การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

วิทยานิพนธ์

ของ

ศราวุทธ โพธิ์สิงห์

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต สาขาวิชาวิจัยและประเมินผลการศึกษา

มกราคม 2565

ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์
ระดับชั้นมัธยมศึกษาปีที่ 3

วิทยานิพนธ์
ของ
ศราวุทธ โพธิ์สิงห์

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต สาขาวิชาวิจัยและประเมินผลการศึกษา
มกราคม 2565
ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

Detecting Differential Item Functioning of Science Literacy Mathayomsuksa 3

Sarawut Posing

A Thesis Submitted in Partial Fulfillment of Requirements

for Master of Education (Educational Research and Evaluation)

January 2022

Copyright of Mahasarakham University

คณะกรรมการสอบวิทยานิพนธ์ ได้พิจารณาวิทยานิพนธ์ของนายศราวุทธ โพธิ์สิงห์ แล้วเห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาวิจัยและประเมินผลการศึกษา ของมหาวิทยาลัยมหาสารคาม

คณะกรรมการสอบวิทยานิพนธ์

....ประธานกรรมการ
(ผศ. ดร. ญาณภัทร สีหะมงคล)

....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผศ. ดร. วราพร เอราวรรณ์)

....กรรมการ
(ผศ. ดร. อรนุช วราอัศวปติ ศรีสะอาด)

....กรรมการผู้ทรงคุณวุฒิภายนอก
(ผศ. ดร. ศักดิ์สิทธิ์ ฤทธิลัน)

มหาวิทยาลัยอนุมัติให้รับวิทยานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา การศึกษามหาบัณฑิต สาขาวิชาวิจัยและประเมินผลการศึกษา ของมหาวิทยาลัยมหาสารคาม

....	....
(รศ. ดร. พชรวิทย์ จันทร์ศิริสิร)	(รศ. ดร. กริสน์ ชัยมูล)
คณบดีคณะศึกษาศาสตร์	คณบดีบัณฑิตวิทยาลัย

ชื่อเรื่อง การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3
ผู้วิจัย ศราวุทธ โพธิ์สิงห์
อาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร. วราพร เอราวรรณ์
ปริญญา การศึกษามหาบัณฑิต **สาขาวิชา** วิจัยและประเมินผลการศึกษา
มหาวิทยาลัย มหาวิทยาลัยมหาสารคาม **ปีที่พิมพ์** 2565

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อ 1) เพื่อสร้างและหาคุณภาพแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 2) เพื่อวิเคราะห์การทำหน้าที่ต่างกัน (DIF) ของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC โดยใช้โปรแกรมประยุกต์ Mplus จำแนกตามเพศ กลุ่มตัวอย่างในงานวิจัยครั้งนี้ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2561 ในจังหวัดชัยภูมิ จำนวน 689 คน จากการสุ่มแบบหลายขั้นตอน (Multi–Stage Random Sampling) เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบการรู้เรื่องวิทยาศาสตร์ (Scientific Literacy Test) จำนวน 54 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 18 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ

ผลการวิจัยพบว่า

1. การสร้างและหาคุณภาพแบบทดสอบการรู้เรื่องวิทยาศาสตร์ จำนวน 54 ข้อ มีค่าดัชนีความสอดคล้องตั้งแต่ 0.60 ถึง 1 เมื่อพิจารณาคุณภาพของข้อสอบผ่านเกณฑ์การคัดเลือก จำนวน 45 ข้อ

1.1 คุณภาพข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ตามแนวทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) พบว่า มีค่าความยาก (p) ตั้งแต่ 0.22 ถึง 0.79 ค่าอำนาจจำแนก (r) ตั้งแต่ 0.21 ถึง 0.77 และค่าความเชื่อมั่นทั้งฉบับ มีค่าเท่ากับ 0.92

1.2 คุณภาพข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ตามแนวทฤษฎีการตอบสนองของข้อสอบ (Item Response Theory: IRT) โดยใช้โมเดลโลจีสติกแบบ 2 พารามิเตอร์ (Two–Parameter Model) พบว่า ค่าความยากของข้อสอบ (b-parameter) มีค่าตั้งแต่ -0.537 ถึง

4.000 และค่าอำนาจจำแนกของข้อสอบ (a-parameter) มีค่าตั้งแต่ 0.313 ถึง 1.382

2. การวิเคราะห์การทำหน้าที่ต่างกัน (DIF) ของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 45 ข้อ ด้วยวิธี MIMIC โดยใช้โปรแกรมประยุกต์ Mplus จำแนกตามเพศ พบว่า เพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 18 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 7 ข้อ

คำสำคัญ : การตรวจสอบการทำหน้าที่ต่างกัน, การรู้เรื่องวิทยาศาสตร์, วิธี MIMIC

TITLE Detecting Differential Item Functioning of Science Literacy Mathayomsuksa 3
AUTHOR Sarawut Posing
ADVISORS Assistant Professor Waraporn Erawan , Ph.D.
DEGREE Master of Education MAJOR Educational Research and Evaluation
UNIVERSITY Mahasarakham University YEAR 2022

ABSTRACT

The aims of this research are 1) to construct and develop a Science Literacy test Mathayomsuksa 3. 2) to analyze the Differential Item Functioning (DIF) by using Multiple Indicators and Multiple Causes (MIMIC) of the Science Literacy Exam through Mplus Program classified by the gender. The sample group was 689 students of Mathayomsuksa 3 in academic year 2018 in Chaiyaphum province through Multistage Random Sampling. The research instrument consisted of 54 items from Science Literacy Test. There were 3 parts of knowledge content as follows: Physical Systems; 18 items, Living Systems; 18 items and Earth and Space Systems; 18 items.

The research findings showed that:

1. The creation and quality of a 54-item science literacy test, with a consistency index ranging from 0.60 to 1, considering the quality of the exams through the selection criteria for 45 items.

1.1 The exam quality of the science literacy test according to the Classical Test Theory (CTT) was found that the difficulty (p) ranged from 0.22 to 0.79. The discriminant (r) ranged from 0.21 to 0.77 and the overall reliability was 0.92.

1.2 The test quality of the Science Literacy Test according to the Item

Response Theory (IRT) by using a two-parameter logistic model (Two-Parameter Model) found that the difficulty of the test (b-parameter) ranged from -0.537 to 4.000 and the power of discrimination of the exam (a-parameter) ranged from 0.313 to 1.382.

2. Differential Item Functioning (DIF) of the science literacy Exam in Mathayomsuksa 3 for 45 items through MIMIC method, classified by gender, using the Mplus application. It was found that males have a higher possibility of answering correctly than females for 18 items and females have a higher possibility of answering correctly than males for 7 items.

Keyword : Differential Item Functioning (DIF), Science Literacy Test, Multiple Indicators and Multiple Causes (MIMIC)

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จสมบูรณ์ได้ด้วยความกรุณาและความช่วยเหลืออย่างสูงยิ่งจาก ผู้ช่วยศาสตราจารย์ ดร.วราพร เอราวรรณ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก ผู้ช่วยศาสตราจารย์ ดร.ญาณภัทร สีหะมงคล ประธานกรรมการสอบ ผู้ช่วยศาสตราจารย์ ดร.อรนุช วราอัศวปติ ศรีสะอาด กรรมการสอบ และผู้ช่วยศาสตราจารย์ ดร.ศักดิ์สิทธิ์ ฤทธิลัน ผู้ทรงคุณวุฒิภายนอก ที่ได้ให้คำแนะนำ ให้ความคิดเห็น ช่วยเหลือ และตรวจแก้ไขข้อบกพร่องด้วยความเอาใจใส่ ผู้วิจัยรู้สึกซาบซึ้งในพระคุณเป็นอย่างยิ่ง และขอขอบพระคุณเป็นอย่างสูง ณ ที่นี้

ขอขอบคุณคณาจารย์ภาควิชาวิจัยและพัฒนาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคามทุกท่าน ที่กรุณาให้ความรู้และคำแนะนำในการทำวิทยานิพนธ์แก่ผู้วิจัย และผู้เชี่ยวชาญ ผู้ช่วยศาสตราจารย์ ดร.สมทรง สิทธิ ดร.ตวัญจลักษณ์ พวงนิล อาจารย์รัตนา เพ็งเพราะ อาจารย์วีรพล เลพล และนายตันติกร ขุนาพรม ที่กรุณาเป็นผู้เชี่ยวชาญที่ช่วยตรวจเครื่องมือการวิจัย

ขอขอบคุณผู้อำนวยการโรงเรียนในจังหวัดชัยภูมิ คณะครู และนักเรียนที่ให้ความอนุเคราะห์ในการเก็บรวบรวมขอมูลการวิจัยในครั้งนี้

ขอขอบคุณเพื่อนนิสิตปริญญาโท สาขาวิจัยและประเมินผลการศึกษา ที่ให้ความช่วยเหลือในการเก็บรวบรวมข้อมูล ขอขอบคุณเจ้าหน้าที่สำนักงานมหาวิทยาลัยมหาสารคามทุกท่านที่กรุณาให้ความอนุเคราะห์ และให้ความสะดวกในการประสานงานเอกสารต่าง ๆ

ขอกราบขอบพระคุณคุณพ่อบุญเพ็ง และคุณแม่ติ๋ว โพธิ์สิงห์ และขอขอบคุณเพื่อนร่วมงานที่ช่วยเหลือและให้กำลังใจมาโดยตลอด คุณค่าและประโยชน์ของวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบบูชาพระคุณบิดา มารดา บุพการี และบูรพาจารย์ ตลอดจนผู้มีพระคุณทุกท่าน

ศราวุทธ โพธิ์สิงห์

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... ฉ
กิตติกรรมประกาศ ..... ซ
สารบัญ ..... ฌ
สารบัญตาราง ..... ฎ
สารบัญภาพประกอบ ..... ฐ
บทที่ 1 บทนำ 1
ภูมิหลัง 1
ความมุ่งหมายของการวิจัย 5
ความสำคัญของการวิจัย 5
ขอบเขตของการวิจัย 6
นิยามศัพท์เฉพาะ 7
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง 10
ตัวชี้วัดและสาระการเรียนรู้แกนกลางการศึกษา กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานพุทธศักราช 2551 (ฉบับปรับปรุง พ.ศ. 2560) 10
โครงการประเมินผลนักเรียนร่วมกับนานาชาติ 17
การรู้เรื่องวิทยาศาสตร์ 18
แบบทดสอบ 31
ทฤษฎีการตอบสนองของข้อสอบ 38
การทำหน้าที่ต่างกันของข้อสอบ 42
งานวิจัยที่เกี่ยวข้อง 62

งานวิจัยในประเทศ 62
งานวิจัยต่างประเทศ 66
กรอบแนวคิดในการวิจัย 71
บทที่ 3 วิธีดำเนินการวิจัย 72
ประชากรและกลุ่มตัวอย่าง 72
เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล 74
ขั้นตอนดำเนินการสร้างและหาคุณภาพของเครื่องมือ 74
การเก็บรวบรวมข้อมูล 80
การวิเคราะห์ข้อมูล 81
สถิติที่ใช้ในการวิเคราะห์ข้อมูล 83
บทที่ 4 ผลการวิเคราะห์ข้อมูล 87
สัญลักษณ์ที่ใช้ในการเสนอผลการวิเคราะห์ข้อมูล 87
ลำดับขั้นตอนในการเสนอผลการวิเคราะห์ข้อมูล 88
ผลการวิเคราะห์ข้อมูล 88
บทที่ 5 สรุปผล อภิปรายผล และข้อเสนอแนะ 106
ความมุ่งหมายของการวิจัย 106
สรุปผล 106
อภิปรายผล 108
ข้อเสนอแนะ 112
บรรณานุกรม 113
ภาคผนวก 120
ภาคผนวก ก รายชื่อผู้เชี่ยวชาญในการวิเคราะห์ความสอดคล้อง (IOC) ระหว่างนิยามและข้อคำถาม แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 121

ภาคผนวก ข ตัวอย่างหนังสือขอความอนุเคราะห์เป็นผู้เชี่ยวชาญตรวจสอบเครื่องมือที่ใช้ในงานวิจัย 123

ภาคผนวก ค ตัวอย่างข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 125

ภาคผนวก ง ตัวอย่าง Print Out ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC โดยโปรแกรม Mplus 180

ประวัติผู้เขียน 187

สารบัญตาราง

หน้า

ตาราง 1 ระดับของบริบทสำหรับการประเมินการรู้เรื่องวิทยาศาสตร์ของ PISA 21

ตาราง 2 จำนวนกลุ่มตัวอย่างที่ใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำแนกตามขนาดของโรงเรียน 73

ตาราง 3 จำนวนข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ที่ใช้จำแนกตามกรอบโครงสร้างการประเมินผลนักเรียนโครงการ PISA 2015.. 78

ตาราง 4 วิธีการพิจารณา Mplus Output และการตรวจสอบดัชนีความสอดคล้องของโมเดล........ 82

ตาราง 5 ดัชนีความสอดคล้องของข้อสอบกับเนื้อหา แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 89

ตาราง 6 การวิเคราะห์ค่าความยาก ค่าอำนาจจำแนก และค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 92

ตาราง 7 การตรวจสอบการแจกแจงแบบปกติคะแนนแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำแนกตามตัวแปรเพศ... 94

ตาราง 8 การวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson’s Product Moment Correlation Coefficient) รายข้อ 95

ตาราง 9 การวิเคราะห์ค่าความยากของข้อสอบ (b) ค่าอำนาจจำแนกของข้อสอบ (a) แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3... 99

ตาราง 10 ผลการตรวจสอบความสอดคล้องกลมกลืนของโมเดล CFA ของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC จำแนกตามเพศ.................................... 101

ตาราง 11 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC 102

ตาราง 12 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC จำแนกตามเพศ.. 105

สารบัญภาพประกอบ

หน้า

ภาพประกอบ 1 แผนภาพสาระการเรียนรู้แกนกลางการศึกษา กลุ่มสาระการเรียนรู้วิทยาศาสตร์ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานพุทธศักราช 2551 (ฉบับปรับปรุง พ.ศ. 2560).................. 12

ภาพประกอบ 2 กรอบการประเมินการรู้เรื่องวิทยาศาสตร์ของ PISA 2015 20

ภาพประกอบ 3 ตัวอย่างข้อสอบตามกรอบโครงสร้าง PISA แบบเลือกตอบ 4 ตัวเลือก.................. 30

ภาพประกอบ 4 โค้งคุณลักษณะของข้อสอบแบบ 1 พารามิเตอร์ 40

ภาพประกอบ 5 โค้งคุณลักษณะของข้อสอบ แบบ 2 พารามิเตอร์.. 41

ภาพประกอบ 6 โค้งคุณลักษณะของข้อสอบแบบ 3 พารามิเตอร์ 42

ภาพประกอบ 7 โมเดลย่อยของ MIMIC 56

ภาพประกอบ 8 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี MIMIC แบบเอกรูป............ 57

ภาพประกอบ 9 โมเดลการวิเคราะห์องค์ประกอบตามแนวคิด IRT.. 59

ภาพประกอบ 10 โมเดลการวิเคราะห์การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ MIMIC Model โดยใช้ตัวแปรสาเหตุ 1 ตัว 60

ภาพประกอบ 11 โมเดลการวิเคราะห์การตรวจสอบการทาหน้าที่ต่างกันของข้อสอบ MIMIC Model 61

ภาพประกอบ 12 กรอบแนวคิดในการวิจัยการสร้างแบบทดสอบการรู้เรื่องวิทยาศาสตร์ที่มีการทำหน้าที่ต่างกันของข้อสอบ.. 71

ภาพประกอบ 13 ขั้นตอนการสร้างและพัฒนาเครื่องมือแบบทดสอบการรู้เรื่องวิทยาศาสตร์........... 75

บทที่ 1

บทนำ

ภูมิหลัง

วิทยาศาสตร์มีบทบาทสำคัญยิ่งในโลกปัจจุบันและอนาคต เพราะวิทยาศาสตร์เกี่ยวข้องกับทุกคน ทั้งในการดำรงชีวิตประจำวันและการงานอาชีพต่าง ๆ ตลอดจนเทคโนโลยี เครื่องมือเครื่องใช้และผลผลิตต่าง ๆ ที่มนุษย์ได้ใช้เพื่ออำนวยความสะดวกในชีวิตและการทำงานเหล่านี้ล้วนเป็นผลของความรู้ทางวิทยาศาสตร์ ผสมผสานกับความคิดสร้างสรรค์และศาสตร์ วิทยาศาสตร์ช่วยให้มนุษย์ได้พัฒนาวิธีคิด ทั้งความคิดเป็นเหตุเป็นผล คิดสร้างสรรค์ คิดวิเคราะห์ วิจารณ์ มีทักษะสำคัญในการค้นคว้าหาความรู้ มีความสามารถในการแก้ปัญหาอย่างเป็นระบบ สามารถตัดสินใจโดยใช้ข้อมูลที่หลากหลายและมีประจักษ์พยานที่ตรวจสอบได้ วิทยาศาสตร์เป็นวัฒนธรรมของโลกสมัยใหม่ ซึ่งเป็นสังคมแห่งฐานความรู้ (Knowledge-based Society) (กระทรวงศึกษาธิการ, 2552) ทุกคนต้องได้รับการพัฒนาด้านวิทยาศาสตร์ เพื่อให้มีความรู้ความเข้าใจในธรรมชาติและเทคโนโลยีที่มนุษย์สร้างขึ้น สามารถนำความรู้ไปใช้อย่างมีเหตุผล สร้างสรรค์ และมีคุณธรรม (กระทรวงศึกษาธิการ, 2551) ซึ่งวิทยาศาสตร์ถือเป็นเนื้อหาวิชาหลักในการจัดการเรียนรู้ในศตวรรษที่ 21 ดังนั้น ทุกคนจึงจำเป็นต้องได้รับการพัฒนาให้รู้วิทยาศาสตร์และสามารถนำวิทยาศาสตร์และเทคโนโลยีไปใช้ให้เกิดประโยชน์กับการดำรงชีวิต และสามารถนำวิทยาศาสตร์ไปใช้ในการแก้ปัญหาของบุคคล สังคม และประเทศชาติได้อย่างเหมาะสม

ความรู้และความเข้าใจทางวิทยาศาสตร์และเทคโนโลยีเป็นเครื่องมือสำคัญอย่างยิ่งในการเตรียมเยาวชนให้สามารถดำเนินชีวิตในโลกปัจจุบันที่มีวิทยาศาสตร์และเทคโนโลยีเป็นพื้นฐานและส่งผลกระทบต่อทุกชีวิตในทุกระดับ ทั้งตัวบุคคล ในอาชีพการงานและในสังคมวัฒนธรรมของทุก ๆ ชีวิต ทำให้บุคคลสามารถรับรู้และตัดสินประเด็นปัญหาของสังคมที่เกิดจากผลกระทบของวิทยาศาสตร์และเทคโนโลยีอย่างมีความรู้ความเข้าใจ มีส่วนร่วมในสังคมระดับชุมชน ระดับประเทศ และระดับโลกอย่างเต็มภาคภูมิ เป้าหมายของการศึกษาวิทยาศาสตร์ คือ การทำให้นักเรียนทุกคนมีการรู้เรื่องวิทยาศาสตร์ (Scientific Literacy) ซึ่งรวมถึงความรู้มิติต่าง ๆ ทางวิทยาศาสตร์ ได้แก่ ความรู้ความสามารถทางสติปัญญา กระบวนการแสวงหาความรู้ทางวิทยาศาสตร์ และความสัมพันธ์ระหว่างวิทยาศาสตร์กับเทคโนโลยีด้วย ในการประเมินการรู้วิทยาศาสตร์ PISA ต้องการหาตัวชี้วัดว่านักเรียนเรียนรู้ทฤษฎีและแนวคิดพื้นฐานทางวิทยาศาสตร์มากน้อยเพียงใด มีสมรรถนะทางวิทยาศาสตร์ และสามารถแก้ปัญหาในชีวิตจริงที่เกี่ยวข้องกับวิทยาศาสตร์และเทคโนโลยีได้ดีเพียงใด

ทั้งนี้เพื่อหาคำตอบให้กับระดับนโยบายและระดับปฏิบัติในการปรับปรุงการเรียนการสอนวิทยาศาสตร์ ประเด็นหลักที่ PISA ให้ความสำคัญในการประเมินผลวิทยาศาสตร์ คือ การประเมินผลต้องให้ความชัดเจนกับสมรรถนะที่เหมาะสมกับนักเรียนวัย 15 ปี กล่าวคือ ต้องตั้งอยู่บนพื้นฐานที่ว่านักเรียนวัย 15 ปี ควรรู้อะไร ทำอะไรได้ และให้คุณค่ากับอะไรบ้าง ภายในขอบเขตของบริบทส่วนตัว บริบทของสังคม และบริบทของโลก ซึ่งประเด็นนี้อาจแตกต่างจากการเรียนการสอนและการประเมินผลตามปกติในโรงเรียน (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2560 ก)

โครงการประเมินผลนักเรียนร่วมกับนานาชาติ (Programme for International Student Assessment หรือ PISA) ริเริ่มโดยองค์การเพื่อความร่วมมือทางเศรษฐกิจและการพัฒนา (Organization for Economic Co-operation and Development หรือ OECD) มีวัตถุประสงค์เพื่อประเมินคุณภาพของระบบการศึกษาทั้งประเทศสมาชิก OECD และประเทศที่ไม่ใช่สมาชิก OECD (Partner Countries) ในการเตรียมความพร้อมให้เยาวชนมีศักยภาพสำหรับการแข่งขันในอนาคต ใช้ความร่วมมือและความเชี่ยวชาญของนานาชาติในการวางกรอบโครงสร้างการประเมิน การสร้างเครื่องมือ และการศึกษาวิจัย เพื่อประกันคุณภาพของการศึกษาวิจัยให้สามารถเปรียบเทียบกันได้ในระดับนานาชาติ และข้อมูลที่ได้สามารถชี้บอกถึงคุณภาพการศึกษาของประเทศต่าง ๆ ได้ ปัจจุบันมีประเทศเข้าร่วมโครงการ PISA มากกว่า 70 ประเทศทั่วโลก โครงการ PISA ประเมินผลนักเรียนในระดับโรงเรียนและดำเนินการอย่างต่อเนื่องทุก ๆ 3 ปี กลุ่มตัวอย่างของ PISA คือ นักเรียนกลุ่มอายุ 15 ปี ซึ่งโดยสากลถือว่าเป็นวัยจบการศึกษาภาคบังคับ การประเมินของ PISA เน้นการประเมินความสามารถของนักเรียนในการใช้ความรู้และทักษะเพื่อแก้ปัญหาในชีวิตจริง หรือที่เรียกว่า “การรู้เรื่อง” (Literacy) ในสามด้าน ได้แก่ การรู้เรื่องการอ่าน (Reading Literacy) การรู้เรื่องคณิตศาสตร์ (Mathematical Literacy) และการรู้เรื่องวิทยาศาสตร์ (Scientific Literacy) ซึ่งการรู้เรื่องทั้งสามด้านนี้ถือว่าเป็นสิ่งจำเป็นสำหรับการเรียนรู้ตลอดชีวิต และเป็นสิ่งที่ประชากรจำเป็นต้องมีเพื่อการพัฒนาและการแข่งขันทางเศรษฐกิจของประเทศ สำหรับประเทศไทยสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยีทำหน้าที่เป็นศูนย์แห่งชาติดำเนินการวิจัยโครงการ PISA โดยได้เข้าร่วมโครงการตั้งแต่ พ.ศ. 2543 หรือใน PISA 2000 ซึ่งเป็นรอบแรกของการประเมินและได้เข้าร่วมโครงการอย่างต่อเนื่อง ในแต่ละรอบการประเมิน PISA จะสอบสามวิชาพื้นฐานดังกล่าว โดยมีสัดส่วนของข้อสอบแต่ละวิชาแตกต่างกัน ขึ้นอยู่กับว่าในรอบการประเมินนั้นเน้นการประเมินวิชาใด สำหรับรอบการประเมิน PISA 2015 เน้นการประเมินวิชาวิทยาศาสตร์ จึงมีสัดส่วนของข้อสอบวิทยาศาสตร์ประมาณ 60% ส่วนวิชาคณิตศาสตร์และการอ่าน มีสัดส่วนของข้อสอบประมาณวิชาละ 20% ซึ่งใน PISA 2015 ประเมินวิทยาศาสตร์เป็นวิชาหลักนับเป็นครั้งที่สองต่อจาก PISA 2006 (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2560 ก)

