิชาการ กระทรวงศึกษาธิการ. หลักสูตรสถานศึกษาขั้นพื้นฐานพุทธศักราช 2544.
กรุงเทพฯ : โรงพิมพ์องค์การรับส่งสินค้าและพัสดุภัณฑ์ (ร.ส.พ.), 2545.
- เกษร ห่วงจิตร. การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบสำหรับแบบสอบคัดเลือกระดับ
บัณฑิตศึกษาวิชาภาษาไทยและภาษาอังกฤษด้วยวิธีแมนเทิล-เฮนส์เซล. จุฬาลงกรณ์
มหาวิทยาลัย, 2539.
- ขวาล แพรรัตน์กุล (มปป.) เทคนิคการเขียนข้อทดสอบ. มปท. กรุงเทพฯ.
- รักชนก ยี่สุนศรี. การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและแบบสอบด้วย
กระบวนการ ดีเอฟ ไอ ที่ สำหรับแบบสอบคัดเลือกบุคคลเข้าศึกษาในสถาบันอุดมศึกษา
วิชาภาษาอังกฤษและวิชาคณิตศาสตร์. วิทยานิพนธ์ (ค.ม.). กรุงเทพฯ : จุฬาลงกรณ์
มหาวิทยาลัย, 2544.
- ล้วน สายยศ และ อังคณา สายยศ. เทคนิคการสร้างและสอบข้อสอบความถนัดทางการเรียน.
กรุงเทพฯ : สุวีริยาสาส์น, 2541.
- ศิริชัย กาญจนวาสี. ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theories). พิมพ์ครั้งที่ 2.
กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2545.
- Allen, Mary J. & Yen, Wendy M. Introduction to Measurement Theory. Belmont
California : Brook/Code Publishing, 1979.
- Bloom, Benjamin S. ; et all (1956). Taxonomy of Educational Objectives Book :
Cognitive Domain. Longman Group Limited.
- Lord, F.M. Application of Item Response Theory to Practical Testing Problems.
N.J. : Hillsdale, Erlbaum, 1980.
- Good, Carter.V. Dictionary of Education. 3rd ed. New York : McGraw-Hill, Inc,
1973.
- Hambleton, Ronald K. & Swaminathan, Hariharan and Rogers H.Jane.
Fundamentals of Item Response Theory. Newbury Park : London :
Sage, 1991.