ผลการประเมิน PISA 2015 ของประเทศไทย แนวโน้มจากการประเมิน PISA 2000 ถึง PISA 2015 พบว่า ผลการประเมินทั้งสามด้านมีแนวโน้มลดลง ผลการประเมินการรู้เรื่องวิทยาศาสตร์ คะแนนเฉลี่ย OECD ของวิทยาศาสตร์ใน PISA 2015 เป็นคะแนนมาตรฐานที่ 493 คะแนน คะแนนเฉลี่ยวิทยาศาสตร์ของนักเรียนไทย คือ 421 คะแนน ซึ่งต่ำกว่าค่าเฉลี่ย OECD แนวโน้มคะแนนวิทยาศาสตร์ของนักเรียนไทยโดยรวมลดต่ำลง จาก PISA 2012 ถึง PISA 2015 คะแนนวิทยาศาสตร์ลดลงอย่างมีนัยสำคัญ (23 คะแนน) และคะแนนลดลงจนเท่ากับการประเมินรอบ PISA 2006 ที่วิทยาศาสตร์เป็นวิชาหลัก และในภาพรวมระดับนานาชาติความแตกต่างระหว่างเพศด้านวิทยาศาสตร์ พบว่า นักเรียนชายมีคะแนนวิทยาศาสตร์สูงกว่านักเรียนหญิงอย่างมีนัยสำคัญทางสถิติ (4 คะแนน) (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2560 ข)

ผลการประเมินนักเรียนจะเป็นข้อมูลสำคัญช่วยปรับปรุงและพัฒนาคุณภาพการศึกษา โดยพิจารณาจากผลการประเมินผลสัมฤทธิ์ทางการเรียนที่มีความสำคัญทั้งระดับผู้เรียน ระดับสถานศึกษา ระดับเขตพื้นที่การศึกษา และระดับชาติ แบบทดสอบจึงเป็นเครื่องมือที่ใช้ในการวัดผลประเมินผลทางการศึกษา โดยตรวจสอบว่าผู้เข้าสอบนั้นจะมีคุณลักษณะแฝงหรือความสามารถอยู่ในระดับใด ดังนั้นการสร้างและการตรวจสอบคุณภาพของแบบทดสอบจะต้องคำนึงถึงความตรง (Validity) เป็นสำคัญ ทั้งนี้เพราะว่าความตรงเป็นคุณสมบัติที่แสดงถึงความสามารถในการวัดได้อย่างถูกต้องแม่นยำ ถ้าผลการวัดได้ค่าที่ใกล้เคียงกับค่าคุณลักษณะที่แท้จริงเพียงใดก็ถือว่าการวัดมีความตรงมากขึ้นเพียงนั้น

การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning : DIF) เป็นการตรวจสอบคุณภาพด้านความตรง โดยเป็นการตรวจสอบในประเด็นของความยุติธรรมของข้อสอบและแบบทดสอบ (Item and Test Unfairness) แต่เดิมใช้คำว่า ความลำเอียงของข้อสอบ (Item Bias) หรือความลำเอียงของแบบทดสอบ (Test Bias) ซึ่งต่อมามีการเปลี่ยนมาใช้คำที่เหมาะสมกว่าเป็นการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) จากผลการตรวจข้อสอบของผู้สอบกลุ่มต่าง ๆ ในประชากรมีมานานแล้ว แต่การศึกษาคุณภาพด้านความยุติธรรมของข้อสอบหรือแบบทดสอบระหว่างผู้สอบกลุ่มต่าง ๆ ปลายปี ค.ศ. 1960 มีการเสนอวิธีการต่าง ๆ เพื่อตรวจสอบความลำเอียงของข้อสอบ (Item Bias) ความลำเอียงของแบบทดสอบ (Test Bias) และความลำเอียงในการคัดเลือก (Selection Bias) โดยนิยามความลำเอียงว่าเป็นความคลาดเคลื่อนอย่างเป็นระบบ (Systematic Error) ซึ่งเป็นการขจัดข้อสอบที่ทำให้เกิดปัญหาความไม่ยุติธรรมระหว่างกลุ่มข้อสอบกลุ่มต่าง ๆ ที่มีลักษณะบางอย่างแตกต่างกัน (ศิริชัย กาญจนวาสี, 2555)

การทดสอบแต่ละครั้ง ผู้สอบอาจจะมีลักษณะแตกต่างกัน เช่น เชื้อชาติ ศาสนา วัฒนธรรม ภูมิลำเนา สังคม เพศ ภาษา อายุ สภาพทางเศรษฐกิจ และประสบการณ์ เป็นต้น ทำให้ผู้สอบไม่ได้รับความยุติธรรมในการทำข้อสอบ โดยข้อสอบบางข้ออาจมีความลำเอียงเข้าข้างผู้สอบกลุ่มย่อยบางกลุ่ม

ของผู้เข้าสอบทั้งหมด ทำให้เกิดการได้เปรียบหรือเสียเปรียบระหว่างกลุ่มผู้สอบด้วยกันทั้ง ๆ ที่สอบด้วยข้อสอบข้อเดียวกันหรือแบบทดสอบฉบับเดียวกัน แสดงว่าแบบทดสอบหรือข้อสอบฉบับนั้นขาดความตรง คือ ไม่ได้วัดความสามารถหลักที่ต้องการวัด (Target Ability) เพียงอย่างเดียว แต่ยังวัดความสามารถแทรกซ้อนที่ไม่ต้องการวัด (Nuisance Ability) อีกด้วย เช่น แบบทดสอบวัดความสามารถด้านคำศัพท์ภาษาไทยฉบับหนึ่ง ข้อสอบบางข้ออาจถามความรู้สำหรับผู้ชายเป็นพิเศษ เช่น ความรู้เรื่องกีฬา ในบางข้ออาจถามความรู้สำหรับผู้หญิงเป็นพิเศษ เช่น ความรู้เรื่องการตัดเย็บ จากสถานการณ์นี้แบบทดสอบวัดความสามารถคำศัพท์ในวิชาภาษาไทยเป็นความสามารถหลักที่ต้องการวัด ส่วนความสามารถด้านกีฬาและด้านการตัดเย็บเป็นความสามารถแทรกซ้อน ทำให้การตอบข้อสอบกลุ่มย่อยมีโอกาสการตอบถูกไม่เท่ากันขึ้นอยู่กับกลุ่มใดมีความสามารถแทรกซ้อนสูงกว่ากัน ทั้ง ๆ ที่ระดับความสามารถหลักที่ต้องการวัดเท่ากัน จึงทำให้ข้อสอบนั้นทำหน้าที่ต่างกัน

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เป็นการเปรียบเทียบผลการตอบข้อสอบระหว่างกลุ่มผู้สอบอย่างน้อย 2 กลุ่มขึ้นไป กลุ่มแรก เรียกว่า กลุ่มเปรียบเทียบ (Focal Group : F) เป็นกลุ่มที่ผู้วิจัยสนใจศึกษา และคาดว่าเป็นกลุ่มที่เสียประโยชน์ในการตอบข้อสอบ และกลุ่มที่สองเรียกว่า กลุ่มอ้างอิง (Reference Group : R) เป็นกลุ่มที่คาดว่าจะได้ประโยชน์จากการตอบข้อสอบได้ถูกต้อง วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่สำคัญ ๆ ได้แก่ การวิเคราะห์ความแปรปรวน (Analysis of Variance : ANOVA) วิธีการวิเคราะห์การถดถอยโลจิสติก (Logistic Regression : LR) วิธีแปลงค่าความยากของข้อสอบ (Transformed Item Difficulty : TID) วิธีวัดพื้นที่ความแตกต่างระหว่างโค้งการตอบสนองข้อสอบ (Item Response Theory–D2 : IRT-D2) วิธีแมนเทล แฮนส์เซล (Mantel-Haenszel : MH) วิธีไคสแควร์ของลอร์ด (Lord’s Chi Square) วิธีอัตราส่วนไลด์ลิฮูด ลอกลิเนียร์ (Loglinear Likelihood Ratio) วิธี SIBTEST และวิธีมิมิค (Multiple Indicators and Multiple Causes : MIMIC) (ศิริชัย กาญจนวาสี, 2555)

วิธี MIMIC เป็นวิธีหนึ่งที่ใช้ในโปรแกรม Mplus สำหรับการวิเคราะห์คุณภาพของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) ซึ่งวิธี MIMIC เป็นโมเดลลิสเรลที่มีตัวแปรแฝงเป็นตัวแปรเดียว โดยที่ตัวแปรแฝงนั้นได้รับอิทธิพลจากตัวแปรภายนอกสังเกตได้หลายตัวแปรและส่งอิทธิพลไปยังตัวแปรภายในสังเกตได้หลายตัวแปร กล่าวอีกอย่างหนึ่งคือเป็นโมเดลลิสเรลของคุณลักษณะแฝงที่มีหลายสาเหตุและวัดได้จากตัวบ่งชี้หลายตัว ลักษณะโมเดลจะเห็นว่าการวัดตัวแปรภายนอกสังเกตได้ต้องมีข้อตกลงข้างต้นว่าไม่มีความคลาดเคลื่อนในการวัด วิธี MIMIC นี้เป็นประโยชน์มากในการตรวจสอบความเป็นเอกมิติ (Unidimensionality) การวัดผลการศึกษาสามารถวิเคราะห์ค่าพารามิเตอร์คุณลักษณะข้อสอบ และค่าความสามารถของผู้สอบไม่สามารถสังเกตโดยตรงจึงต้องประมาณจากการตอบข้อสอบ การประมาณค่าพารามิเตอร์ในทฤษฎีการตอบสนองข้อสอบ

การวิจัยครั้งนี้ ผู้วิจัยมุ่งที่จะพัฒนาแบบทดสอบการรู้เรื่องวิทยาศาสตร์ เพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี MIMIC โดยใช้โปรแกรมประยุกต์ Mplus โดยใช้ข้อมูลในการสร้างแบบทดสอบตามกรอบโครงสร้างการประเมินการรู้เรื่องวิทยาศาสตร์ โครงการประเมินผลนักเรียนนานาชาติ (PISA) ค.ศ. 2015 ได้แก่ สถานการณ์ บริบท สมรรถนะความรู้ทางวิทยาศาสตร์ : แนวคิดและเนื้อหาที่ครอบคลุม ซึ่งเป็นแบบทดสอบที่เป็นแบบเลือกตอบ 4 ตัวเลือก โดยไม่คำนึงความแตกต่างของคุณลักษณะพื้นฐานในระดับนักเรียน เพื่อเตรียมความพร้อมให้กับนักเรียนได้ฝึกการคิดวิเคราะห์สถานการณ์ สามารถนำความรู้ทางวิทยาศาสตร์มาแก้ปัญหาตามสถานการณ์ที่กำหนด ซึ่งเป็นการพัฒนาองค์ความรู้ นำเสนอสารสนเทศที่เป็นประโยชน์ทางด้านวัดผลการศึกษา และการนำความรู้ไปใช้ปฏิบัติจริง

ความมุ่งหมายของการวิจัย

การวิจัยครั้งนี้มีจุดมุ่งหมาย เพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 โดยมีจุดมุ่งหมายเฉพาะดังนี้

1. เพื่อสร้างและหาคุณภาพแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3
2. เพื่อตรวจสอบการทำหน้าที่ต่างกัน (DIF) ของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

ความสำคัญของการวิจัย

1. ได้แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ตามแนวการประเมินผลนักเรียนนานาชาติ (PISA) ที่มีคุณภาพ
2. เป็นข้อมูลในการตัดสินใจของผู้ที่จะนำวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบไปใช้ในการตรวจสอบคุณภาพของข้อสอบ เพื่อขจัดความลำเอียงของข้อสอบ ตลอดจนพัฒนาแบบทดสอบให้มีคุณภาพมากยิ่งขึ้น
3. เพื่อเป็นแนวทางสำหรับคุณครูและผู้สนใจ สามารถนำแบบทดสอบในด้านการรู้เรื่องทางวิทยาศาสตร์ไปใช้ในการประเมินผลผู้เรียนได้

ขอบเขตของการวิจัย

การวิจัยในครั้งนี้ ผู้วิจัยกำหนดขอบเขตการวิจัยตามขั้นตอนการดำเนินการวิจัย โดยพิจารณาขอบเขตด้านเนื้อหา ขอบเขตด้านแหล่งข้อมูล ดังนี้

1. ประชากรและกลุ่มตัวอย่าง

1.1 ประชากร

นักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2561 ในจังหวัดชัยภูมิ จำนวน 6,890 คน จาก 66 โรงเรียน

1.2 กลุ่มตัวอย่าง

กลุ่มตัวอย่างในงานวิจัยครั้งนี้ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2561 ในจังหวัดชัยภูมิ จำนวน 689 คน จากการสุ่มแบบหลายขั้นตอน (Multi–Stage Random Sampling)

2. กรอบโครงสร้างการประเมินการรู้เรื่องวิทยาศาสตร์ ตามวัตถุประสงค์ของการประเมิน PISA 2015 จึงได้กำหนดกรอบโครงสร้างการประเมินผลการรู้เรื่องวิทยาศาสตร์ ซึ่งประกอบด้วย 4 องค์ประกอบที่เกี่ยวข้องกัน ได้แก่

2.1 บริบท หมายถึง การรับรู้ถึงสถานการณ์ในชีวิต ในระดับส่วนตัว ระดับชาติ และระดับโลก ทั้งที่เป็นเรื่องในปัจจุบันหรือในอดีตที่ผ่านมา ซึ่งจำเป็นต้องมีความเข้าใจเรื่องวิทยาศาสตร์และเทคโนโลยี

2.2 ความรู้ทางวิทยาศาสตร์ หมายถึง ความเข้าใจในข้อเท็จจริง แนวคิดหลัก และทฤษฎีสำคัญที่ทำให้เกิดความรู้พื้นฐานทางวิทยาศาสตร์ ความรู้ประกอบด้วยความรู้เกี่ยวกับธรรมชาติของโลกและสิ่งประดิษฐ์ทางเทคโนโลยี (ความรู้ด้านเนื้อหา) ความรู้เกี่ยวกับวิธีการในการสร้างแนวคิดต่าง ๆ (ความรู้ด้านกระบวนการ) และความเข้าใจในเหตุผลพื้นฐานของกระบวนการสร้างความรู้ (ความรู้เกี่ยวกับการได้มาของความรู้)

2.3 สมรรถนะทางวิทยาศาสตร์ หมายถึง ความสามารถในการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ และการแปลความหมายข้อมูลและใช้ประจักษ์พยานเชิงวิทยาศาสตร์

2.4 เจตคติต่อวิทยาศาสตร์ หมายถึง การแสดงการตอบสนองต่อวิทยาศาสตร์และเทคโนโลยีด้วยความสนใจ ให้ความสำคัญกับกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ และรับรู้และตระหนักถึงปัญหาสิ่งแวดล้อม ในการวิจัยครั้งนี้ไม่นำเจตคติต่อวิทยาศาสตร์มาใช้ในการออกข้อสอบ เนื่องจาก PISA 2015 ประเมินเจตคติต่อวิทยาศาสตร์ของนักเรียนโดยใช้แบบสอบถาม

3. ตัวแปรที่ใช้ในการวิจัย

3.1 ตัวแปรอิสระ ได้แก่ เพศ

3.2 ตัวแปรตาม ได้แก่ ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีมิมิค (Multiple Indicators and Multiple Causes : MIMIC)

4. ระยะเวลา

ระยะเวลาที่ใช้ในการศึกษาค้นคว้าวิจัยในครั้งนี้ คือ 1 พฤษภาคม 2561 ถึงวันที่ 30 มิถุนายน 2564

นิยามศัพท์เฉพาะ

1. การรู้เรื่องวิทยาศาสตร์ (Scientific Literacy) หมายถึง ความสามารถในการเชื่อมโยงสิ่งต่าง ๆ เข้ากับประเด็นที่เกี่ยวข้องกับวิทยาศาสตร์ และแนวคิดทางวิทยาศาสตร์ได้อย่างไตร่ตรอง บุคคลที่รู้เรื่องวิทยาศาสตร์ (Scientifically Literate Person) จะสื่อสารพูดคุยในประเด็นที่เกี่ยวข้องกับวิทยาศาสตร์และเทคโนโลยีอย่างเป็นเหตุเป็นผล ซึ่งการประเมินผลการรู้เรื่องวิทยาศาสตร์ประกอบด้วย 3 องค์ประกอบที่เกี่ยวข้องกัน ได้แก่

1.1 สมรรถนะมีอยู่ 3 สมรรถนะ คือ 1) การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ 2) การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ และ 3) การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์

1.1.1 การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ หมายถึง มีความสามารถในการรับรู้ เสนอ และประเมินคำอธิบายที่เกี่ยวข้องกับปรากฏการณ์ทางธรรมชาติและเทคโนโลยี

1.1.2 การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ หมายถึง มีความสามารถในการอธิบายและประเมินคุณค่าของการสำรวจตรวจสอบทางวิทยาศาสตร์และนำเสนอแนวทางในการตอบคำถามอย่างเป็นวิทยาศาสตร์

1.1.3 การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์ หมายถึง มีความสามารถในการวิเคราะห์และประเมินข้อมูล คำกล่าวอ้าง และข้อโต้แย้งในหลากหลายรูปแบบ และลงข้อสรุปทางวิทยาศาสตร์ได้อย่างเหมาะสม

1.2 ความรู้ทางวิทยาศาสตร์ (Scientific Knowledge) ที่ PISA 2015 กำหนดไว้นั้นครอบคลุมความรู้ทางวิทยาศาสตร์ 3 ด้าน ได้แก่ 1) ความรู้ด้านเนื้อหา (Content Knowledge) 2) ความรู้ด้านกระบวนการ (Procedural Knowledge) และ 3) ความรู้เกี่ยวกับการได้มาของความรู้ (Epistemic Knowledge) สำหรับงานวิจัยนี้ผู้วิจัยมุ่งเน้นศึกษาเฉพาะความรู้ด้านเนื้อหา

1.2.1 ความรู้ด้านเนื้อหา เป็นความรู้เกี่ยวกับข้อเท็จจริง แนวความคิดหลัก แนวคิด และทฤษฎีเกี่ยวกับโลกธรรมชาติ โดย PISA เลือกประเมินความรู้ในสาขาวิชาหลัก 3 เนื้อหา ประกอบด้วย 1) ระบบทางกายภาพ 2) ระบบสิ่งมีชีวิต และ 3) ระบบของโลกและอวกาศ

1.3 บริบท หมายถึง สถานการณ์หรือบริบทที่ปัญหาทางวิทยาศาสตร์ตั้งอยู่การที่วิทยาศาสตร์ได้เข้าไปเกี่ยวข้องกับการแก้ปัญหาที่อยู่ในบริบทหนึ่ง ซึ่งเป็นบริบทในโลกชีวิตจริงที่มีปัญหานั้นอยู่ ซึ่งมีอยู่ 3 ระดับ คือ 1) บริบทระดับส่วนตัว 2) บริบทระดับท้องถิ่น/ชาติ และ 3) บริบทระดับโลก

2. แบบทดสอบการรู้เรื่องวิทยาศาสตร์ หมายถึง เครื่องมือที่ใช้วัดการรู้เรื่องวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 54 ข้อ ที่มีองค์ประกอบสัมพันธ์ซึ่งกันและกัน ได้แก่ บริบท สมรรถนะทางวิทยาศาสตร์ และความรู้ทางวิทยาศาสตร์ (ความรู้ด้านเนื้อหา) 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 18 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ มีการกำหนดหลักเกณฑ์การให้คะแนนในข้อสอบแต่ละข้อที่ชัดเจน โดยเมื่อทำถูกให้ 1 คะแนน ทำผิดได้ 0 คะแนน และเป็นแบบปรนัยชนิดเลือกตอบ 4 ตัวเลือก

3. คุณภาพของแบบทดสอบ หมายถึง คุณลักษณะที่ดีของแบบทดสอบ ซึ่งได้แก่ ค่าความเที่ยงตรง ค่าอำนาจจำแนก ค่าความยาก และค่าความเชื่อมั่น ดังนี้

3.1 ความเที่ยงตรง (Validity) หมายถึง คุณสมบัติของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ที่วัดได้ตรงตามลักษณะหรือจุดมุ่งหมาย ซึ่งในการศึกษาค้นคว้าครั้งนี้ตรวจสอบคุณภาพด้านความเที่ยงตรงตามเนื้อหา (Content Validity) เป็นการตรวจสอบว่าข้อความวัดได้ตรงตามเนื้อหาการรู้เรื่องวิทยาศาสตร์ที่กำหนดให้หรือไม่ หาได้จากค่าเฉลี่ยที่เกิดจากการนำค่าเฉลี่ยของความคิดเห็นผู้เชี่ยวชาญ จำนวน 5 คน ซึ่งต้องมีค่าดัชนีความสอดคล้องตั้งแต่ 0.50 ขึ้นไป จึงจะถือว่ามีความเที่ยงตรงเชิงเนื้อหา

3.2 ค่าความยาก (Difficulty) หมายถึง อัตราส่วนระหว่างจำนวนนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ที่ตอบแบบทดสอบข้อนั้นถูกกับจำนวนนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ที่ตอบข้อนั้นทั้งหมด ในการวิจัยนี้คัดเลือกข้อสอบที่มีค่าความยาก (p) ตั้งแต่ 0.20 ถึง 0.80 ตามแนวทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) และกำหนดค่าความยาก (b-parameter) มีค่าระหว่าง -3.00 ถึง +3.00 ตามแนวทฤษฎีการตอบสนองของข้อสอบ (Item Response Theory)

3.3 ค่าอำนาจจำแนก (Discrimination) หมายถึง เป็นการวิเคราะห์ค่าอำนาจจำแนกเป็นการดูความเหมาะสมของรายข้อว่า ข้อคำถามสามารถจำแนกกลุ่มเก่งและกลุ่มอ่อนได้จริงหรือจำแนกผู้ที่มีคุณลักษณะสูงจากผู้มีคุณลักษณะต่ำได้ ในการวิจัยนี้คัดเลือกข้อสอบที่มีค่าอำนาจจำแนก (r) ตั้งแต่ 0.20 ถึง 1.00 ตามแนวทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory)

และกำหนดค่าอำนาจจำแนก (a-parameter) มีค่ามากกว่า 0.30 ตามแนวทฤษฎีการตอบสนองของข้อสอบ (Item Response Theory)

3.4 ค่าความเชื่อมั่น (Reliability) หมายถึง คุณสมบัติของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ได้คงที่แน่นอน ภายใต้สถานการณ์และเงื่อนไขเดียวกัน หาได้จากการใช้วิธีคิดของคูเดอร์-ริชาร์ดสัน KR-20 (Kuder-Richardson Method) ซึ่งเป็นการหาจากการดำเนินการทดสอบเพียงครั้งเดียว

4. การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning : DIF) หมายถึง ข้อสอบที่ผู้สอบต่างกลุ่มกันที่มีระดับความสามารถของผู้เข้าสอบเท่ากัน แต่มีโอกาสในการตอบข้อสอบได้ถูกต้องแตกต่างกัน ในการวิจัยครั้งนี้เพื่อตรวจสอบการทำหน้าที่ต่างกัน (DIF) ของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 โดยตรวจสอบคุณลักษณะตัวแปร ได้แก่ เพศ

5. กลุ่มอ้างอิง (Reference Group) หมายถึง ผู้สอบที่คาดว่าจะเป็นกลุ่มที่ได้ประโยชน์จากข้อสอบลำเอียงและเป็นกลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบที่ได้ถูกต้องสูงกว่าผู้สอบอีกกลุ่มที่ได้คะแนนจากข้อสอบเท่ากัน สำหรับการวิจัยครั้งนี้ ได้แก่ เพศชาย

6. กลุ่มเปรียบเทียบ (Focal Group) หมายถึง ผู้สอบที่คาดว่าจะเป็นกลุ่มที่เสียประโยชน์จากข้อสอบลำเอียง คือ กลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบที่ลำเอียงได้ถูกต้องต่ำกว่าผู้สอบกลุ่มอ้างอิงที่มีคะแนนจากการสอบเท่ากัน สำหรับการวิจัยครั้งนี้ ได้แก่ เพศหญิง

7. วิธีมิมิค (Multiple Indicators and Multiple Causes : MIMIC) หมายถึง โมเดลสมการโครงสร้างที่มีตัวแปรแฝงเพียงตัวแปรเดียว ซึ่งก็คือ แบบทดสอบการรู้เรื่องวิทยาศาสตร์ โดยที่ตัวแปรแฝงจะได้รับอิทธิพลจากตัวแปรภายนอก ซึ่งได้แก่ เพศ และส่งอิทธิพลไปยังตัวแปรสังเกตได้ภายใน ซึ่งได้แก่ ข้อสอบแต่ละข้อในแบบทดสอบนั้น

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในงานวิจัยครั้งนี้ ผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง โดยได้นำเสนอตามหัวข้อต่อไปนี้

1. ตัวชี้วัดและสาระการเรียนรู้แกนกลางการศึกษา กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานพุทธศักราช 2551 (ฉบับปรับปรุง พ.ศ. 2560)
2. โครงการประเมินผลนักเรียนร่วมกับนานาชาติ
3. การรู้เรื่องวิทยาศาสตร์
4. แบบทดสอบ
5. ทฤษฎีการตอบสนองของข้อสอบ
6. การทำหน้าที่ต่างกันของข้อสอบ
7. งานวิจัยที่เกี่ยวข้อง
    7.1 งานวิจัยในประเทศ
    7.2 งานวิจัยต่างประเทศ
8. กรอบแนวคิดในการวิจัย

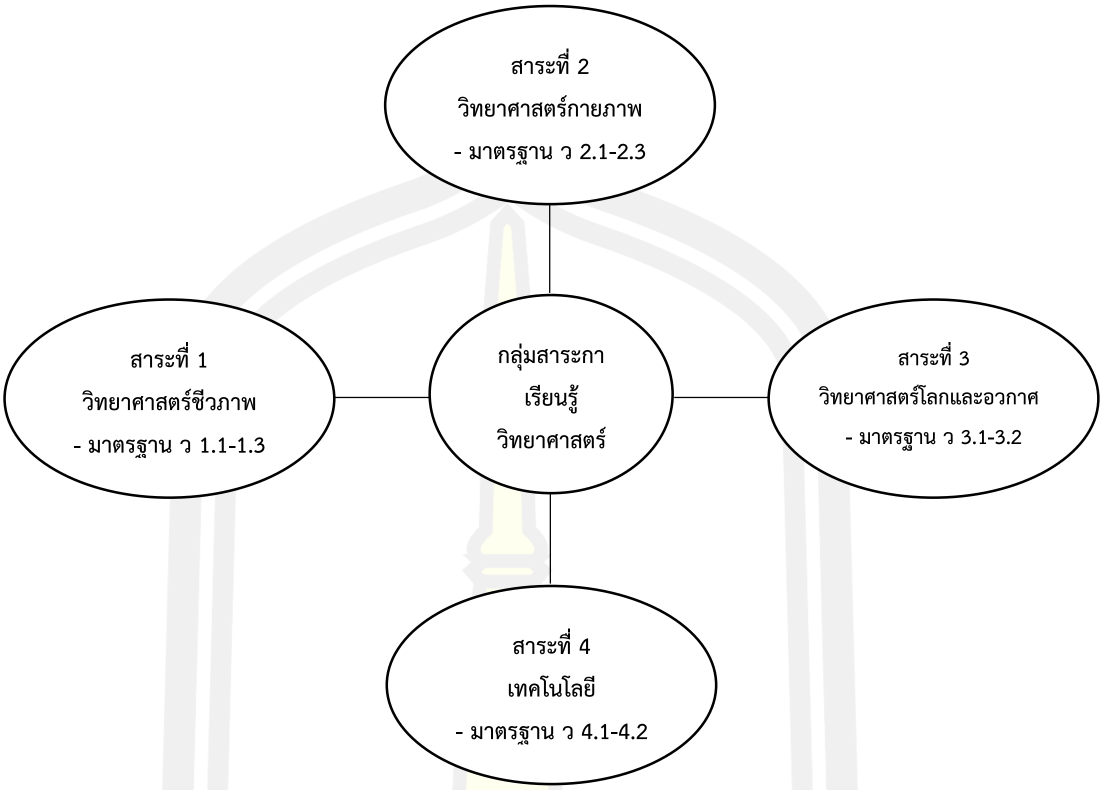
ตัวชี้วัดและสาระการเรียนรู้แกนกลางการศึกษา กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานพุทธศักราช 2551 (ฉบับปรับปรุง พ.ศ. 2560)

ตัวชี้วัดและสาระการเรียนรู้แกนกลางการศึกษา กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานพุทธศักราช 2551 (ฉบับปรับปรุง พ.ศ. 2560) (กระทรวงศึกษาธิการ, 2551) ได้กำหนดสาระการเรียนรู้ออกเป็น 4 สาระ ได้แก่ สาระที่ 1 วิทยาศาสตร์ชีวภาพ สาระที่ 2 วิทยาศาสตร์กายภาพ สาระที่ 3 วิทยาศาสตร์โลกและอวกาศ และสาระที่ 4 เทคโนโลยี มีสาระเพิ่มเติม 4 สาระ ได้แก่ สาระชีววิทยา สาระเคมี สาระฟิสิกส์ สาระโลก ดาราศาสตร์ และอวกาศ ซึ่งองค์ประกอบของหลักสูตรทั้งในด้านของเนื้อหา การจัดการเรียนการสอน และการวัดและประเมินผลการเรียนรู้นั้น มีความสำคัญอย่างยิ่งในการวางรากฐานการเรียนรู้วิทยาศาสตร์ของผู้เรียนในแต่ละระดับชั้นให้มีความต่อเนื่องเชื่อมโยงกัน ตั้งแต่ชั้นประถมศึกษาปีที่ 1 จนถึงชั้นมัธยมศึกษาปีที่ 6 สำหรับกลุ่มสาระการเรียนรู้วิทยาศาสตร์ได้กำหนดตัวชี้วัดและสาระการเรียนรู้แกนกลางที่ผู้เรียนจำเป็นต้องเรียนเป็นพื้นฐาน เพื่อให้สามารถนำความรู้นี้

ไปใช้ในการดำรงชีวิตหรือศึกษาต่อในวิชาชีพที่ต้องใช้วิทยาศาสตร์ได้ โดยจัดเรียงลำดับความยากง่ายของเนื้อหาแต่ละสาระในแต่ละระดับชั้นให้มีการเชื่อมโยงความรู้กับกระบวนการเรียนรู้ และการจัดกิจกรรมการเรียนรู้ที่ส่งเสริมให้ผู้เรียนพัฒนาความคิดทั้งความคิดเป็นเหตุเป็นผล คิดสร้างสรรค์ คิดวิเคราะห์วิจารณ์ มีทักษะที่สำคัญทั้งทักษะกระบวนการทางวิทยาศาสตร์และทักษะในศตวรรษที่ 21 ในการค้นคว้าและสร้างองค์ความรู้ด้วยกระบวนการสืบเสาะหาความรู้ สามารถแก้ปัญหาอย่างเป็นระบบ สามารถตัดสินใจโดยใช้ข้อมูลหลากหลายและประจักษ์พยานที่ตรวจสอบได้

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) ตระหนักถึงความสำคัญของการจัดการเรียนรู้วิทยาศาสตร์ที่มุ่งหวังให้เกิดผลสัมฤทธิ์ต่อผู้เรียนมากที่สุด จึงได้จัดทำตัวชี้วัดและสาระการเรียนรู้แกนกลางกลุ่มสาระการเรียนรู้วิทยาศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานพุทธศักราช 2551 ขึ้น เพื่อให้สถานศึกษา ครูผู้สอน ตลอดจนหน่วยงานต่าง ๆ ได้ใช้เป็นแนวทางในการพัฒนาหนังสือเรียน คู่มือครู สื่อประกอบการเรียนการสอน ตลอดจนการวัดและประเมินผล โดยตัวชี้วัดและสาระการเรียนรู้แกนกลางกลุ่มสาระการเรียนรู้วิทยาศาสตร์ (ฉบับปรับปรุง พ.ศ. 2560) ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ที่จัดทำขึ้นนี้ได้ปรับปรุงเพื่อให้มีความสอดคล้องและเชื่อมโยงกันภายในสาระการเรียนรู้เดียวกัน และระหว่างสาระการเรียนรู้ในกลุ่มสาระการเรียนรู้วิทยาศาสตร์ ตลอดจนการเชื่อมโยงเนื้อหาความรู้ทางวิทยาศาสตร์กับคณิตศาสตร์ด้วย นอกจากนี้ยังได้ปรับปรุงเพื่อให้มีความทันสมัยต่อการเปลี่ยนแปลง และความเจริญก้าวหน้าของวิทยาการต่าง ๆ และทัดเทียมกับนานาชาติ กลุ่มสาระการเรียนรู้วิทยาศาสตร์ สรุปเป็นแผนภาพได้ดังนี้

พหุนํ ปณฺ ฑิโต ชีเว



วิทยาศาสตร์เพิ่มเติม – สาระชีววิทยา - สาระเคมี - สาระฟิสิกส์
- สาระโลก ดาราศาสตร์ และอวกาศ

ภาพประกอบ 1 แผนภาพสาระการเรียนรู้แกนกลางการศึกษา กลุ่มสาระการเรียนรู้วิทยาศาสตร์ตาม หลักสูตรแกนกลางการศึกษาขั้นพื้นฐานพุทธศักราช 2551 (ฉบับปรับปรุง พ.ศ. 2560)

1. เป้าหมายของวิทยาศาสตร์

ในการเรียนการสอนวิทยาศาสตร์มุ่งเน้นให้ผู้เรียนได้ค้นพบความรู้ด้วยตนเองมากที่สุด เพื่อให้ได้ทั้งกระบวนการและความรู้จากวิธีการสังเกต การสำรวจตรวจสอบ การทดลอง แล้วนำผลที่ได้มาจัดระบบเป็นหลักการ แนวคิด และองค์ความรู้

การจัดการเรียนการสอนวิทยาศาสตร์จึงมีเป้าหมายที่สำคัญ ดังนี้

1. เพื่อให้เข้าใจหลักการ ทฤษฎี และกฎที่เป็นพื้นฐานในวิชาวิทยาศาสตร์

2. เพื่อให้เข้าใจขอบเขตของธรรมชาติของวิชาวิทยาศาสตร์และข้อจำกัดในการศึกษาวิชาวิทยาศาสตร์

3. เพื่อให้มีทักษะที่สำคัญในการศึกษาค้นคว้าและคิดค้นทางเทคโนโลยี

4. เพื่อให้ตระหนักถึงความสัมพันธ์ระหว่างวิชาวิทยาศาสตร์ เทคโนโลยี มวลมนุษย์ และสภาพแวดล้อมในเชิงที่มีอิทธิพลและผลกระทบซึ่งกันและกัน

5. เพื่อนำความรู้ ความเข้าใจในวิชาวิทยาศาสตร์ และเทคโนโลยีไปใช้ให้เกิดประโยชน์ต่อสังคมและการดำรงชีวิต

6. เพื่อพัฒนากระบวนการคิดและจินตนาการ ความสามารถในการแก้ปัญหา และการจัดการทักษะในการสื่อสาร และความสามารถในการตัดสินใจ

7. เพื่อให้เป็นผู้ที่มีจิตวิทยาศาสตร์ มีคุณธรรม จริยธรรม และค่านิยมในการใช้วิทยาศาสตร์และเทคโนโลยีอย่างสร้างสรรค์

2. เรียนรู้อะไรในวิทยาศาสตร์

กลุ่มสาระการเรียนรู้วิทยาศาสตร์มุ่งหวังให้ผู้เรียนได้เรียนรู้วิทยาศาสตร์ที่เน้นการเชื่อมโยงความรู้กับกระบวนการมีทักษะสำคัญในการค้นคว้าและสร้างองค์ความรู้ โดยใช้กระบวนการในการสืบเสาะหาความรู้ และแก้ปัญหาที่หลากหลายให้ผู้เรียนมีส่วนร่วมในการเรียนรู้ทุกขั้นตอน มีการทำกิจกรรมด้วยการลงมือปฏิบัติจริงอย่างหลากหลาย เหมาะสมกับระดับชั้น โดยกำหนดสาระสำคัญดังนี้

วิทยาศาสตร์ชีวภาพ เรียนรู้เกี่ยวกับชีวิตในสิ่งแวดล้อม องค์ประกอบของสิ่งมีชีวิต การดำรงชีวิตของมนุษย์และสัตว์ การดำรงชีวิตของพืช พันธุกรรม ความหลากหลายทางชีวภาพ และวิวัฒนาการของสิ่งมีชีวิต

วิทยาศาสตร์กายภาพ เรียนรู้เกี่ยวกับธรรมชาติของสาร การเปลี่ยนแปลงของสาร การเคลื่อนที่พลังงานและคลื่น

วิทยาศาสตร์โลกและอวกาศ เรียนรู้เกี่ยวกับองค์ประกอบของเอกภพ ปฏิสัมพันธ์ภายในระบบสุริยะ เทคโนโลยีอวกาศ ระบบโลก การเปลี่ยนแปลงทางธรณีวิทยา กระบวนการเปลี่ยนแปลงลมฟ้าอากาศ และผลต่อสิ่งมีชีวิตและสิ่งแวดล้อมเทคโนโลยี

การออกแบบและเทคโนโลยี เรียนรู้เกี่ยวกับเทคโนโลยีเพื่อการดำรงชีวิตในสังคมที่มีการเปลี่ยนแปลงอย่างรวดเร็ว ใช้ความรู้และทักษะทางด้านวิทยาศาสตร์ คณิตศาสตร์ และศาสตร์อื่น ๆ เพื่อแก้ปัญหาหรือพัฒนางานอย่างมีความคิดสร้างสรรค์ ด้วยกระบวนการออกแบบเชิงวิศวกรรมเลือกใช้เทคโนโลยีอย่างเหมาะสมโดยคำนึงถึงผลกระทบต่อชีวิต สังคม และสิ่งแวดล้อม

วิทยาการคำนวณ เรียนรู้เกี่ยวกับการคิดเชิงคำนวณ การคิดวิเคราะห์ แก้ปัญหาเป็นขั้นตอนและเป็นระบบ ประยุกต์ใช้ความรู้ด้านวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศและการสื่อสาร ในการแก้ปัญหาที่พบในชีวิตจริงได้อย่างมีประสิทธิภาพ

3. สาระและมาตรฐานการเรียนรู้

สาระที่ 1 วิทยาศาสตร์ชีวภาพ

มาตรฐาน ว 1.1 เข้าใจความหลากหลายของระบบนิเวศ ความสัมพันธ์ระหว่างสิ่งไม่มีชีวิตกับสิ่งมีชีวิต และความสัมพันธ์ระหว่างสิ่งมีชีวิตกับสิ่งมีชีวิตต่าง ๆ ในระบบนิเวศ การถ่ายทอดพลังงาน การเปลี่ยนแปลงแทนที่ในระบบนิเวศ ความหมายของประชากร ปัญหาและผลกระทบที่มีต่อทรัพยากรธรรมชาติและสิ่งแวดล้อม แนวทางในการอนุรักษ์ทรัพยากรธรรมชาติและการแก้ไขปัญหาสิ่งแวดล้อมรวมทั้งนำความรู้ไปใช้ประโยชน์

มาตรฐาน ว 1.2 เข้าใจสมบัติของสิ่งมีชีวิต หน่วยพื้นฐานของสิ่งมีชีวิต การลำเลียงสารเข้าและออกจากเซลล์ ความสัมพันธ์ของโครงสร้างและหน้าที่ของระบบต่าง ๆ ของสัตว์และมนุษย์ที่ทำงานสัมพันธ์กัน ความสัมพันธ์ของโครงสร้าง และหน้าที่ของอวัยวะต่าง ๆ ของพืชที่ทำงานสัมพันธ์กันรวมทั้งนำความรู้ไปใช้ประโยชน์

มาตรฐาน ว 1.3 เข้าใจกระบวนการและความสำคัญของการถ่ายทอดลักษณะทางพันธุกรรม สารพันธุกรรม การเปลี่ยนแปลงทางพันธุกรรมที่มีผลต่อสิ่งมีชีวิต ความหลากหลายทางชีวภาพและวิวัฒนาการของสิ่งมีชีวิต รวมทั้งนำความรู้ไปใช้ประโยชน์

สาระที่ 2 วิทยาศาสตร์กายภาพ

มาตรฐาน ว 2.1 เข้าใจสมบัติของสสาร องค์ประกอบของสสาร ความสัมพันธ์ระหว่างสมบัติของสสารกับโครงสร้างและแรงยึดเหนี่ยวระหว่างอนุภาค หลักและธรรมชาติของการเปลี่ยนแปลงสถานะของสสาร การเกิดสารละลาย และการเกิดปฏิกิริยาเคมี

มาตรฐาน ว 2.2 เข้าใจธรรมชาติของแรงในชีวิตประจำวัน ผลของแรงที่กระทำต่อวัตถุลักษณะการเคลื่อนที่แบบต่าง ๆ ของวัตถุ รวมทั้งนำความรู้ไปใช้ประโยชน์

มาตรฐาน ว 2.3 เข้าใจความหมายของพลังงาน การเปลี่ยนแปลงและการถ่ายโอนพลังงาน ปฏิสัมพันธ์ระหว่างสสารและพลังงาน พลังงานในชีวิตประจำวัน ธรรมชาติของคลื่น ปรากฏการณ์ที่เกี่ยวข้องกับเสียง แสง และคลื่นแม่เหล็กไฟฟ้า รวมทั้งนำความรู้ไปใช้ประโยชน์

สาระที่ 3 วิทยาศาสตร์โลก และอวกาศ

มาตรฐาน ว 3.1 เข้าใจองค์ประกอบ ลักษณะกระบวนการเกิด และวิวัฒนาการของเอกภพ กาแล็กซี ดาวฤกษ์ และระบบสุริยะ รวมทั้งปฏิสัมพันธ์ภายในระบบสุริยะที่ส่งผลต่อสิ่งมีชีวิต และการประยุกต์ใช้เทคโนโลยีอวกาศ

มาตรฐาน ว 3.2 เข้าใจองค์ประกอบและความสัมพันธ์ของระบบโลก กระบวนการเปลี่ยนแปลงภายในโลก และบนผิวโลก ธรณีพิบัติภัย กระบวนการเปลี่ยนแปลงลมฟ้าอากาศ และภูมิอากาศโลก รวมทั้งผลต่อสิ่งมีชีวิตและสิ่งแวดล้อม

สาระที่ 4 เทคโนโลยี

มาตรฐาน ว 4.1

เข้าใจแนวคิดหลักของเทคโนโลยีเพื่อการดำรงชีวิตในสังคมที่มีการเปลี่ยนแปลงอย่างรวดเร็ว ใช้ความรู้และทักษะทางด้านวิทยาศาสตร์ คณิตศาสตร์ และศาสตร์อื่น ๆ เพื่อแก้ปัญหาหรือพัฒนางานอย่างมีความคิดสร้างสรรค์ ด้วยกระบวนการออกแบบเชิงวิศวกรรม เลือกใช้เทคโนโลยีอย่างเหมาะสม โดยคำนึงถึงผลกระทบต่อชีวิต สังคม และสิ่งแวดล้อม

มาตรฐาน ว 4.2 เข้าใจและใช้แนวคิดเชิงคำนวณในการแก้ปัญหาที่พบในชีวิตจริงอย่างเป็นขั้นตอนและเป็นระบบ ใช้เทคโนโลยีสารสนเทศและการสื่อสารในการเรียนรู้การทำงาน และการแก้ปัญหาได้อย่างมีประสิทธิภาพ รู้เท่าทัน และมีจริยธรรม

4. คุณภาพผู้เรียนเมื่อจบชั้นมัธยมศึกษาปีที่ 3

4.1 เข้าใจลักษณะและองค์ประกอบที่สำคัญของเซลล์สิ่งมีชีวิต ความสัมพันธ์ของการทำงานของระบบต่าง ๆ ในร่างกายมนุษย์ การดำรงชีวิตของพืช การถ่ายทอดลักษณะทางพันธุกรรม การเปลี่ยนแปลงของยีนหรือโครโมโซม และตัวอย่างโรคที่เกิดจากการเปลี่ยนแปลงทางพันธุกรรม ประโยชน์และผลกระทบของสิ่งมีชีวิตดัดแปรพันธุกรรม ความหลากหลายทางชีวภาพ ปฏิสัมพันธ์ขององค์ประกอบของระบบนิเวศ และการถ่ายทอดพลังงานในสิ่งมีชีวิต

4.2 เข้าใจองค์ประกอบและสมบัติของธาตุ สารละลาย สารบริสุทธิ์ สารผสม หลักการแยกสาร การเปลี่ยนแปลงของสารในรูปแบบของการเปลี่ยนสถานะ การเกิดสารละลาย และการเกิดปฏิกิริยาเคมี และสมบัติทางกายภาพ และการใช้ประโยชน์ของวัสดุประเภทพอลิเมอร์ เซรามิกส์ และวัสดุผสม

4.3 เข้าใจการเคลื่อนที่ แรงลัพธ์และผลของแรงลัพธ์กระทำต่อวัตถุ โมเมนต์ของแรง แรงที่ปรากฏในชีวิตประจำวัน สนามของแรง ความสัมพันธ์ของงาน พลังงานจลน์ พลังงานศักย์โน้มถ่วง กฎการอนุรักษ์พลังงาน การถ่ายโอนพลังงาน สมดุลความร้อน ความสัมพันธ์ของปริมาณทางไฟฟ้า การต่อวงจรไฟฟ้าในบ้าน พลังงานไฟฟ้า และหลักการเบื้องต้นของวงจรอิเล็กทรอนิกส์

4.4 เข้าใจสมบัติของคลื่น และลักษณะของคลื่นแบบต่าง ๆ แสง การสะท้อน การหักเหของแสงและทัศนอุปกรณ์

4.5 เข้าใจการโคจรของดาวเคราะห์รอบดวงอาทิตย์ การเกิดฤดู การเคลื่อนที่ปรากฏของดวงอาทิตย์ การเกิดข้างขึ้นข้างแรม การขึ้นและตกของดวงจันทร์ การเกิดน้ำขึ้นน้ำลงประโยชน์ของเทคโนโลยีอวกาศ และความก้าวหน้าของโครงการสำรวจอวกาศ

4.6 เข้าใจลักษณะของชั้นบรรยากาศ องค์ประกอบและปัจจัยที่มีผลต่อ ลม ฟ้าอากาศการเกิดและผลกระทบของพายุฟ้าคะนอง พายุหมุนเขตร้อน การพยากรณ์อากาศ สถานการณ์การเปลี่ยนแปลงภูมิอากาศโลก กระบวนการเกิดเชื้อเพลิงซากดึกดำบรรพ์และการใช้ประโยชน์

พลังงานทดแทนและการใช้ประโยชน์ ลักษณะโครงสร้างภายในโลก กระบวนการเปลี่ยนแปลงทางธรณีวิทยาบนผิวโลก ลักษณะชั้นหน้าตัดดิน กระบวนการเกิดดิน แหล่งน้ำผิวดิน แหล่งน้ำใต้ดิน กระบวนการเกิด และผลกระทบของภัยธรรมชาติและธรณีพิบัติภัย

4.7 เข้าใจแนวคิดหลักของเทคโนโลยี ได้แก่ ระบบทางเทคโนโลยี การเปลี่ยนแปลงของเทคโนโลยี ความสัมพันธ์ระหว่างเทคโนโลยีกับศาสตร์อื่น โดยเฉพาะวิทยาศาสตร์ หรือคณิตศาสตร์ วิเคราะห์ เปรียบเทียบ และตัดสินใจเพื่อเลือกใช้เทคโนโลยี โดยคำนึงถึงผลกระทบต่อชีวิต สังคม และสิ่งแวดล้อม ประยุกต์ใช้ความรู้ ทักษะ และทรัพยากรเพื่อออกแบบและสร้างผลงานสำหรับการแก้ปัญหาในชีวิตประจำวันหรือการประกอบอาชีพ โดยใช้กระบวนการออกแบบเชิงวิศวกรรม รวมทั้งเลือกใช้วัสดุ อุปกรณ์ และเครื่องมือได้อย่างถูกต้อง เหมาะสม ปลอดภัย รวมทั้งคำนึงถึงทรัพย์สินทางปัญญา

4.8 นำข้อมูลปฐมภูมิเข้าสู่ระบบคอมพิวเตอร์ วิเคราะห์ ประเมิน นำเสนอข้อมูลและสารสนเทศได้ตามวัตถุประสงค์ ใช้ทักษะการคิดเชิงคำนวณในการแก้ปัญหาที่พบในชีวิตจริงและเขียนโปรแกรมอย่างง่ายเพื่อช่วยในการแก้ปัญหา ใช้เทคโนโลยีสารสนเทศ และการสื่อสารอย่างรู้เท่าทัน และรับผิดชอบต่อสังคม

4.9 ตั้งคำถามหรือกำหนดปัญหาที่เชื่อมโยงกับพยานหลักฐาน หรือหลักการทางวิทยาศาสตร์ที่มีการกำหนดและควบคุมตัวแปร คิดคาดคะเนคำตอบหลายแนวทาง สร้างสมมติฐานที่สามารถนำไปสู่การสำรวจตรวจสอบ ออกแบบและลงมือสำรวจตรวจสอบ โดยใช้วัสดุและเครื่องมือที่เหมาะสม เลือกใช้เครื่องมือและเทคโนโลยีสารสนเทศที่เหมาะสมในการเก็บรวบรวมข้อมูลทั้งในเชิงปริมาณและคุณภาพที่ได้ผลเที่ยงตรงและปลอดภัย

4.10 วิเคราะห์และประเมินความสอดคล้องของข้อมูลที่ได้จากการสำรวจตรวจสอบจากพยานหลักฐาน โดยใช้ความรู้และหลักการทางวิทยาศาสตร์ในการแปลความหมายและลงข้อสรุป และสื่อสารความคิด ความรู้ จากผลการสำรวจตรวจสอบหลากหลายรูปแบบ หรือใช้เทคโนโลยีสารสนเทศเพื่อให้ผู้อื่นเข้าใจได้อย่างเหมาะสม

4.11 แสดงถึงความสนใจ มุ่งมั่น รับผิดชอบ รอบคอบ และซื่อสัตย์ ในสิ่งที่จะเรียนรู้ มีความคิดสร้างสรรค์เกี่ยวกับเรื่องที่จะศึกษาตามความสนใจของตนเอง โดยใช้เครื่องมือและวิธีการที่ให้ได้ผลถูกต้อง เชื่อถือได้ ศึกษาค้นคว้าเพิ่มเติมจากแหล่งความรู้ต่าง ๆ แสดงความคิดเห็นของตนเอง รับฟังความคิดเห็นผู้อื่น และยอมรับการเปลี่ยนแปลงความรู้ที่ค้นพบ เมื่อมีข้อมูลและประจักษ์พยานใหม่เพิ่มขึ้นหรือโต้แย้งจากเดิม

4.12 ตระหนักในคุณค่าของความรู้วิทยาศาสตร์และเทคโนโลยีที่ใช้ในชีวิตประจำวัน ใช้ความรู้และกระบวนการทางวิทยาศาสตร์และเทคโนโลยีในการดำรงชีวิต และการประกอบอาชีพ แสดงความชื่นชม ยกย่อง และเคารพสิทธิในผลงานของผู้คิดค้น เข้าใจผลกระทบทั้งด้านบวกและ

ด้านลบของการพัฒนาทางวิทยาศาสตร์ต่อสิ่งแวดล้อมและต่อบริบทอื่น ๆ และศึกษาหาความรู้เพิ่มเติม ทำโครงงานหรือสร้างชิ้นงานตามความสนใจ

4.13 แสดงถึงความซาบซึ้ง ห่วงใย มีพฤติกรรมเกี่ยวกับการดูแลรักษาความสมดุลของระบบนิเวศ และความหลากหลายทางชีวภาพ

โครงการประเมินผลนักเรียนร่วมกับนานาชาติ

โครงการประเมินผลนักเรียนร่วมกับนานาชาติ (Programme for International Student Assessment หรือ PISA) ริเริ่มโดยองค์การเพื่อความร่วมมือทางเศรษฐกิจและการพัฒนา (Organization for Economic Co-operation and Development หรือ OECD) มีวัตถุประสงค์เพื่อประเมินคุณภาพของระบบการศึกษาทั้งประเทศสมาชิก OECD และประเทศที่ไม่ใช่สมาชิก OECD (Partner Countries) ในการเตรียมความพร้อมให้เยาวชนมีศักยภาพสำหรับการแข่งขันในอนาคต PISA ใช้ความร่วมมือและความเชี่ยวชาญของนานาชาติในการวางกรอบโครงสร้างการประเมิน การสร้างเครื่องมือ และการศึกษาวิจัย เพื่อประกันคุณภาพของการศึกษาวิจัยให้สามารถเปรียบเทียบกันได้ในระดับนานาชาติ และข้อมูลที่ได้สามารถชี้บอกถึงคุณภาพการศึกษาของประเทศต่าง ๆ ได้ ปัจจุบันมีประเทศเข้าร่วมโครงการ PISA มากกว่า 70 ประเทศทั่วโลก

โครงการ PISA ประเมินผลนักเรียนในระดับโรงเรียนและดำเนินการอย่างต่อเนื่องทุก ๆ 3 ปี กลุ่มตัวอย่างของ PISA คือ นักเรียนกลุ่มอายุ 15 ปี ซึ่งโดยสากลถือว่าเป็นวัยจบการศึกษาภาคบังคับ การประเมินของ PISA เน้นการประเมินความสามารถของนักเรียนในการใช้ความรู้และทักษะเพื่อแก้ปัญหาในชีวิตจริงหรือที่เรียกว่า “การรู้เรื่อง” (Literacy) ในสามด้าน ได้แก่ การรู้เรื่องการอ่าน (Reading Literacy) การรู้เรื่องคณิตศาสตร์ (Mathematical Literacy) และการรู้เรื่องวิทยาศาสตร์ (Scientific Literacy) ซึ่งการรู้เรื่องทั้งสามด้านนี้ถือว่าเป็นสิ่งจำเป็นสำหรับการเรียนรู้ตลอดชีวิต และเป็นสิ่งที่ประชากรจำเป็นต้องมีเพื่อการพัฒนาและการแข่งขันทางเศรษฐกิจของประเทศ

สำหรับประเทศไทย สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยีทำหน้าที่เป็นศูนย์แห่งชาติดำเนินการวิจัยโครงการ PISA โดยได้เข้าร่วมโครงการตั้งแต่ พ.ศ. 2543 หรือใน PISA 2000 ซึ่งเป็นรอบแรกของการประเมินและได้เข้าร่วมโครงการอย่างต่อเนื่องใน PISA 2003 PISA 2006 PISA 2009 PISA 2012 และ PISA 2015 ในแต่ละรอบการประเมิน PISA จะสอบสามวิชาพื้นฐานดังกล่าว โดยมีสัดส่วนของข้อสอบแต่ละวิชาแตกต่างกัน ขึ้นอยู่กับว่าในรอบการประเมินนั้นเน้นการประเมินวิชาใด หากเป็นวิชาหลักในรอบการประเมินจะมีสัดส่วนของข้อสอบประมาณ 60% ส่วนวิชารองจะมีสัดส่วนของข้อสอบประมาณ 20% สำหรับรอบการประเมิน PISA 2015 เน้นการประเมินวิชาวิทยาศาสตร์ จึงมีสัดส่วนของข้อสอบวิทยาศาสตร์ ประมาณ 60% ส่วนวิชาคณิตศาสตร์

และการอ่าน มีสัดส่วนของข้อสอบประมาณวิชาละ 20% ในบางรอบการประเมินนอกจากสามวิชาพื้นฐานแล้วยังอาจมีการประเมินสมรรถนะอื่นเพิ่มเติม เช่น PISA 2015 มีการประเมินเพิ่มเติมด้านการแก้ปัญหาแบบร่วมมือ (Collaborative Problem Solving หรือ CPS) ทั้งนี้การประเมินเพิ่มเติมด้านใดจะเป็นไปตามที่ OECD กำหนด (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2560 ก)

การรู้เรื่องวิทยาศาสตร์

ความรู้และความเข้าใจทางวิทยาศาสตร์ และเทคโนโลยีเป็นเครื่องมือสำคัญอย่างยิ่งในการเตรียมเยาวชนให้สามารถดำเนินชีวิตในโลก ปัจจุบันที่มีวิทยาศาสตร์และเทคโนโลยีเป็นพื้นฐานและส่งผลกระทบต่อทุกชีวิตในทุกระดับ ทั้งตัวบุคคลในอาชีพการงาน และในสังคมวัฒนธรรมของทุก ๆ ชีวิต ทำให้บุคคลสามารถรับรู้และตัดสินประเด็นปัญหาของสังคมที่เกิดจากผลกระทบของวิทยาศาสตร์และเทคโนโลยีอย่างมีความรู้ ความเข้าใจ มีส่วนร่วมในสังคม ระดับชุมชน ระดับประเทศและระดับโลก อย่างเต็มภาคภูมิ

เป้าหมายของการศึกษาวิทยาศาสตร์ คือ การทำให้นักเรียนทุกคนมีการรู้เรื่องวิทยาศาสตร์ (Scientific Literacy) ซึ่งรวมถึงความรู้มิติต่าง ๆ ทางวิทยาศาสตร์ ได้แก่ ความรู้ความสามารถทางสติปัญญา กระบวนการแสวงหาความรู้ทางวิทยาศาสตร์ และความสัมพันธ์ระหว่างวิทยาศาสตร์กับเทคโนโลยีด้วย ในการประเมินการรู้เรื่องวิทยาศาสตร์ PISA ต้องการหาตัวชี้วัดว่านักเรียนเรียนรู้ทฤษฎีและแนวคิดพื้นฐานทางวิทยาศาสตร์มากน้อยเพียงใด มีสมรรถนะทางวิทยาศาสตร์ และสามารถแก้ปัญหาในชีวิตจริงที่เกี่ยวข้องกับวิทยาศาสตร์และเทคโนโลยีได้ดีเพียงใด ทั้งนี้เพื่อหาคำตอบให้กับระดับนโยบายและระดับปฏิบัติในการปรับปรุงการเรียนการสอนวิทยาศาสตร์

ประเด็นหลักที่ PISA ให้ความสำคัญในการประเมินผลวิทยาศาสตร์ คือ การประเมินผลต้องให้ความชัดเจนกับสมรรถนะที่เหมาะสมกับนักเรียนวัย 15 ปีว่า ควรรู้อะไร ทำอะไรได้ และให้คุณค่ากับอะไรบ้าง ภายในขอบเขตของบริบทส่วนตัว บริบทของสังคม และบริบทของโลก ซึ่งประเด็นนี้อาจแตกต่างจากการเรียนการสอนและการประเมินผลตามปกติในโรงเรียน

1. นิยามของการรู้เรื่องวิทยาศาสตร์

PISA ให้นิยาม “การรู้เรื่องวิทยาศาสตร์” ไว้ว่า การรู้เรื่องวิทยาศาสตร์ (Scientific Literacy) หมายถึง ความสามารถในการเชื่อมโยงสิ่งต่าง ๆ เข้ากับประเด็นที่เกี่ยวข้องกับวิทยาศาสตร์และแนวคิดทางวิทยาศาสตร์ได้อย่างไตร่ตรอง

บุคคลที่รู้เรื่องวิทยาศาสตร์ (Scientifically Literate Person) จะสื่อสารพูดคุยในประเด็นที่เกี่ยวข้องกับวิทยาศาสตร์และเทคโนโลยีอย่างเป็นเหตุเป็นผล ซึ่งจำเป็นต้องใช้สมรรถนะดังต่อไปนี้

1. การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ หมายถึง มีความสามารถในการรับรู้ เสนอ และประเมินคำอธิบายที่เกี่ยวข้องกับปรากฏการณ์ทางธรรมชาติและเทคโนโลยี

2. การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ หมายถึง มีความสามารถในการอธิบายและประเมินคุณค่าของการสำรวจตรวจสอบทางวิทยาศาสตร์ และนำเสนอแนวทางในการตอบคำถามอย่างเป็นวิทยาศาสตร์

3. การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์ หมายถึง มีความสามารถในการวิเคราะห์และประเมินข้อมูล คำกล่าวอ้าง และข้อโต้แย้งในหลากหลายรูปแบบ และลงข้อสรุปทางวิทยาศาสตร์ได้อย่างเหมาะสม

2. กรอบโครงสร้างการประเมินการรู้เรื่องวิทยาศาสตร์

ตามวัตถุประสงค์ของการประเมิน PISA 2015 จึงได้กำหนดกรอบโครงสร้างการประเมินผลการรู้เรื่องวิทยาศาสตร์ ซึ่งประกอบด้วยสี่องค์ประกอบที่เกี่ยวข้องกัน ได้แก่

1. บริบท หมายถึง การรับรู้ถึงสถานการณ์ในชีวิต ในระดับส่วนตัว ระดับชาติ และระดับโลก ทั้งที่เป็นเรื่องในปัจจุบัน หรือในอดีตที่ผ่านมา ซึ่งจำเป็นต้องมีความเข้าใจเรื่องวิทยาศาสตร์และเทคโนโลยี

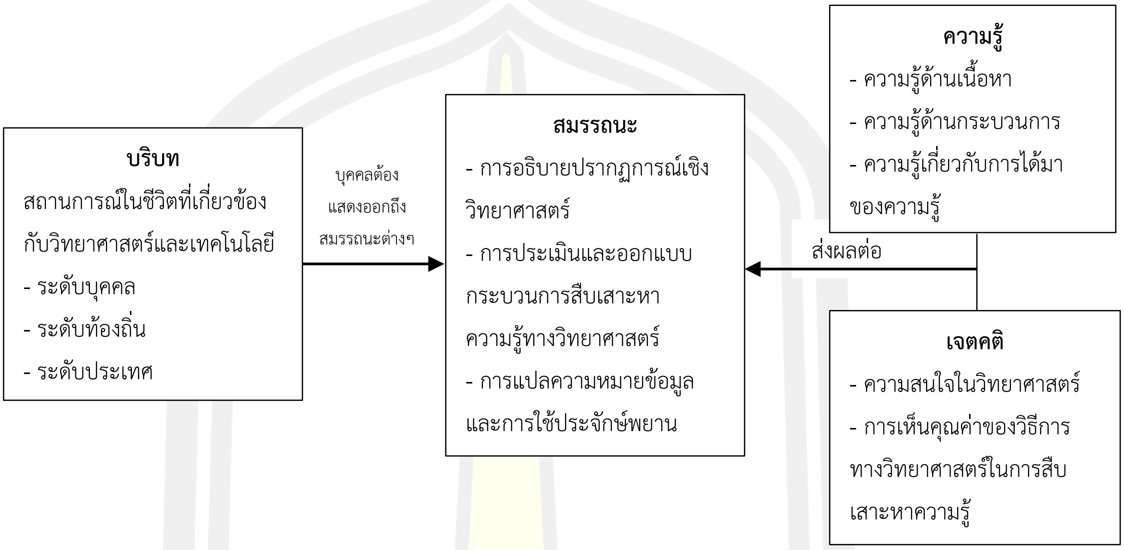
2. สมรรถนะทางวิทยาศาสตร์ หมายถึง ความสามารถในการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ และการแปลความหมายข้อมูลและใช้ประจักษ์พยานเชิงวิทยาศาสตร์

3. ความรู้ทางวิทยาศาสตร์ หมายถึง ความเข้าใจในข้อเท็จจริง แนวคิดหลักและทฤษฎี สำคัญที่ทำให้เกิดความรู้พื้นฐานทางวิทยาศาสตร์ ความรู้ ประกอบด้วย ความรู้เกี่ยวกับธรรมชาติของโลกและสิ่งประดิษฐ์ทางเทคโนโลยี (ความรู้ด้านเนื้อหา) ความรู้เกี่ยวกับวิธีการในการสร้างแนวคิดต่าง ๆ (ความรู้ด้านกระบวนการ) และความเข้าใจในเหตุผลพื้นฐานของกระบวนการสร้างความรู้ (ความรู้เกี่ยวกับการได้มาของความรู้)

4. เจตคติต่อวิทยาศาสตร์ หมายถึง การแสดงการตอบสนองต่อวิทยาศาสตร์และเทคโนโลยีด้วยความสนใจ ให้ความสำคัญกับกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ และรับรู้และตระหนักถึงปัญหาสิ่งแวดล้อม

องค์ประกอบทั้งสี่มีความสัมพันธ์กัน กล่าวคือ ในการดำเนินชีวิตคนเราต้องเผชิญสถานการณ์ที่หลากหลายในชีวิตจริงที่เกี่ยวข้องกับทั้งตนเอง ท้องถิ่น ประเทศ หรือสถานการณ์ของโลก เราจึงต้องมีและใช้สมรรถนะเพื่อตอบสนองและแก้ปัญหาต่าง ๆ ได้อย่างสมเหตุสมผล

ซึ่งการตอบสนองจะทำได้ดีเพียงใดนั้นขึ้นอยู่กับความรู้และเจตคติต่าง ๆ ที่แต่ละคนมีอยู่ ดังความสัมพันธ์ที่แสดงในรูปต่อไปนี้



ภาพประกอบ 2 กรอบการประเมินการรู้เรื่องวิทยาศาสตร์ของ PISA 2015

2.1 สถานการณ์ และบริบทของวิทยาศาสตร์

สิ่งหนึ่งที่ PISA ให้ความสำคัญในการประเมิน คือ การใช้วิทยาศาสตร์ในสถานการณ์ต่าง ๆ อย่างหลากหลายในการจัดการกับประเด็นทางวิทยาศาสตร์ การเลือกวิธีการที่ใช้มักจะขึ้นอยู่กับสถานการณ์ของประเด็นปัญหานั้น ปัญหาแบบเดียวกันแต่ถ้าอยู่ในสถานการณ์ที่ต่างกัน วิธีการที่เลือกใช้ก็จะต่างกัน ดังนั้นในการสร้างข้อสอบจึงมีการจัดสถานการณ์ หรือจำกัดบริบทของภารกิจในการประเมินข้อคำถามของ PISA จะเป็นการทดสอบความรู้ความเข้าใจในแนวคิดหลักการทางวิทยาศาสตร์ที่ได้เรียนรู้มาจากหลักสูตร เพื่อนำมาใช้ในการตอบคำถามเรื่องวิทยาศาสตร์ที่เกิดขึ้นในชีวิตจริง เช่น เกิดกับตัวเอง ครอบครัว หรือเพื่อน (บริบทส่วนตัว) ประเด็นที่ส่งผลกระทบต่อสังคม วัฒนธรรม สุขภาพ หรือชีวิตมนุษย์ (บริบทสังคม) ประเด็นทางวิทยาศาสตร์ที่เป็นข่าวในสื่อหรือมีผลกระทบสืบเนื่องถึงสังคมโลกหรือต่อโลกอนาคต (บริบทโลก) เป็นต้น

คำถามของการประเมินผล PISA จึงอยู่ในสถานการณ์ที่เป็นส่วนหนึ่งในโลกชีวิตจริงของนักเรียน และไม่จำกัดอยู่เฉพาะสถานการณ์ในโรงเรียนเท่านั้น แต่จะเป็นสถานการณ์ที่อาจเกี่ยวข้องกับตัวเอง ครอบครัว ชุมชน หรือสถานการณ์ของโลกก็ได้ หรือแม้กระทั่งคำถามที่อยู่ในบริบทประวัติศาสตร์ก็สามารถนำมาใช้ประเมินความเข้าใจเกี่ยวกับกระบวนการและความก้าวหน้าของความรู้ทางวิทยาศาสตร์ได้

ตาราง 1 ระดับของบริบทสำหรับการประเมินการรู้เรื่องวิทยาศาสตร์ของ PISA

บริบท	ระดับส่วนตัว (ตัวเอง ครอบครัว เพื่อน)	ระดับท้องถิ่น/ระดับชาติ (ชุมชน สังคม)	ระดับโลก (ชีวิตทั่วโลก)
สุขภาพและโรคภัย	การดูแลรักษาสุขภาพ อุบัติเหตุ โภชนาการ	การควบคุมโรค แพร่เชื้อในสังคม การเลือกอาหาร สุขภาพชุมชน	โรคระบาด การระบาดข้ามประเทศ
ทรัพยากรธรรมชาติ	การใช้วัสดุ และพลังงาน	การรักษาจำนวนประชากรให้คงที่ คุณภาพชีวิต ความมั่นคง การผลิตและการกระจายอาหาร การจัดหาพลังงาน	แหล่งทรัพยากรที่เกิดใหม่ได้และแหล่งทรัพยากรที่เกิดใหม่ไม่ได้ การเพิ่มจำนวนประชากร การใช้ประโยชน์จากสิ่งมีชีวิตชนิดต่าง ๆ อย่างยั่งยืน
คุณภาพสิ่งแวดล้อม	พฤติกรรมเป็นมิตรกับสิ่งแวดล้อม การใช้และการกำจัดวัสดุและอุปกรณ์	การกระจายของประชากร การกำจัดขยะ ผลกระทบต่อสิ่งแวดล้อม	ความหลากหลายทางชีววิทยา ความยั่งยืนของระบบนิเวศ การควบคุมมลพิษ การเกิดและการสูญเสียผิวดิน/ชีวมวล
อันตราย	การประเมินความเสี่ยงภัยจากทางเลือกการดำเนินชีวิต	การเปลี่ยนแปลงกะทันหัน (แผ่นดินไหว สภาพอากาศเลวร้าย) การเปลี่ยนแปลงอย่างช้าๆ และต่อเนื่อง (การกัดเซาะชายฝั่ง การตกตะกอน) การประเมินความเสี่ยง	การเปลี่ยนแปลงสภาพภูมิอากาศ ผลกระทบจากการสื่อสารสมัยใหม่

ตาราง 1 (ต่อ)

บริบท	ระดับส่วนตัว (ตัวเอง ครอบครัว เพื่อน)	ระดับท้องถิ่น/ ระดับชาติ (ชุมชน สังคม)	ระดับโลก (ชีวิตทั่วโลก)
ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี	แง่มุมทางวิทยาศาสตร์เกี่ยวกับงานอดิเรก เทคโนโลยีที่ใช้ส่วนบุคคล กิจกรรมทางดนตรีและกีฬา	วัสดุ เครื่องมือและกระบวนการใหม่ การดัดแปลงพันธุกรรม เทคโนโลยีเกี่ยวกับสุขภาพ การคมนาคมขนส่ง	การสูญพันธุ์ของสิ่งมีชีวิต การสำรวจอวกาศ การเกิดและโครงสร้างของจักรวาล

2.2 สมรรถนะทางวิทยาศาสตร์

PISA ประเมินการรู้เรื่องวิทยาศาสตร์โดยให้ความสำคัญเป็นพิเศษกับสมรรถนะทางวิทยาศาสตร์ (Scientific Competencies) และนิยามการประเมินสมรรถนะทางวิทยาศาสตร์ว่าเป็นการประเมินความสามารถของนักเรียนในการทำสิ่งต่อไปนี้

1. การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ (Explain Phenomena Scientifically)

2. การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ (Evaluate and Design Scientific Enquiry)

3. การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์ (Interpret Data and Evidence Scientifically)

2.2.1 การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์

การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์เป็นสมรรถนะที่จำเป็นสำหรับการรู้เรื่องวิทยาศาสตร์ การแสดงออกถึงสมรรถนะนี้บุคคลที่รู้เรื่องวิทยาศาสตร์ต้องสามารถระลึกถึงความรู้ด้านเนื้อหาที่เหมาะสมในสถานการณ์ที่กำหนดให้ และใช้ความรู้เพื่อแปลความหมายและให้คำอธิบายต่อปรากฏการณ์ต่าง ๆ สมรรถนะนี้รวมถึงการวาดแบบจำลองทางวิทยาศาสตร์เพื่อใช้อธิบายปรากฏการณ์ที่เกิดในชีวิตประจำวัน การบรรยายและการตีความปรากฏการณ์ การคาดการณ์หรือการพยากรณ์การเปลี่ยนแปลงที่อาจจะเกิดขึ้น รวมถึงการให้นักเรียนระบุว่าคำบรรยาย คำอธิบายใดสมเหตุสมผลหรือไม่ อย่างไร คำคาดการณ์จะเป็นไปได้หรือไม่ ด้วยเหตุผลอะไร เป็นต้น

โดยสรุปแล้ว สมรรถนะการอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ หมายถึง การมีความสามารถในการรับรู้ เสนอ และประเมินคำอธิบายที่เกี่ยวข้องกับปรากฏการณ์ทางธรรมชาติ และเทคโนโลยี โดยสามารถทำสิ่งต่อไปนี้

1. นำความรู้ทางวิทยาศาสตร์มาใช้สร้างคำอธิบายที่สมเหตุสมผล
2. ระบุ ใช้ และสร้างแบบจำลองและการนำเสนอข้อมูลเพื่อใช้ในการอธิบาย
3. เสนอสมมติฐานเพื่อใช้ในการอธิบาย
4. พยากรณ์การเปลี่ยนแปลงในเชิงวิทยาศาสตร์และให้เหตุผลที่สมเหตุสมผล
5. อธิบายถึงศักยภาพของความรู้ทางวิทยาศาสตร์ที่สามารถนำไปใช้เพื่อสังคม

2.2.2 การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์

บุคคลที่รู้เรื่องวิทยาศาสตร์ต้องมีความสามารถในการประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ เพื่อใช้ในการสร้างความรู้ที่เชื่อถือได้เกี่ยวกับโลกธรรมชาติ การแสดงออกถึงสมรรถนะด้านนี้ บุคคลต้องสามารถประเมินข้อค้นพบทางวิทยาศาสตร์อย่างมีวิจารณญาณ แยกแยะคำถามทางวิทยาศาสตร์ว่าคำถามใดสามารถตอบได้ด้วยการสำรวจตรวจสอบทางวิทยาศาสตร์ สมรรถนะนี้จำเป็นต้องใช้ความรู้เกี่ยวกับลักษณะสำคัญของการสำรวจตรวจสอบทางวิทยาศาสตร์ เช่น การทดสอบที่เที่ยงตรงต้องทำอย่างไร ต้องเปรียบเทียบอะไร ควบคุมตัวแปรใด และเปลี่ยนแปลงตัวแปรใด ต้องค้นคว้าสาระ และข้อมูลอะไรเพิ่มเติมอีก และต้องทำอะไรอย่างไรจึงจะเก็บข้อมูลที่ต้องการได้ นอกจากนี้ ยังต้องรู้ถึงความสำคัญและคุณค่าของงานวิจัยที่ผ่านมาที่ส่งผลต่อการค้นคว้าทางวิทยาศาสตร์เรื่องอื่น ๆ ต่อไป รวมถึงการเข้าใจถึงความสำคัญของการตั้งข้อสงสัยในการรายงานของสื่อ เกี่ยวกับเรื่องวิทยาศาสตร์ว่า ข้อค้นพบจากงานวิจัยต่าง ๆ อาจมีความคลุมเครือ ไม่แน่นอน หรือมีความลำเอียงได้ เป็นต้น

โดยสรุปแล้ว สมรรถนะการประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ หมายถึง การมีความสามารถในการอธิบายและประเมินคุณค่าของการสำรวจตรวจสอบทางวิทยาศาสตร์ และนำเสนอแนวทางในการตอบคำถามอย่างเป็นวิทยาศาสตร์ โดยสามารถทำสิ่งต่อไปนี้

1. ระบุปัญหาที่ต้องการสำรวจตรวจสอบจากการศึกษาทางวิทยาศาสตร์ที่กำหนดให้
2. แยกแยะได้ว่าประเด็นปัญหาหรือคำถามใดสามารถตรวจสอบได้ด้วยวิธีการทางวิทยาศาสตร์

3. เสนอวิธีสำรวจตรวจสอบปัญหาทางวิทยาศาสตร์ที่กำหนดให้

4. ประเมินวิธีสำรวจตรวจสอบปัญหาทางวิทยาศาสตร์ที่กำหนดให้

5. บรรยายและประเมินวิธีการต่าง ๆ ที่นักวิทยาศาสตร์ใช้ในการยืนยันถึงความน่าเชื่อถือของข้อมูล และความเป็นกลางและการสรุปอ้างอิงจากคำอธิบาย

2.2.3 การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์

บุคคลที่มีสมรรถนะการแปลความหมายข้อมูลและใช้ประจักษ์พยานเชิงวิทยาศาสตร์ ต้องแสดงออกถึงความสามารถในการตีความข้อมูลและหลักฐานทางวิทยาศาสตร์ที่ใช้ในการสร้างคำกล่าวอ้างหรือลงข้อสรุปนำเสนอข้อมูลที่ได้รับในรูปแบบอื่น เช่น ใช้คำพูดของตนเอง แผนภาพ หรือการแสดงแทนอื่น ๆ ได้ ซึ่งสมรรถนะนี้จำเป็นต้องใช้เครื่องมือทางคณิตศาสตร์ในการวิเคราะห์หรือสรุปข้อมูล และใช้ความสามารถในการใช้วิธีการพื้นฐานในการแปลงข้อมูลเป็นการแสดงแทนในรูปแบบอื่น ๆ นอกจากนี้ ยังต้องสร้างข้อสรุปที่สมเหตุสมผลบนพื้นฐานของประจักษ์พยานข้อมูล หรือประเมินข้อสรุปที่ผู้อื่นสร้างขึ้นว่าสอดคล้องกับประจักษ์พยานที่มีหรือไม่ รวมถึงสามารถให้เหตุผลสนับสนุนหรือโต้แย้งข้อสรุปอย่างสมเหตุสมผล

สำหรับนักเรียนวัย 15 ปี สมรรถนะการแปลความหมายข้อมูลและใช้ประจักษ์พยานเชิงวิทยาศาสตร์ หมายถึง การที่บุคคลต้องมีความสามารถในการวิเคราะห์และประเมิน ข้อมูล คำกล่าวอ้าง และข้อโต้แย้งและลงข้อสรุปทางวิทยาศาสตร์ได้อย่างเหมาะสม โดยสามารถทำสิ่งต่อไปนี้

1. แปลงข้อมูลที่นำเสนอในรูปแบบหนึ่งไปสู่รูปแบบอื่น

2. วิเคราะห์และแปลความหมายข้อมูลทางวิทยาศาสตร์ และลงข้อสรุป

3. ระบุข้อสันนิษฐาน ประจักษ์พยาน และเหตุผล ในเรื่องที่เกี่ยวกับวิทยาศาสตร์

4. แยกแยะระหว่างข้อโต้แย้งที่มาจากประจักษ์พยานและทฤษฎีทางวิทยาศาสตร์กับที่มาจากการพิจารณาจากสิ่งอื่น

5. ประเมินข้อโต้แย้งทางวิทยาศาสตร์และประจักษ์พยานจากแหล่งที่มาที่หลากหลาย (เช่น หนังสือพิมพ์ อินเทอร์เน็ต และวารสาร)

2.3 ความรู้ทางวิทยาศาสตร์: แนวคิดและเนื้อหาที่ครอบคลุม

การประเมินการรู้เรื่องวิทยาศาสตร์ของ PISA 2015 มีส่วนของความรู้ทางวิทยาศาสตร์แตกต่างจากการประเมินในรอบที่ผ่านมา การประเมินความรู้ทางวิทยาศาสตร์ (Scientific Knowledge) ที่ PISA 2015 กำหนดไว้นั้นครอบคลุมความรู้ทางวิทยาศาสตร์สามด้าน ได้แก่

1. ความรู้ด้านเนื้อหา (Content Knowledge)

2. ความรู้ด้านกระบวนการ (Procedural Knowledge)

3. ความรู้เกี่ยวกับการได้มาของความรู้ (Epistemic Knowledge)

2.3.1 ความรู้ด้านเนื้อหา (Content Knowledge)

เป็นความรู้เกี่ยวกับข้อเท็จจริง แนวความคิดหลัก แนวคิด และทฤษฎีเกี่ยวกับโลกธรรมชาติ โดย PISA เลือกประเมินความรู้ในสาขาวิชาหลัก ได้แก่ ฟิสิกส์ เคมี ชีววิทยา และวิทยาศาสตร์ โลกและอวกาศ ทั้งนี้มีเกณฑ์การเลือกแนวคิดทางวิทยาศาสตร์ที่ใช้ในการประเมิน ดังนี้

1. เป็นเรื่องที่เกี่ยวข้องกับสถานการณ์ในชีวิตจริง

2. แสดงให้เห็นถึงแนวความคิดหลักทางวิทยาศาสตร์ หรือทฤษฎีที่สำคัญซึ่งใช้ได้อย่างยาวนานและ

3. เหมาะสมกับระดับพัฒนาการของนักเรียนอายุ 15 ปี

ความรู้ด้านเนื้อหาที่ PISA ประเมินนั้นครอบคลุมความรู้เกี่ยวกับโลกธรรมชาติ ได้แก่

ระบบทางกายภาพ (Physical Systems) ใช้ความรู้เกี่ยวกับ :

1. โครงสร้างของสสาร (เช่น แบบจำลองอนุภาค และพันธะ)

2. สมบัติของสสาร (เช่น การเปลี่ยนสถานะ การนำความร้อนและการนำไฟฟ้า)

3. การเปลี่ยนแปลงทางเคมี (เช่น ปฏิกิริยาเคมี การถ่ายโอนพลังงานและกรด/เบส)

4. การเคลื่อนที่และแรง (เช่น ความเร็ว และความเสียดทาน) แรงที่เกิดขึ้นเมื่อวัตถุอยู่ห่างกัน (เช่น แรงแม่เหล็ก แรงโน้มถ่วง และแรงไฟฟ้าสถิตย์)

5. พลังงานและการเปลี่ยนรูปพลังงาน (เช่น การอนุรักษ์พลังงานการสูญเสียพลังงาน และปฏิกิริยาเคมี)

6. การปฏิสัมพันธ์ระหว่างพลังงานและสสาร (เช่น คลื่นแสงและคลื่นวิทยุ และคลื่นเสียง และคลื่นแผ่นดินไหว)

ระบบสิ่งมีชีวิต (Living Systems) ใช้ความรู้เกี่ยวกับ :

1. เซลล์ (เช่น โครงสร้างและหน้าที่ DNA และพืชและสัตว์)

2. แนวความคิดเรื่องสิ่งมีชีวิต (เช่น สิ่งมีชีวิตเซลล์เดียว และสิ่งมีชีวิตหลายเซลล์)

3. มนุษย์ (เช่น สุขภาพ โภชนาการ ระบบต่าง ๆ ในร่างกาย ซึ่งรวมทั้งการย่อย การหายใจ การหมุนเวียนเลือด การขับถ่าย การสืบพันธุ์ และความสัมพันธ์ของระบบต่าง ๆ)

4. ประชากร (เช่น สายพันธุ์ การวิวัฒนาการ ความหลากหลายทางชีววิทยา และความแปรผันทางพันธุกรรม)

5. ระบบนิเวศ (เช่น โซ่อาหาร การถ่ายทอดสารและพลังงาน)

6. ไบโอสเฟีย (เช่น ประโยชน์ที่ได้รับจากระบบนิเวศ และความยั่งยืนของระบบนิเวศ)

ระบบของโลกและอวกาศ (Earth and Space Systems) ใช้ความรู้เกี่ยวกับ :

1. โครงสร้างของโลกทั้งระบบ (เช่น พื้นดิน พื้นน้ำ และบรรยากาศ)

2. พลังงานในระบบโลก (เช่น แหล่งพลังงาน และภูมิอากาศของโลก)

3. การเปลี่ยนแปลงในระบบโลก (เช่น การเคลื่อนที่ของแผ่นธรณีภาค วัฏจักรธรณี เคมีและแรงดึงและแรงอัด)

4. ประวัติศาสตร์ของโลก (เช่น ฟอสซิล และกำเนิดและวิวัฒนาการของโลก)

5. โลกในอวกาศ (เช่น ความโน้มถ่วง ระบบสุริยะ และกาแล็กซี)

6. ประวัติศาสตร์และขนาดของจักรวาล (เช่น ปีแสง และทฤษฎีบิกแบง)

2.3.2 ความรู้ด้านกระบวนการ

เป็นความรู้เกี่ยวกับกระบวนการที่นักวิทยาศาสตร์ใช้ในการสร้างความรู้วิทยาศาสตร์และเป็นความรู้ในเรื่องการปฏิบัติและแนวความคิดเกี่ยวกับการสืบเสาะหาความรู้ เช่น การตรวจสอบซ้ำเพื่อลดความผิดพลาดและความไม่แน่นอน การควบคุมตัวแปร และการมีกระบวนการมาตรฐานเพื่อนำเสนอและสื่อสารข้อมูล

ลักษณะทั่วไปของความรู้ด้านกระบวนการที่จะทดสอบนักเรียน เช่น

1. แนวคิดเรื่องตัวแปร ได้แก่ ตัวแปรต้น ตัวแปรตาม และตัวแปรควบคุม

2. แนวคิดเรื่องการวัด เช่น การวัดเชิงปริมาณ การวัดเชิงคุณภาพ การวัดตัวแปรต่อเนื่องและไม่ต่อเนื่อง และการใช้มาตรวัด

3. วิธีการประเมินและลดข้อผิดพลาด เช่น การทำซ้ำ และการเฉลี่ยผลจากการวัด

4. กลไกที่ทำให้เกิดความน่าเชื่อถือในการทำซ้ำและความถูกต้องของข้อมูล

5. การสรุปและนำเสนอข้อมูลโดยใช้ตาราง กราฟ และแผนภูมิได้เหมาะสม

6. วิธีการกำหนดและควบคุมตัวแปร และบทบาทของตัวแปรในการออกแบบการทดลอง

7. ลักษณะของการออกแบบที่เหมาะสมเพื่อตอบคำถามทางวิทยาศาสตร์ เช่น การทดลอง การสำรวจตรวจสอบในภาคสนาม หรือการสืบค้นข้อสนเทศจากแหล่งต่าง ๆ

2.3.3 ความรู้เกี่ยวกับการได้มาของความรู้ (Epistemic Knowledge)

เป็นความรู้เกี่ยวกับบทบาทและลักษณะที่จำเป็นต่อกระบวนการสร้างความรู้ทางวิทยาศาสตร์ รวมถึงความเข้าใจบทบาทและหน้าที่ของสิ่งต่าง ๆ ที่มีต่อวิทยาศาสตร์ เช่น คำถาม การสังเกต ทฤษฎี สมมติฐาน แบบจำลอง การอภิปรายโต้แย้ง การยอมรับรูปแบบที่หลากหลายในกระบวนการสืบเสาะหาความรู้ และบทบาทในการตรวจสอบจากผู้อื่นที่ทำให้ความรู้ที่สร้างขึ้นนั้นน่าเชื่อถือ

ความรู้เกี่ยวกับการได้มาของความรู้ที่จำเป็นต่อการรู้เรื่องวิทยาศาสตร์มีลักษณะสำคัญดังนี้

1. การสร้างและการระบุลักษณะของวิทยาศาสตร์ ซึ่งหมายรวมถึง

1.1 ธรรมชาติของการสังเกต ข้อเท็จจริง สมมติฐาน แบบจำลอง และทฤษฎีทางวิทยาศาสตร์

1.2 วัตถุประสงค์และเป้าหมายของวิทยาศาสตร์เพื่อสร้างคำอธิบายธรรมชาติของโลก ซึ่งต่างจากวัตถุประสงค์และเป้าหมายของเทคโนโลยี ซึ่งเป็นการสร้างทางแก้ปัญหาที่ตรงตามความต้องการของมนุษย์ให้มากที่สุด จึงต้องพิจารณาถึงคำถามและข้อมูลที่เหมาะสมสำหรับวิทยาศาสตร์ หรือเทคโนโลยี

1.3 คุณค่าของวิทยาศาสตร์ เช่น ความมุ่งมั่นในการตีพิมพ์ผลงาน การไม่เอาเรื่องส่วนตัวมาเกี่ยวข้อง และการขจัดอคติ

1.4 ธรรมชาติของการให้เหตุผลทางวิทยาศาสตร์ เช่น การอนุมาน การอุปมาน การลงข้อสรุป เพื่อหาคำอธิบายที่ดีที่สุด การเปรียบเทียบความคล้ายคลึง การใช้แบบจำลอง

2. ลักษณะที่ใช้ในการตัดสินความรู้ที่สร้างจากวิทยาศาสตร์ ซึ่งหมายรวมถึง

2.1 คำกล่าวอ้างทางวิทยาศาสตร์ ต้องได้รับการสนับสนุนจากข้อมูลและเหตุผลทางวิทยาศาสตร์

2.2 บทบาทของการใช้กระบวนการสืบเสาะหาความรู้หลากหลายรูปแบบในการสร้างความรู้ กำหนดเป้าหมาย (เพื่อตรวจสอบสมมติฐานและระบุรูปแบบต่าง ๆ) และการออกแบบ (การสังเกต การควบคุมการทดลอง การวิจัยเชิงความสัมพันธ์)

2.3 ความผิดพลาดในการตรวจวัดส่งผลต่อความเชื่อมั่นในความรู้วิทยาศาสตร์

2.4 การใช้ บทบาท และข้อจำกัดของแบบจำลองที่เป็นรูปธรรม แบบจำลองที่เป็นระบบ และแบบจำลองที่เป็นนามธรรม

2.5 บทบาทของการทำงานแบบร่วมมือกัน การวิพากษ์วิจารณ์ และการตรวจสอบคุณภาพจากผู้อื่นในการสร้างความน่าเชื่อถือเกี่ยวกับคำกล่าวอ้างทางวิทยาศาสตร์

2.6 ความรู้ทางวิทยาศาสตร์มีบทบาทในการระบุถึงปัญหาทางสังคมและเทคโนโลยี

2.4 เจตคติต่อวิทยาศาสตร์

เจตคติต่อวิทยาศาสตร์มีบทบาทสำคัญที่จะทำให้เกิดความสนใจในเรื่องราวของวิทยาศาสตร์และเทคโนโลยี และประเด็นที่ส่งผลกระทบต่อตนเองโดยตรง เป้าหมายหนึ่งของการศึกษาวิทยาศาสตร์ คือ การพัฒนาให้นักเรียนเกิดเจตคติที่ดีต่อวิทยาศาสตร์ ซึ่งจะนำไปสู่การส่งเสริมสนับสนุนวิทยาศาสตร์ หาความรู้ และใช้ความรู้อย่างเหมาะสม เพื่อประโยชน์ต่อตนเอง ท้องถิ่น ประเทศ และสังคมโลก และนำไปสู่การพัฒนาการรับรู้ความสามารถในตนเองต่อไป

การประเมินการรู้วิทยาศาสตร์ของ PISA ตั้งอยู่บนแนวคิดว่า การรู้เรื่องวิทยาศาสตร์ของคนต้องมีทั้งเจตคติ ความเชื่อ แรงบันดาลใจ ความเชื่อในตนเอง การให้คุณค่า และแสดงออกด้วยการกระทำ PISA 2015 ประเมินเจตคติต่อวิทยาศาสตร์ของนักเรียนโดยใช้แบบสอบถามและประเมินเจตคติในสามด้าน ได้แก่

1. ความสนใจในวิทยาศาสตร์
2. การให้ความสำคัญกับวิธีการทางวิทยาศาสตร์ที่นำมาสู่การสืบเสาะหาความรู้
3. ความตระหนักถึงสิ่งแวดล้อม

รายละเอียดของการประเมินเจตคติต่อวิทยาศาสตร์ในแต่ละด้าน มีดังนี้

2.4.1 ความสนใจในวิทยาศาสตร์

1) ความอยากรู้อยากเห็นเกี่ยวกับวิทยาศาสตร์ และประเด็นที่เกี่ยวข้องกับวิทยาศาสตร์

2) ความตั้งใจที่จะหาความรู้วิทยาศาสตร์และทักษะเพิ่มเติม โดยใช้แหล่งข้อมูล และวิธีการที่หลากหลาย

3) ความสนใจในวิทยาศาสตร์อย่างต่อเนื่อง รวมถึงตระหนักถึงอาชีพการงานทางวิทยาศาสตร์

2.4.2 การให้ความสำคัญกับวิธีการทางวิทยาศาสตร์ที่นำมาสู่การสืบเสาะหาความรู้

1) การยึดถือว่าหลักฐานเป็นข้อมูลสำคัญที่นำมาสู่การสร้างคำอธิบายในเรื่องต่าง ๆ

2) การยึดมั่นในการใช้วิธีการทางวิทยาศาสตร์ในสถานการณ์ที่เหมาะสมเพื่อสืบเสาะหาความรู้

3) การเห็นความสำคัญของการวิพากษ์วิจารณ์ว่าเป็นเครื่องมือในการสร้างความน่าเชื่อถือของแนวคิดต่าง ๆ

2.4.3 ความตระหนักถึงสิ่งแวดล้อม

1) การแสดงออกถึงความห่วงใยในสิ่งแวดล้อมและการรักษาสิ่งแวดล้อมให้ยั่งยืน

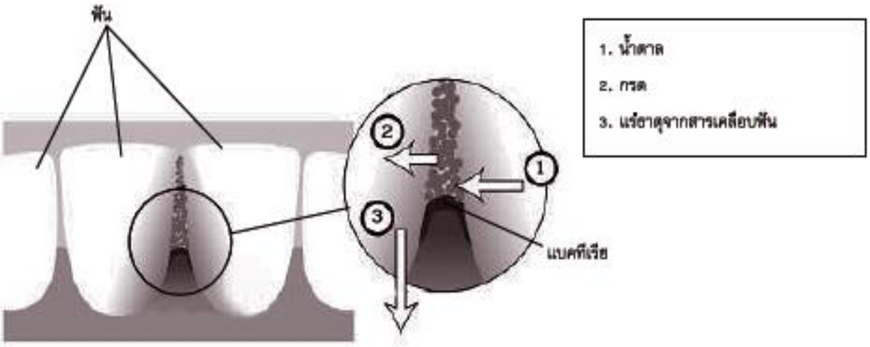
2) การมีแนวคิดในการใช้และส่งเสริมพฤติกรรมการรักษาสิ่งแวดล้อมให้ยั่งยืน

ฟันผุ

แบคทีเรียที่อยู่ในปากเป็นสาเหตุของฟันผุ ฟันผุเป็นปัญหามาตั้งแต่ปี ค.ศ.1700 นับตั้งแต่มีน้ำตาลจากการขยายอุตสาหกรรมน้ำตาลจากอ้อย

ปัจจุบันนี้ เรามีความรู้มากเกี่ยวกับฟันผุ ตัวอย่างเช่น:

- แบคทีเรียที่เป็นสาเหตุของฟันผุกินน้ำตาลเป็นอาหาร
- น้ำตาลถูกเปลี่ยนไปเป็นกรด
- กรดทำลายผิวของฟัน
- การแปรงฟันช่วยป้องกันฟันผุ



คำถามที่ 1 : ฟันผุ

แบคทีเรียมีบทบาทใดที่ทำให้ฟันผุ

1. แบคทีเรียสร้างสารเคลือบฟัน
2. แบคทีเรียสร้างน้ำตาล
3. แบคทีเรียสร้างแร่ธาตุ
4. แบคทีเรียสร้างกรด

คะแนนเต็ม

ข้อ 4. แบคทีเรียสร้างกรด

ไม่มีคะแนน

คำตอบอื่น ๆ

ลักษณะเฉพาะของข้อสอบ

สมรรถนะ : การใช้ประจักษ์พยานทางวิทยาศาสตร์

ความรู้ : การอธิบายเชิงวิทยาศาสตร์ (ความรู้เกี่ยวกับวิทยาศาสตร์)

การใช้ความรู้ : สุขภาพ

สถานการณ์ : ส่วนตัว

ลักษณะของข้อสอบ : เลือกตอบ

ภาพประกอบ 3 ตัวอย่างข้อสอบตามกรอบโครงสร้าง PISA แบบเลือกตอบ 4 ตัวเลือก

แบบทดสอบ

การพัฒนาแบบสอบในแต่ละประเภทให้มีคุณภาพและสามารถวัดคุณลักษณะได้ตรงตามกับสิ่งที่ต้องการวัดนั้น นักวัดผลมีความจำเป็นอย่างยิ่งที่จะต้องทำความเข้าใจเกี่ยวกับแนวคิดพื้นฐาน หลักการเพื่อจะนำไปสู่การสร้างและพัฒนาเครื่องมือวัดผลให้มีคุณภาพ สำหรับในส่วนนี้ผู้วิจัยจะนำเสนอในประเด็นเกี่ยวกับความหมายของแบบทดสอบ การออกแบบการทดสอบ การสร้างและพัฒนาแบบทดสอบ ลักษณะรูปแบบของข้อสอบแบบเลือกตอบ และคุณภาพของเครื่องมือ ดังนี้

1. ความหมายของแบบสอบและการทดสอบ

สำหรับแบบสอบนักวิชาการบางคนอาจเรียกว่า แบบทดสอบก็ได้ แต่ใช้ภาษาอังกฤษที่เหมือนกัน “test” มีความหมายดังนี้

Cronbach (1970) ได้ให้ความหมายของแบบสอบ ว่าหมายถึง กลุ่มคำถาม หรือกลุ่มงาน (Tasks) ที่ผู้ทดสอบได้เสนอให้ผู้ ถูกทดสอบตอบสนอง

พิชิต ฤทธิ์จรูญ (2550) ได้ให้ความหมายของแบบทดสอบว่า หมายถึง ชุดของคำถามหรือกลุ่มงานใด ๆ ที่สร้างขึ้นเพื่อชักนำให้ผู้ถูกทดสอบแสดงพฤติกรรมอย่างใดอย่างหนึ่งออกมาให้ผู้สอบสังเกตได้และวัดได้

ศิริชัย กาญจนวาสี (2552) ได้ให้ความหมายของแบบสอบว่า เป็นเครื่องมือวัดผลชนิดหนึ่ง ซึ่งประกอบด้วยชุดของข้อคำถามที่ใช้วัดกลุ่มตัวอย่างพฤติกรรมเกี่ยวกับความสามารถทางสมอง หรือความรู้สึกนึกคิดทางจิตใจหรือทักษะการดำเนินงานของบุคคลหรือกลุ่มบุคคล ภายใต้สถานการณ์ที่เป็นมาตรฐาน และมีการกำหนดหลักเกณฑ์ การใช้คะแนนที่ชัดเจน

จากความหมายของแบบสอบที่กล่าวมาข้างต้นสามารถสรุปได้ว่า แบบสอบหมายถึง ชุดของข้อคำถาม (Item) ที่ถูกรวบรวมไว้เป็นชุด โดยแต่ละข้อคำถามนั้นจะกระตุ้นหรือเร้าพฤติกรรมภายในของผู้สอบให้เกิดการตอบสนองหรือแสดงเป็นพฤติกรรม หรือสัญลักษณ์ต่าง ๆ ออกมา ทำให้สามารถมองเห็นได้อย่างชัดเจนหรือมีความเป็นรูปธรรมมากขึ้น เพื่อที่จะนำผลที่ได้ไปสู่การตัดสินคุณค่า

การทดสอบ (Testing) นั้นเป็นกระบวนการสำคัญสำหรับการวัดผลทางการศึกษา ซึ่ง Cronbach (1970) ได้กล่าวว่า แนวคิดเกี่ยวกับการทดสอบจะถูกต้องมากน้อยเพียงใดขึ้นอยู่กับการให้คำนิยามของ “การทดสอบ” ได้มีผู้ให้นิยามของการทดสอบไว้มากมาย สามารถสรุปได้ ดังนี้

Cronbach (1970) ได้ให้นิยาม การทดสอบว่า หมายถึง กระบวนการที่มีระบบในการสังเกตพฤติกรรมของมนุษย์ แล้วบรรยายผลการสังเกตด้วยมาตรแสดงจำนวน หรือด้วยการจำแนกประเภทอย่างเป็นระบบ

ล้วน สายยศ และอังคณา สายยศ (2543) ได้ให้นิยาม การทดสอบ หมายถึง การนำแบบสอบซึ่งสร้างขึ้นอย่างเป็นกระบวนการ และมีระบบไปตรวจสอบตัวอย่างของคุณลักษณะของสิ่งที่ต้องการวัด

ศิริชัย กาญจนวาสี (2552) ได้ให้นิยาม การทดสอบ หมายถึง กระบวนการใช้แบบสอบสำหรับกำหนดหรือบรรยายคุณลักษณะหรือคุณภาพเฉพาะอย่างของบุคคลหรือกลุ่มบุคคล เพื่อใช้เป็นสารสนเทศสำหรับการตัดสินใจ

จากความหมายของการทดสอบที่กล่าวมาข้างต้นสามารถสรุปได้ว่า การทดสอบ หมายถึง กระบวนการที่นำชุดของข้อคำถาม (Item) ที่ถูกสร้างมาอย่างเป็นระบบที่นำไปกำหนดคุณลักษณะหรือพฤติกรรมต่าง ๆ ของผู้สอบเพื่อที่จะนำผลที่ได้ไปสู่การตัดสินคุณค่า ดังนั้นจะเห็นได้ว่าแบบสอบและการทดสอบมีความสัมพันธ์กัน เพราะว่าแบบสอบเป็นเครื่องมือวัดผลชนิดหนึ่งที่ใช้ในการทดสอบ

2. การออกแบบการทดสอบ

ในการออกแบบการทดสอบนั้น นักวัดผลต้องตอบคำถามหลัก ๆ 3 ประการ ได้แก่ วัดอะไร วัดอย่างไร และควบคุมตัวแปรแทรกซ้อนได้อย่างไร โดยการออกแบบการทดสอบนั้นมีแนวทางดังนี้

วัดอะไร ซึ่งเป็นคำถามแรกที่นักวัดผลต้องทำการตีความหมายให้ได้ว่าจะวัดอะไร ต้องทราบเป้าหมายหรือสิ่งที่ต้องการจะทำการวัด เช่น การวิจัยครั้งนี้มุ่งที่จะวัดการรู้เรื่องวิทยาศาสตร์ของนักเรียน ดังนั้นต้องศึกษาและทำความเข้าใจเกี่ยวกับการรู้เรื่องวิทยาศาสตร์และนำไปสู่การกำหนดนิยามเชิงปฏิบัติการ

วัดอย่างไร เมื่อทำความเข้าใจเกี่ยวกับการรู้เรื่องวิทยาศาสตร์ และกำหนดนิยามเชิงปฏิบัติการเรียบร้อยแล้ว ในขั้นต่อไปคือ ให้ตรวจสอบดูว่าในแต่ละคุณลักษณะที่มุ่งวัดนั้นสามารถใช้เครื่องมือหรือวิธีการวัดที่มีคุณภาพได้อย่างไรบ้าง และมีเครื่องมือชนิดใดบ้าง เช่น ในกรณีงานวิจัยนี้การทดสอบการรู้เรื่องวิทยาศาสตร์ของนักเรียน การที่จะรู้ได้ว่านักเรียนนั้นมีการรู้เรื่องวิทยาศาสตร์หรือไม่ นักเรียนต้องทำแบบทดสอบ ซึ่งเป็นแบบทดสอบที่จะมุ่งวัดองค์ประกอบของการรู้เรื่องวิทยาศาสตร์ ด้านพฤติกรรมเนื้อหาวิทยาศาสตร์ โดยเป็นแบบทดสอบที่มีตัวเลือก 4 ตัวเลือก เมื่อสร้างแบบสอบแล้วต้องมีการตรวจสอบคุณภาพของแบบสอบ เพื่อดูว่าแบบสอบมีความตรง (Validity) ซึ่งวิธีการตรวจสอบมีหลายวิธีควรเลือกแต่ละวิธีให้เหมาะสมกับประเภทเครื่องมือ เมื่อดำเนินการตรวจสอบคุณภาพแล้วจนได้เครื่องมือที่มีความตรง จึงนำไปใช้กับผู้สอบที่เป็นกลุ่มเป้าหมาย

ควบคุมตัวแปรแทรกซ้อนอย่างไร ตัวแปรแทรกซ้อนในที่นี้ คือ ตัวแปรที่ส่งผลต่อการทดสอบที่ผู้วิจัย และนักวัดผลไม่ต้องการให้เกิดขึ้นในระหว่างการทดสอบ เช่น แบบสอบ

ไม่สอดคล้องกับกลุ่มผู้สอบ เนื้อหาในข้อสอบผู้สอบสามารถเดาคำตอบถูกได้ หรือความไม่พร้อมของสถานที่สอบ เป็นต้น หากการทดสอบนั้นมีตัวแปรแทรกซ้อนเกิดขึ้นอาจส่งผลต่อการทดสอบ ทำให้ผลสอบที่ได้ไม่สอดคล้องกับความสามารถที่แท้จริงของผู้สอบ

3. การสร้างและพัฒนาแบบทดสอบ

Gronlund และ Linn (2009) ได้เสนอแนวทางในการวางแผนการพัฒนาแบบทดสอบเพื่อใช้ในการทดสอบ สามารถสรุปได้ ดังนี้

1. กำหนดจุดมุ่งหมายของการทดสอบ (Specification of Purpose) ซึ่งการทดสอบในครั้งนี้เป็นการทดสอบอะไร ถ้าเป็นการสอบเพื่อดูความพร้อมก็ควรเป็นการทดสอบก่อนเรียน ถ้าเป็นการดูพัฒนาการระหว่างเรียนเราจะเรียกว่าการทดสอบย่อย ถ้าเป็นการทดสอบหลังเสร็จสิ้นกระบวนการเรียนการสอนเราจะเรียกว่า ทดสอบรวม ซึ่งเป็นการสอบประมวลผลความรู้ทั้งหมดที่ได้เรียนผ่านมาแล้วทั้งหมด ดังนั้นในการกำหนดจุดมุ่งหมายนั้นผู้ออกข้อสอบต้องดำเนินการวิเคราะห์หลักสูตร หรือวิเคราะห์องค์ประกอบของมวลประสบการณ์ทั้งหมดของการจัดการเรียนการสอน

2. สร้างตารางกำหนดคุณลักษณะของแบบสอบ (Table of Specifications) ซึ่งเป็นตารางสำหรับการวิเคราะห์หลักสูตรในวิชาหนึ่ง ๆ ว่า ก่อนดำเนินการออกข้อสอบนั้นจะต้องดูด้านเนื้อหาให้สัมพันธ์กับด้านพฤติกรรมที่มุ่งวัด

3. เลือกรูปแบบของแบบทดสอบให้เหมาะสมรูปแบบการเขียน แบบทดสอบมีหลายประเภท ได้แก่ แบบเลือกตอบที่มีคำตอบถูกเพียงคำตอบเดียว แบบเลือกตอบที่มีคำตอบถูกมากกว่า 1 คำตอบ แบบสอบแบบความเรียง เป็นต้น เมื่อพิจารณาตารางการกำหนดคุณลักษณะของแบบสอบแล้วควรเลือกรูปแบบให้เหมาะสมกับวัตถุประสงค์ เนื้อหาที่มุ่งวัด หรือพฤติกรรมที่ต้องการวัด

4. กำหนดแบบแผนของข้อสอบแต่ละข้อ (Item Specification) ซึ่งแบบแผนของข้อสอบคือ รูปแบบทั่วไปของข้อสอบแต่ละข้อที่สามารถใช้วัดสมรรถภาพและความสามารถของผู้สอบได้ตรงตามจุดประสงค์การเรียนรู้หรือสิ่งที่มุ่งวัด แบบแผนข้อสอบจะช่วยประหยัดเวลาในการเขียนข้อสอบครั้งต่อ ๆ ไป ซึ่งสามารถใช้สร้างข้อสอบเพื่อรวมเป็นแบบทดสอบได้หลาย ๆ ฉบับ และเป็นแนวทางในการพัฒนาแบบสอบคู่ขนานได้ แบบแผนของข้อสอบจะประกอบด้วย จุดประสงค์การเรียนรู้ที่มุ่งวัดสมรรถภาพ ความสามารถหรือระดับพฤติกรรมที่มุ่งวัดลักษณะคำถาม และลักษณะคำตอบที่ต้องการ เมื่อกำหนดแบบแผนของข้อสอบเสร็จแล้วจึงร่างข้อสอบ

5. รวมข้อสอบทำเป็นแบบทดสอบ ในการรวมข้อสอบแต่ละข้อมาทำเป็นแบบทดสอบชุดหนึ่งนั้น จุดสำคัญควรตรวจสอบดูว่า แต่ละข้อวัดจุดประสงค์การเรียนรู้นั้นจริงหรือไม่

สำหรับการชุดของแบบทดสอบนั้นมีหลักเกณฑ์ว่า ข้อสอบที่มีลักษณะรูปแบบข้อสอบเหมือนกันจะรวมอยู่ในชุดเดียวกัน การเรียงข้อสอบความเรียงจากข้อที่ง่ายไปหาข้อที่ยาก เป็นต้น

6. การตรวจสอบคุณภาพและประเมินคุณภาพของแบบสอบ เมื่อจัดทำแบบสอบเสร็จแล้วก่อนนำไปทดลองใช้นั้น ผู้ออกข้อสอบควรทำการทบทวนการสร้างข้อสอบก่อน โดยอาศัยผู้เชี่ยวชาญด้านเนื้อหา เพื่อดูว่าข้อคำถามที่สร้างในข้อสอบนั้นมีความสอดคล้องกับจุดประสงค์การเรียนรู้หรือมวลประสบการณ์ต่าง ๆ ที่ต้องการวัดหรือไม่ และเพื่อดูความเหมาะสมของภาษาและนำข้อสอบที่ได้รับข้อเสนอแนะไปปรับปรุง หลังจากนั้นอาจจะนำไปทดลองกับกลุ่มตัวอย่างกลุ่มเล็ก ๆ ที่ไม่ใช่กลุ่มเป้าหมายที่ต้องการศึกษาแต่ต้องมีบริบทที่ใกล้เคียงกันกับกลุ่มเป้าหมาย เพื่อตรวจสอบความยากง่าย อำนาจจำแนก ความตรงและความเที่ยงของข้อสอบว่ามีคุณภาพตามเกณฑ์ที่จะยอมรับได้ หรือไม่ถ้ายังไม่เป็นไปตามเกณฑ์ที่กำหนดนั้นต้องนำมาปรับปรุงอีกครั้งและทำไปทดลองใช้อีกรอบจนกว่าข้อสอบจะเป็นไปตามเกณฑ์ที่ยอมรับได้

7. การนำผลการสอบไปใช้ปรับปรุงเป้าประสงค์ของการเรียนรู้ เมื่อได้ข้อสอบและแบบสอบที่มีคุณภาพตามเกณฑ์ แล้วก็นำไปทดสอบกับกลุ่มเป้าหมายที่ต้องการวัดผลที่ได้จากการทดสอบจะออกมาในรูปของคะแนน ซึ่งคะแนนที่ได้นี้จะสามารถสะท้อนผลการเรียนรู้ ระดับความสามารถ และสมรรถนะของผู้สอบได้ว่าอยู่ในระดับใดหรือเป็นไปตามเป้าประสงค์ที่หลักสูตรต้องการหรือไม่ หากเป็นเช่นนี้ก็จะทำให้ผู้ที่จัดการศึกษาได้สารสนเทศที่เป็นประโยชน์ในการควบคุมคุณภาพของการศึกษา

4. ลักษณะรูปแบบของข้อสอบแบบเลือกตอบ

ได้มีผู้กล่าวถึงลักษณะของข้อสอบแบบเลือกตอบไว้ดังนี้

ล้วน สายยศ และอังคณา สายยศ (2539) กล่าวถึงลักษณะของข้อสอบแบบเลือกตอบว่า ประกอบด้วย ส่วนที่สำคัญที่สุด 2 ส่วน คือ ส่วนข้อคำถาม (Stem) และส่วนตัวเลือก (Alternative หรือ Choice) ตัวเลือกยังแบ่งออกเป็น 2 ส่วน คือ ตัวเลือกที่เป็นตัวถูก กับตัวเลือกที่เป็นตัวลวง

สมนึก ภัททิยธนี (2549) กล่าวว่า ข้อสอบแบบเลือกตอบ นอกจากจะต้องคำนึงถึงคุณภาพคุณภาพของคำถามและตัวเลือกแล้ว ควรคำนึงถึงรูปแบบ (Style) ของข้อสอบที่เลือกใช้ด้วย เพื่อช่วยให้ข้อสอบมีคุณภาพมากยิ่งขึ้น โดยพยายามเลือกใช้รูปแบบที่เห็นว่าเหมาะสมกับเนื้อหาหรือเรื่องราวที่ต้องการถาม แต่วิธีที่ดีของการออกข้อสอบแต่ละครั้ง ควรใช้รูปแบบต่าง ๆ กัน จะช่วยให้ข้อสอบมีคุณภาพยิ่งขึ้น ซึ่งรูปแบบเลือกตอบที่นิยมใช้กันมากมี 3 ชนิด คือ

1. รูปแบบคำถามเดี่ยว
2. รูปแบบตัวเลือกคงที่
3. รูปแบบสถานการณ์

จากเอกสารของลักษณะของข้อสอบแบบเลือกตอบ สรุปได้ว่าข้อสอบแบบเลือกตอบจะประกอบด้วย 2 ส่วน คือ ส่วนที่เป็นคำถาม และส่วนที่เป็นตัวเลือก ตัวเลือกยังแบ่งออกเป็นตัวเลือกที่เป็นตัวถูก และตัวเลือกที่เป็นตัวลวง และการสร้างรูปแบบข้อสอบเลือกตอบแต่ละวิชา หรือแต่ละฉบับควรใช้รูปแบบ ทั้ง 3 รูปแบบ ดังนั้นผู้วิจัยจึงสร้างข้อสอบแบบเลือกตอบโดยการใช้รูปแบบทั้ง 3 รูปแบบ ผสมผสานกันไป เพื่อให้แบบทดสอบมีคุณภาพยิ่งขึ้น

5. คุณภาพของเครื่องมือ

การที่นักวัดผลพัฒนาเครื่องมือที่มีคุณภาพนั้นผลที่ตามมา คือ ทำให้เกิดความมั่นใจในการนำเครื่องมือไปใช้ว่าสามารถวัดได้ตรงตามสิ่งที่ต้องการวัด และทำให้ทราบความสามารถที่แท้จริงของผู้สอบ หลังจากนั้นจึงนำผลที่ได้จากการสอบไปเป็นข้อมูลย้อนกลับในการปรับปรุงและพัฒนากระบวนการเรียนการสอน หรือปรับปรุงหลักสูตรให้ดีขึ้น รวมถึงการจัดกิจกรรมส่งเสริมต่าง ๆ เป็นต้น ตัวบ่งชี้คุณภาพของเครื่องมือวัดผลประกอบด้วย (ศิริชัย กาญจนวาสี, 2552)

1. ความเที่ยงตรง (Validity) หมายถึง ความถูกต้องแม่นยำของเครื่องมือในการวัดสิ่งที่ต้องการวัด หรือความสอดคล้องหรือความเหมาะสมของผลการวัดกับเนื้อเรื่องหรือเกณฑ์หรือทฤษฎีที่เกี่ยวข้องกับลักษณะที่มุ่งวัด การตรวจสอบความตรงเป็นกระบวนการรวบรวมและวิเคราะห์หลักฐาน เพื่อการสนับสนุนความเหมาะสมและความถูกต้องของการนำคะแนนจากเครื่องมือวัดไปสรุปอ้างอิงคุณลักษณะที่มุ่งวัด

ศิริชัย กาญจนวาสี (2552) กล่าวว่า การตรวจสอบความตรงของคะแนนสอบจากแบบสอบจะต้องมีความชัดเจนตั้งแต่ต้นว่า “ลักษณะที่ต้องการสรุปอ้างอิงไปถึงนั้นคืออะไร” เพื่อทำการเก็บรวบรวมหลักฐานอันเป็นประโยชน์ต่อการตัดสินความเหมาะสมและความถูกต้องได้ตรงตามเป้าหมายของการนำคะแนนนั้นไปใช้ สำหรับการตรวจสอบคุณภาพด้านความตรง สามารถจำเป็นตามเป้าหมายสำคัญได้ 3 ประเภท สามารถสรุปแต่และประเภทได้ ดังนี้

1. ความเที่ยงตรงตามเนื้อหา (Content Validity) คือ ความสามารถในการวัดกลุ่มตัวอย่าง เนื้อเรื่องที่วัดได้ครอบคลุมและเป็นตัวแทนของมวลประสบการณ์ทั้งหมดที่ต้องการจะวัดโดยมีวิธีการตรวจสอบความตรงตามเนื้อหา ได้แก่ 1) ให้ผู้เชี่ยวชาญตรวจสอบความเหมาะสมของนิยามและขอบเขตมวลเนื้อหาหรือประสบการณ์ที่มุ่งวัด 2) ตรวจสอบกลุ่มตัวอย่างของพฤติกรรมที่นำมาทำการทดสอบว่า มีความครอบคลุมตามเนื้อหาหรือมวลประสบการณ์ทั้งหมดหรือไม่ และ 3) เปรียบเทียบสัดส่วนของข้อคำถามว่าสอดคล้องกับน้ำหนักความสำคัญของแต่ละลักษณะเนื้อเรื่องที่มุ่งวัดมากน้อยเพียงใด

2. ความเที่ยงตรงตามเกณฑ์สัมพันธ์ (Criterion-Related Validity) แบ่งออกเป็น 2 ประเภท คือ ความตรงตามสภาพ (Concurrent Validity) และ ความตรงเชิงทำนาย (Predictive Validity) ซึ่งความตรงตามสภาพ คือ ความสามารถในการวัดคุณลักษณะที่สนใจ

ได้ตรงตามสมรรถนะของสิ่งนั้น ในสภาพปัจจุบันมีวิธีการตรวจสอบ ได้แก่ การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่วัดได้จากเครื่องมือที่สร้างขึ้นใหม่กับคะแนนที่วัดได้จากเครื่องมือมาตรฐาน ซึ่งวัดได้ในสภาพปัจจุบันสำหรับความตรงเชิงทำนาย คือ ความสามารถในการวัดคุณลักษณะที่สนใจได้ตรงตามสมรรถนะของสิ่งนั้นที่จะเกิดขึ้นในอนาคต มีวิธีการตรวจสอบ ได้แก่ การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่วัดได้จากเครื่องมือที่สร้างขึ้นใหม่กับคะแนนที่วัดได้จากเครื่องมือมาตรฐาน ซึ่งสามารถวัดได้ ในเวลาต่อมาหรือในอนาคต

3. ความเที่ยงตรงเชิงโครงสร้าง (Construct Validity) คือ ความสามารถในการวัดได้ตรงตามลักษณะที่มุ่งวัดโดยผลการวัดมีความสอดคล้องกับโครงสร้างและความหมายทางทฤษฎีของลักษณะที่มุ่งวัด มีวิธีการตรวจสอบได้โดยการศึกษาความสัมพันธ์ระหว่างผลการวัดที่ได้จากเครื่องมือกับโครงสร้าง โดยอาศัยการสนับสนุนจากวิธีการวิเคราะห์ต่าง ๆ ได้แก่ 1) วิธีตัดสินโดยผู้เชี่ยวชาญ 2) วิธีเปรียบเทียบคะแนนระหว่างกลุ่มที่ทราบผล (Comparing the Scores of Know Groups) 3) วิธีเปรียบเทียบจากกลุ่มทดลอง (Comparing Scores from Experiment) 4) การวิเคราะห์เมตริกซ์พหุลักษณะ-พหุวิธี (Multitrait-Multimethod) 5) วิธีวิเคราะห์องค์ประกอบ (Factor Analysis) นอกจากนั้น Wolfe และ Smith (2007) ได้นำเสนอวิธีการแสดงหลักฐานความตรงเชิงโครงสร้างด้วยการวิเคราะห์ความเป็นพหุมิติ (Multidimensionality) โดยเป็นการแสดงหลักฐานเมื่อการทดสอบมีลักษณะที่มุ่งวัดเป็นพหุมิติ ดังนั้นจึงต้องมีการใช้โมเดลการวัดแบบพหุมิติช่วยยืนยันโครงสร้างที่วัดด้วย

2. ความเชื่อมั่น (Reliability) หมายถึง ความคงที่หรือความคงเส้นคงวาของผลที่ได้จากการวัดซ้ำ ถ้ามีการวัดสิ่งเดียวกันหลาย ๆ ครั้งได้ค่าที่ค่อนข้างคงเส้นคงวาสูงขึ้นเพียงใดถือว่าการวัดมีความเที่ยงมากขึ้นเพียงนั้น ถ้ามองตามนัยของทฤษฎี ความเชื่อมั่นในเชิงคณิตศาสตร์จะหมายถึง อัตราส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่สังเกตได้ (ศิริชัย กาญจนวาสี, 2552) การประมาณค่าความเชื่อมั่นสามารถกระทำได้หลายวิธี ศิริชัย กาญจนวาสี (2552) ได้จำแนกเป็น 4 ประเภท ดังรายละเอียดต่อไปนี้

2.1 ความเชื่อมั่นแบบความคงที่ (Measure of Stability) เป็นการหาความคงเส้นคงวาของคะแนนจากการวัดในช่วงเวลาที่ต่างกัน โดยวิธีสอบซ้ำด้วยแบบสอบเดิม (Test-Retest Method) โดยการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่วัดได้จากคนกลุ่มเดียวกัน ด้วยเครื่องมือเดียวกัน โดยทำการวัดซ้ำสองครั้งในเวลาที่ต่างกัน

2.2 ความเชื่อมั่นแบบความสมมูล (Measure of Equivalence) เป็นการหาความสอดคล้องของคะแนนจากการวัดในช่วงเวลาเดียวกัน โดยใช้แบบสอบที่สมมูลกัน (Equivalent Forms Method) โดยการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่วัดในเวลาเดียวกันจากกลุ่มคนเดียวกัน โดยใช้เครื่องมือ 2 ฉบับ ที่ทัดเทียมกัน

2.3 ความเชื่อมั่นแบบความคงที่และสมมูล (Measure of Stability and Equivalence) เป็นการหาความสอดคล้องของคะแนนจากการวัดในช่วงเวลาที่ต่างกัน โดยวิธีสอบซ้ำด้วยแบบสอบที่สมมูล (Test-Retest with Equivalent) โดยคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่วัดได้ในช่วงเวลาที่ต่างกัน จากกลุ่มเดียวกัน โดยใช้เครื่องมือ 2 ฉบับ ที่ทัดเทียมกัน

2.4 ความเชื่อมั่นแบบสอดคล้องภายใน (Internal Consistency Coefficient) เป็นวิธีการหาความสอดคล้องกันระหว่างคะแนนรายข้อหรือความเป็นเอกพันธ์ของเนื้อหารายข้อที่เป็นตัวแทนของคุณลักษณะเด่นเดียวกันที่ต้องการวัด โดยใช้วิธีต่างกัน ได้แก่ วิธีแบ่งครึ่งข้อสอบ (Split-half Method) วิธีของคูเดอร์ ริชาร์ดสัน (Kuder-Richardson Method) วิธีสัมประสิทธิ์แอลฟาของครอนบัค (Cronbach's Alpha Method) และวิธีวิเคราะห์ความแปรปรวนของฮอยท์ (Hoyt's Analysis of Variance Method)

3. อำนาจจำแนกของข้อสอบ (Discrimination Power of the Item) หมายถึง ความสามารถของข้อสอบในการจำแนกหรือแยกให้เห็นความแตกต่างระหว่างผู้สอบที่มีผลสัมฤทธิ์ต่างกัน เช่น ระหว่างผู้ที่มีสมรรถนะกับผู้ที่ไม่มีสมรรถนะ เป็นต้น หรืออีกความหมายหนึ่ง คือ ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนคุณลักษณะที่มุ่งวัดกับคะแนนรายข้อ

4. ความยากง่าย (Difficulty) หมายถึง จำนวนคนตอบข้อสอบได้ถูกมากน้อยเพียงใด หรืออัตราส่วนของจำนวนคนตอบถูกกับจำนวนคนทั้งหมดที่เข้าสอบ ตามทฤษฎีการวัดผลแบบอิงกลุ่ม ข้อสอบที่ดีคือข้อสอบที่ไม่ยากไม่ง่ายเกินไป เรียกว่า มีความยากพอเหมาะ เพราะคุณค่าของข้อสอบดังกล่าวจะช่วยจำแนกผู้สอบได้ว่าใครเก่ง ใครอ่อน ข้อสอบข้อใดที่ไม่มีใครทำถูก หรือข้อสอบที่ทุกคนทำถูก ต่างก็ไม่สามารถจำแนกผู้สอบได้ว่าใครเก่ง ใครอ่อน จึงไม่มีคุณค่าในการจำแนก ส่วนทฤษฎีในการวัดผลแบบอิงเกณฑ์ถือว่าข้อสอบที่ดี คือ สามารถวัดว่าผู้เรียนได้บรรลุจุดประสงค์หรือไม่ ถ้าวัดได้จริงก็ถือว่าเป็นข้อสอบที่ดี

จากการศึกษาเอกสารเกี่ยวกับแบบทดสอบและการพัฒนาแบบทดสอบ จึงทำให้ผู้วิจัยสนใจที่พัฒนาแบบทดสอบการรู้เรื่องวิทยาศาสตร์ โดยแบบทดสอบที่พัฒนาขึ้นเป็นแบบทดสอบแบบเลือกตอบ 4 ตัวเลือก เนื่องจากการออกข้อสอบวิทยาศาสตร์ตามกรอบการประเมินการรู้เรื่องวิทยาศาสตร์นั้นประกอบไปด้วย ข้อสอบแบบเลือกตอบ ข้อสอบแบบตอบเชิงซ้อน และข้อสอบแบบเขียนคำตอบแบบสั้น ๆ ซึ่งสัดส่วนของข้อสอบแบบเลือกตอบมีสัดส่วนมากที่สุดในจำนวนข้อที่ทำการทดสอบการรู้เรื่องวิทยาศาสตร์ของนักเรียน และในส่วนของการหาคุณภาพของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ผู้วิจัยจึงได้ทำการหาคุณภาพของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ เพื่อเป็นการหาตัวบ่งชี้ว่าแบบทดสอบการรู้เรื่องวิทยาศาสตร์ สามารถวัดได้ตรงตามสิ่งที่ต้องการวัด และทำให้ทราบความสามารถที่แท้จริงของผู้สอบได้มากน้อยเพียงใด ประกอบด้วย ตัวบ่งชี้คุณภาพของแบบทดสอบ

การรู้เรื่องวิทยาศาสตร์ ได้แก่ ความเที่ยงตรงตามเนื้อหา ค่าอำนาจจำแนกของข้อสอบ และความยากง่าย

ทฤษฎีการตอบสนองของข้อสอบ

Hambleton, Swaminathan และ Rogers (1991) กล่าวว่า ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT) อยู่บนแนวคิด 2 ประการ ดังนี้ 1) พฤติกรรมในการตอบข้อสอบของผู้เข้าสอบนำไปใช้ทำนายความสามารถ (Ability) หรือลักษณะภายใน (Trait) ของผู้เข้าสอบ โดยมีข้อตกลงเบื้องต้น ดังนี้

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT) เป็นทฤษฎีที่พัฒนาขึ้นเพื่อแก้ไขจุดด้อยของทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory : CTT) ซึ่งจุดด้อยที่สำคัญ คือ ค่าพารามิเตอร์ของข้อสอบจะแปรผันตามกลุ่มของผู้สอบ และคะแนนหรือการประมาณค่าความสามารถของผู้เข้าสอบไม่เป็นอิสระ ขึ้นอยู่กับข้อสอบหรือแบบทดสอบที่นำมาใช้ สำหรับทฤษฎีการตอบสนองข้อสอบสามารถแบ่งออกเป็น 2 ประเภท คือ 1) ทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนนแบบ 2 ค่า (Dichotomous Item Response Theory) และ 2) ทฤษฎีการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous Item Response Theory) สำหรับงานวิจัยนี้กล่าวถึงเฉพาะในส่วนของทฤษฎีการตอบสนองข้อสอบ แบบตรวจให้คะแนน 2 ค่า (Dichotomous Item Response Theory) เท่านั้น โดยมีรายละเอียด ดังนี้

ทฤษฎีการตอบสนองข้อสอบ แบบตรวจให้คะแนนแบบ 2 ค่า (Dichotomous Item Response Theory) เป็นทฤษฎีที่อธิบายถึงความสัมพันธ์ระหว่างความสามารถของผู้สอบกับการตอบข้อสอบโดยใช้โค้งคุณลักษณะข้อสอบ (Item Characteristic Curve : ICC) ซึ่งมีการกำหนดคุณลักษณะข้อสอบด้วยค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าการเดา (c) มีหลักการตรวจให้คะแนนเพียง 2 ค่า เช่น ถูก-ผิด ใช่-ไม่ใช่ หรือ 0,1 เป็นต้น

1. ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

1.1 ความเป็นมิติเดียว (Unidimension) หมายถึง ผลการตอบข้อสอบของผู้เข้าสอบสามารถอธิบายความสามารถหรือคุณลักษณะภายในเพียงด้านใดด้านหนึ่งของผู้เข้าสอบและคุณลักษณะภายในด้านเดียวนี้มีความหมายเหมือนเป็นมิติเดียว ซึ่งข้อตกลงนี้ชี้ให้เห็นว่าอาจมีคุณลักษณะของข้อสอบบางประการที่ส่งผลร่วมต่อการตอบข้อสอบเข้ามาเกี่ยวข้อง ดังนั้น จึงควรกำหนดความเป็นมิติเดียวให้เป็นลักษณะเด่น (Dominant) หรือลักษณะหลัก เพื่อที่จะนำไปอธิบายผลการตอบข้อสอบของผู้เข้าสอบได้

1.2 ความเป็นอิสระในการตอบข้อสอบ (Local Independence) หมายถึง เมื่อค่าความสามารถของผู้เข้าสอบเป็นค่าแน่นอน การตอบข้อสอบแต่ละข้อของผู้เข้าสอบคนหนึ่งจะมีความเป็นอิสระจากกัน กล่าวได้ว่า การตอบข้อสอบข้อใด ๆ ของผู้เข้าสอบจะไม่มีผลต่อข้อสอบข้ออื่น ๆ เลย แต่สิ่งที่ส่งผลต่อการตอบข้อสอบแต่ละข้อเป็นผลมาจากความสามารถของผู้เข้าสอบเท่านั้น ความเป็นอิสระในการตอบข้อสอบทำให้ค่าพารามิเตอร์ข้อสอบยังเป็นค่าคงที่ ไม่ว่าข้อสอบข้อนั้นอยู่ตำแหน่งใด ๆ ก็ตาม

1.3 โค้งคุณลักษณะของข้อสอบ (Item Characteristic Curve : ICC) หรือฟังก์ชันการตอบสนองข้อสอบ (Item Response Function : IRF) เป็นฟังก์ชันทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างความน่าจะเป็นของการตอบข้อสอบถูกกับระดับความสามารถของผู้เข้าสอบ

2. โมเดลการตอบสนองข้อสอบ (Item Response Models)

โมเดลการตอบสนองข้อสอบ เป็นโมเดลแสดงความสัมพันธ์ระหว่างโอกาสตอบข้อสอบถูกกับความสามารถของผู้เข้าสอบในรูปแบบของโค้งคุณลักษณะเฉพาะของข้อสอบ ซึ่งมีลักษณะเป็นฟังก์ชั่นโลจิสติก (Logistic Function) หรือฟังก์ชันปกติสะสม (Normal Ogive Function) สามารถเรียกอีกอย่างหนึ่งว่า “โมเดลโลจิสติกหรือโมเดลปกติสะสม” โมเดลการตอบสนองข้อสอบ มี 3 รูปแบบ ดังต่อไปนี้

2.1 โมเดลการตอบสนองข้อสอบแบบ 1 พารามิเตอร์ (One-Parameter Model) เป็นโมเดลที่อธิบายผลการวิเคราะห์ข้อสอบด้วยค่าความยากของข้อสอบ (b) เรียกอีกอย่างว่า “Rasch Model” สามารถเขียนฟังก์ชันโลจิสติก ตามสมการที่ 1

$$P_i(\theta) = \frac{1}{1+e^{-(\theta-b_i)}} \qquad i = 1, 2, 3,\ldots n \qquad (1)$$

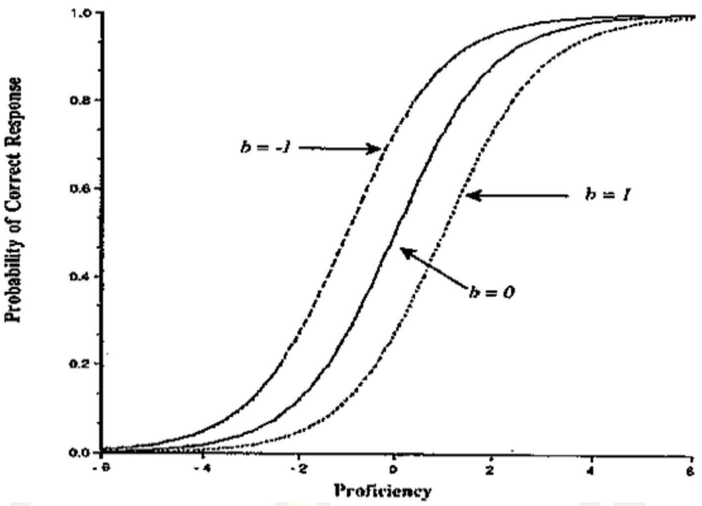
เมื่อ

$P_i(\theta)$ คือ ความน่าจะเป็นของผู้เข้าสอบที่มีความสามารถ (θ) จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

b_i คือ ค่าความยากของข้อสอบข้อที่ i

θ คือ ความสามารถของผู้เข้าสอบ

e คือ 2.72



ภาพประกอบ 4 โค้งคุณลักษณะของข้อสอบแบบ 1 พารามิเตอร์

2.2 โมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ (Two-Parameter Model) เป็นโมเดลการวิเคราะห์ข้อสอบที่ใช้ค่าพารามิเตอร์แบบ 2 พารามิเตอร์ คือ ค่าความยากของข้อสอบ (b) และค่าอำนาจจำแนกของข้อสอบ (a) เขียนเป็นฟังก์ชันโลจิสติก ตามสมการที่ 2

$$P_i(\theta) = \frac{1}{1+e^{-Da_i(\theta-b_i)}} \quad i = 1, 2, 3,\ldots n \qquad (2)$$

เมื่อ

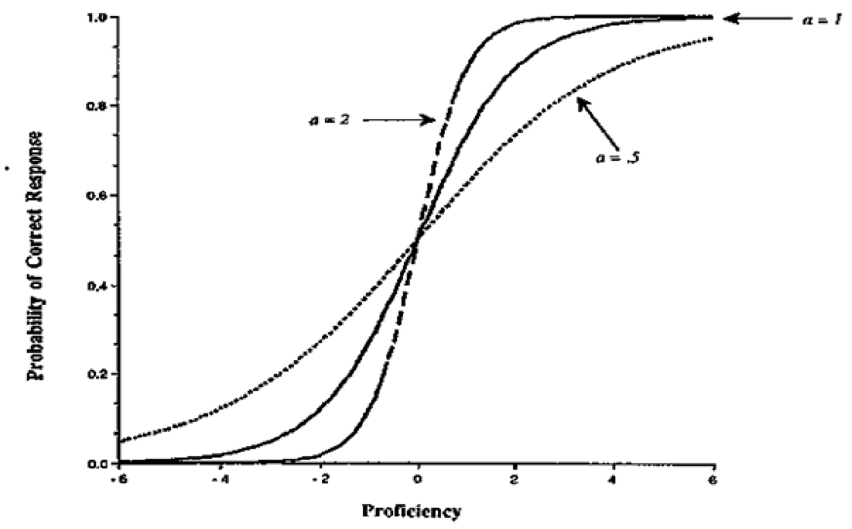
$P_i(\theta)$ คือ ความน่าจะเป็นของผู้เข้าสอบที่มีความสามารถ (θ) จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

b_i คือ ค่าความยากของข้อสอบข้อที่ i

a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i

θ คือ ความสามารถของผู้เข้าสอบ

D คือ 1.70



ภาพประกอบ 5 โค้งคุณลักษณะของข้อสอบ แบบ 2 พารามิเตอร์

2.3 โมเดลการตอบสนองข้อสอบ 3 พารามิเตอร์ (Three-Parameter Model) เป็นโมเดลพัฒนามาจากโมเดลการวิเคราะห์ข้อสอบที่ใช้ค่าพารามิเตอร์ 3 พารามิเตอร์ คือ ค่าความยากของข้อสอบ (b) ค่าอำนาจจำแนกของข้อสอบ (a) และค่าการเดาของข้อสอบ (c) เขียนเป็นฟังก์ชันโลจิสติก ตามสมการที่ 3

$$Pi(\theta) = c_i + \frac{1-c_i}{1+e^{-Da_i(\theta-bi)}} \quad i = 1, 2, 3,\ldots n \qquad (3)$$

เมื่อ

$P_i(\theta)$ คือ ความน่าจะเป็นของผู้เข้าสอบที่มีความสามารถ (θ) จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

b_i คือ ค่าความยากของข้อสอบข้อที่ i

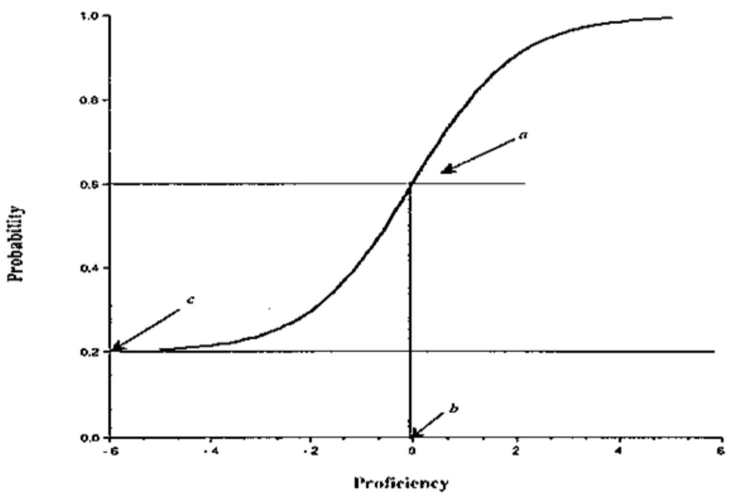
a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i

c_i คือ ค่าการเดาของข้อสอบข้อที่ i

θ คือ ความสามารถของผู้เข้าสอบ

D คือ 1.70

e คือ 2.72



ภาพประกอบ 6 โค้งคุณลักษณะของข้อสอบแบบ 3 พารามิเตอร์

3. ความไม่เปลี่ยนแปลงของพารามิเตอร์

เมื่อโมเดลการตอบสนองข้อสอบมีความสอดคล้องกับข้อมูลที่มีอยู่ จะทำให้ค่าพารามิเตอร์ของข้อสอบ (Item Parameter) และค่าพารามิเตอร์ความสามารถของผู้เข้าสอบ (Ability Parameter) ไม่เปลี่ยนแปลง ซึ่งเป็นคุณสมบัติสำคัญของทฤษฎีการตอบสนองข้อสอบ (IRT) ซึ่งโค้งคุณลักษณะของข้อสอบ (ICC) จะมีลักษณะเดียวกัน (a, b และ c) สำหรับทุกกลุ่มความสามารถของผู้เข้าสอบ นั่นคือ โค้งคุณลักษณะข้อสอบมีความคงที่ข้ามกลุ่มผู้เข้าสอบ

การทำหน้าที่ต่างกันของข้อสอบ

1. ความหมายของการทำหน้าที่ต่างกันของข้อสอบ

ในการศึกษาเรื่องผลการสอบของกลุ่มผู้เข้าสอบย่อยจากกลุ่มผู้เข้าสอบทั้งหมดมีการศึกษามานานแล้ว แต่เพิ่งมีการศึกษาเรื่องของความยุติธรรมในการสอบระหว่างผู้เข้าสอบย่อยต่างกลุ่มกันอย่างจริงจัง ในช่วงปลายทศวรรษที่ 1960 โดยมีการนำเสนอวิธีการต่าง ๆ ในการตรวจสอบความลำเอียงของแบบทดสอบ (Test Bias) และความลำเอียงในการคัดเลือกผู้ที่จะเข้าสอบ (Selection Bias) เพิ่มขึ้นหลายวิธี ในช่วงเวลานั้นนักพัฒนาแบบทดสอบมีความสนใจวิธีการจำแนกข้อสอบที่ไม่เหมาะสมกับผู้เข้าสอบบางกลุ่มออกจากแบบทดสอบ ก่อนที่จะมีการพัฒนาให้เป็นแบบทดสอบฉบับสมบูรณ์ จึงมีการพัฒนาวิธีการตรวจสอบความลำเอียงของข้อสอบ (Item Bias) เพื่อใช้ในการจำแนกข้อสอบที่มีความลำเอียงกับกลุ่มผู้เข้าสอบบางกลุ่มที่มีลักษณะบางอย่างแตกต่างกัน เช่น เชื้อชาติ ศาสนา วัฒนธรรม ภูมิลำเนา สังคม เพศ ภาษา อายุและประสบการณ์

เป็นต้น เพื่อเป็นการพัฒนาแบบทดสอบให้มีคุณภาพที่เหมาะสมสามารถนำไปใช้ในการทดสอบต่อไปได้

การศึกษาเรื่องผลการสอบในช่วงแรก ๆ มีวัตถุประสงค์เพื่อคัดเลือกคนเข้าศึกษาต่อหรือเข้าทำงาน แต่มีหลักฐานปรากฏอย่างชัดเจนที่แสดงให้เห็นว่ามีความลำเอียงเกิดขึ้นกับกลุ่มคนต่างชาติ เพศ ทำให้ต้องมีการศึกษาความลำเอียงในการคัดเลือกผู้เข้าสอบ เพื่อให้การศึกษานี้มีความถูกต้อง ชัดเจนเพิ่มมากขึ้น ในเวลาต่อมาจึงมีการศึกษาในระดับข้อสอบ (Item Level) ที่เรียกว่า ความลำเอียงของข้อสอบ (Item Bias) แต่ในปัจจุบันนักวิจัยทางการวัดผลส่วนใหญ่ ใช้คำว่า ข้อสอบทำหน้าที่ต่างกันกับกลุ่มผู้เข้าสอบย่อยต่างกลุ่มกัน หรือเรียกสั้น ๆ ว่า ข้อสอบทำหน้าที่ต่างกัน (Differential Item Functioning : DIF) โดยเห็นว่าเป็นคำที่มีความหมายกลาง ๆ และมีความเหมาะสมในเชิงวิชาการมากกว่าคำว่าความลำเอียง (Bias) เป็นคำที่ใช้กันในทางสังคมและมีความหมายในเชิงลบ แต่อย่างไรก็ตามคำสองคำนี้มีจุดเน้นที่แตกต่างกัน โดยคำว่าความลำเอียงของข้อสอบจะเน้นที่อิทธิพลที่สังเกตได้ของกลุ่มผู้เข้าสอบย่อยที่มุ่งศึกษา ส่วนคำว่าข้อสอบที่ทำหน้าที่ต่างกันเน้นที่ลักษณะทางสถิติของข้อสอบที่ทำการตรวจสอบได้ด้วยวิธีการวิเคราะห์ทางสถิติ ซึ่งเป็นองค์ประกอบหนึ่งที่แสดงถึงความลำเอียงของข้อสอบ (Scheuneman และ Bleistein, 1989 ; Angoff, 1993 ; Hambleton และคณะ, 1993 ; Zieky, 1993 ; Camilli และ Shepard, 1994) จากจุดเน้นนี้แสดงให้เห็นว่าวิธีการทางสถิติที่นำมาใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเป็นสิ่งที่มีความจำเป็นในการประเมินความลำเอียงของข้อสอบ แต่ถ้ามีการใช้เฉพาะวิธีการทางสถิติอย่างเดียวผลการตรวจสอบพบข้อสอบทำหน้าที่ต่างกัน ที่ไม่อาจจะสรุปได้ว่าข้อสอบข้อนั้นมีความลำเอียงหรือไม่ เนื่องจากการประเมินความลำเอียงของข้อสอบยังต้องรวมถึงการใช้วิธีการตัดสินข้อสอบ (Judgmental Method) โดยมีผู้เชี่ยวชาญพิจารณาเนื้อหาสาระของข้อสอบและจุดมุ่งหมายในการวัดของแบบทดสอบก่อนที่จะสรุปว่าข้อสอบข้อนั้นมีความลำเอียงหรือไม่

ปัจจุบันนี้นักวิจัยทางการวัดผลหลายท่านใช้คำว่าการทำหน้าที่ต่างกันของข้อสอบแทนคำว่า ความลำเอียงของข้อสอบ ซึ่งมีนักวิจัยทางการวัดผลได้ให้ความหมายของการทำหน้าที่ต่างกันของข้อสอบ ไว้ดังนี้

Holland และ Wainer (1993) กล่าวว่า การทำหน้าที่ต่างกันของข้อสอบ หมายถึง สารสนเทศทางสถิติของข้อสอบที่ได้จากกลุ่มผู้เข้าสอบต่างกลุ่มกันและมีความสามารถเท่ากัน แต่มีโอกาสในการตอบข้อสอบได้ถูกต้องแตกต่างกัน

Camilli และ Shepard (1994) กล่าวว่า การทำหน้าที่ต่างกันของข้อสอบ หมายถึง การตรวจสอบความเป็นพหุมิติในการวัดของข้อสอบ ซึ่งแสดงได้จากการแจกแจงความสามารถหลัก (Primary Ability) ของกลุ่มผู้สอบตั้งแต่ 2 กลุ่มขึ้นไปมีความเท่ากัน แต่มีการแจกแจงความสามารถรอง (Secondary Ability) แตกต่างกัน

Narayanan และ Swaminathan (1996) กล่าวว่า การทำหน้าที่ต่างกันของข้อสอบ หมายถึง ผู้สอบมีความสามารถระดับเดียวกัน แต่มาจากกลุ่มย่อยแตกต่างกัน มีโอกาสในการตอบข้อสอบได้ถูกต้องแตกต่างกัน

มีผู้ให้ความหมายของคำว่า การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning : DIF) ไว้หลายความหมายดังกล่าวไว้แล้วข้างต้น แต่ความหมายที่เป็นที่ยอมรับกันอย่างกว้างขวาง คือ ข้อสอบทำหน้าที่ต่างกันภายใต้เงื่อนไขผู้เข้าสอบที่มีความสามารถเท่ากัน แต่มาจากกลุ่มผู้สอบย่อยที่มีลักษณะต่างกัน มีความน่าจะเป็นในการตอบข้อสอบข้อนั้นไม่เท่ากัน

ดังนั้นจึงสรุปได้ว่า การทำหน้าที่ต่างกันของข้อสอบ (DIF) หมายถึง การที่ข้อสอบทำให้ผู้เข้าสอบจากกลุ่มต่างกันที่มีความสามารถหรือคุณลักษณะที่มุ่งวัดเท่ากัน มีโอกาสในการตอบข้อสอบได้ถูกต้องแตกต่างกัน เป็นการเปรียบเทียบผลการตอบระหว่างผู้เข้าสอบ 2 กลุ่ม คือ กลุ่มอ้างอิง (Reference Group : R) และกลุ่มเปรียบเทียบ (Focal Group : F) กลุ่มอ้างอิงเป็นกลุ่มที่คาดว่าจะได้ประโยชน์จากการตอบข้อสอบ มีโอกาสในการตอบข้อสอบถูกได้มากกว่าผู้เข้าสอบกลุ่มเปรียบเทียบและกลุ่มเปรียบเทียบเป็นกลุ่มที่คาดว่าจะเสียประโยชน์จากการตอบข้อสอบ มีโอกาสตอบข้อสอบถูกได้น้อยกว่าผู้เข้าสอบกลุ่มอ้างอิง การทำหน้าที่ต่างกันของข้อสอบเกิดขึ้นเมื่อนำข้อสอบไปทดสอบกับผู้เข้าสอบกลุ่มย่อยต่างกันที่มีความสามารถหลัก (Primary Ability) เท่ากันหรือมีคุณลักษณะแฝง (Secondary Ability) แตกต่างกัน ทำให้ผู้เข้าสอบต่างกลุ่มที่นำมาจับคู่เปรียบเทียบมีโอกาสตอบข้อสอบถูกแตกต่างกัน

การทดสอบแต่ละครั้งผู้สอบระหว่างกลุ่มย่อยอาจมีลักษณะที่แตกต่างกัน เช่น เชื้อชาติ ศาสนา วัฒนธรรม ภูมิลำเนา สังคม เพศ ภาษา อายุ ประสบการณ์ เป็นต้น ผู้สอบกลุ่มย่อยอาจไม่ได้รับความยุติธรรมในการทำข้อสอบ โดยข้อสอบบางข้ออาจมีความลำเอียงเข้าข้างผู้สอบกลุ่มย่อยบางกลุ่มของผู้สอบทั้งหมด ซึ่งทำให้เกิดการได้เปรียบเสียเปรียบระหว่างผู้สอบกลุ่มย่อยด้วยกัน ทั้ง ๆ ที่สอบด้วยข้อสอบฉบับเดียวกัน สาเหตุดังกล่าวอาจเนื่องมาจากแบบสอบไม่ได้วัดความสามารถเป้าหมายที่ต้องการวัดเพียงอย่างเดียว แต่ยังวัดความสามารถแทรกซ้อนที่ไม่ต้องการวัดอีกด้วย ตัวอย่างเช่น แบบสอบวัดคำศัพท์ในวิชาภาษาอังกฤษฉบับหนึ่ง ข้อสอบบางข้ออาจถามความรู้สำหรับผู้ชายเป็นพิเศษ เช่น ความรู้เรื่องกีฬา ในขณะที่ข้อสอบบางข้ออาจถามความรู้สำหรับผู้หญิงโดยเฉพาะ เช่น ความรู้เกี่ยวกับงานในบ้าน จากสถานการณ์ดังกล่าวทักษะวัดคำศัพท์ในวิชาภาษาอังกฤษเป็นความสามารถเป้าหมาย (θ) ส่วนทักษะวัดความรู้ด้านกีฬา (η_1) และงานในบ้าน (η_2) เป็นความสามารถแทรกซ้อน ข้อสอบทุกข้อในแบบสอบจะวัดความสามารถเป้าหมาย ส่วนข้อสอบบางข้อที่ทำหน้าที่ต่างกันจะวัดทั้งความสามารถเป้าหมายและความสามารถแทรกซ้อน นั่นคือ ถ้าผู้สอบกลุ่มย่อยกลุ่มใดมีความแทรกซ้อนสูงกว่าก็มีโอกาสในการตอบข้อสอบได้ถูกต้อง

มากกว่า ทั้ง ๆ ที่ระดับความสามารถเป้าหมายที่ต้องการวัดเท่ากัน จึงมีผลทำให้ข้อสอบทำหน้าที่ต่างกัน

การศึกษาถึงคุณภาพของข้อสอบจากผลการตรวจสอบข้อสอบของผู้สอบกลุ่มต่าง ๆ ในประชากรมีมานานแล้ว แต่การศึกษาคุณภาพด้านความยุติธรรมของข้อสอบหรือแบบสอบระหว่างผู้สอบกลุ่มต่าง ๆ เริ่มศึกษากันอย่างจริงจังในช่วงปลายทศวรรษของปี ค.ศ. 1960 มีเสนอวิธีการต่าง ๆ เพื่อตรวจสอบความลำเอียงของข้อสอบ (Item Bias) ความลำเอียงของแบบสอบ (Test Bias) และความลำเอียงในการคัดเลือก (Selection Bias) โดยนิยามความลำเอียงว่าเป็นความคลาดเคลื่อนอย่างเป็นระบบ (Systematic Error) ที่เกิดขึ้นจากการวัด ความพยายามของการตรวจสอบความลำเอียงดังกล่าวดำเนินไปเพื่อจำแนกข้อสอบที่ทำหน้าที่ไม่เหมาะสมหรือไม่ยุติธรรมสำหรับปรับปรุง หรือตัดข้อสอบนั้นออกจากแบบสอบ เป็นการขจัดข้อสอบที่ทำให้เกิดปัญหาความยุติธรรมระหว่างกลุ่มข้อสอบกลุ่มต่าง ๆ ที่มีลักษณะบางอย่างแตกต่างกัน เช่น เชื้อชาติ ศาสนา วัฒนธรรม ภูมิลาเนา สังคม เพศ ภาษา อายุ ประสบการณ์ เป็นต้น เพื่อพัฒนาแบบสอบให้มีคุณภาพเหมาะสมสำหรับนำไปใช้ทดสอบต่อไป (ศิริชัย กาญจนวาสี, 2555)

ในเวลาต่อมา นักวัดผลการศึกษาได้ทำการศึกษาความลำเอียงของข้อสอบ (Item Bias) กันอย่างกว้างขวาง ทำให้เกิดความสับสนของการใช้คำและความหมาย มีประเด็นโต้แย้งกันว่า ความลำเอียงของข้อสอบเป็นผลการตัดสินว่าข้อสอบมีความยุติธรรมหรือไม่ อันส่งผลต่อการบรรลุจุดมุ่งหมายของการใช้แบบทดสอบหรือความลำเอียงของข้อสอบ เป็นสารสนเทศทางสถิติที่ได้จากข้อสอบเกี่ยวกับกลุ่มผู้สอบต่างกลุ่มกันตอบข้อสอบข้อเดียวกัน ความแตกต่างที่เกิดขึ้นอาจมาจากความไม่เหมาะสมของข้อคำถาม ซึ่งสามารถเกิดขึ้นได้หลายลักษณะ หรือประสบการณ์ของผู้สอบ ซึ่งอาจมีลักษณะพื้นฐานเดิมแตกต่างกันในหลายสถานการณ์จึงไม่เหมาะสมที่จะใช้คำว่าข้อสอบลำเอียง (Biased Item) เนื่องจากเป็นภาษาที่มีความหมายในเชิงลบ ประกอบกับเกณฑ์ที่ใช้สำหรับตัดสินความลำเอียงยังมีความคลุมเครือและค่อนข้างสับสน ดังนั้น จึงควรเปลี่ยนมาใช้คำว่าการทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning: DIF) ซึ่งเป็นคำที่มีความเป็นกลางและ ความเหมาะสมกว่า (Holland และ Thayer, 1988 ; Holland และ Wainer, 1993)

จากการศึกษาเอกสารและงานวิจัยผู้วิจัยสามารถสรุปได้ว่า การทำหน้าที่ต่างกันของข้อสอบ แต่เดิมใช้คำว่า “ความลำเอียงของข้อสอบ” (Item Bias) ซึ่งเป็นภาษาที่ใช้กันในทางสังคมและมีความหมายในทางลบ แต่ระยะหลังนักวิจัยได้เปลี่ยนไปใช้คำใหม่ว่า “การทำหน้าที่ต่างกันของข้อสอบ” (Differential Item Functioning : DIF) แต่อย่างไรก็ตามคำสองคำนี้มีจุดเน้นที่แตกต่างกัน คำว่า “ความลำเอียงของข้อสอบ” เน้นที่อิทธิพลที่สังเกตได้ของกลุ่มผู้สอบย่อยที่มุ่งศึกษา ส่วนคำว่า “ข้อสอบที่ทำหน้าที่ต่างกัน” นั้น เน้นคุณลักษณะทางสถิติของข้อสอบที่ตรวจสอบได้ด้วยวิธีการ

วิเคราะห์ทางสถิติ สำหรับสิ่งที่ต้องให้ความสำคัญลำดับต่อมา ได้แก่ ประเภทของการทำหน้าที่ต่างกันของข้อสอบ

2. ประเภทของการทำหน้าที่ต่างกันของข้อสอบ

การทำหน้าที่ต่างกันของข้อสอบ (DIF) เป็นการเปรียบเทียบผลการตอบข้อสอบระหว่างกลุ่มผู้สอบอย่างน้อย 2 กลุ่มขึ้นไป ปกตินิยมทำการเปรียบเทียบ 2 กลุ่ม ประกอบด้วย กลุ่มแรกเรียกว่า กลุ่มเปรียบเทียบ (Focal Group หรือกลุ่ม F) เป็นกลุ่มที่สนใจศึกษาและคาดว่าจะเป็นกลุ่มที่เสียเปรียบในการตอบข้อสอบ และกลุ่มที่สองเรียกว่ากลุ่มอ้างอิง (Reference Group หรือกลุ่ม R) เป็นกลุ่มที่คาดว่าจะได้เปรียบในการตอบข้อสอบได้ถูกต้อง

ในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ จะพบว่า ข้อสอบสามารถทำหน้าที่ต่างกัน ได้ 2 ประเภท (Mellenbergh, 1982) ได้แก่ การทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป (Uniform) และแบบอเนกรูป (Nonuniform) (ศิริชัย กาญจนวาสี, 2555)

1. ข้อสอบทำหน้าที่ต่างกันแบบเอกรูป (Uniform DIF) หมายถึง ข้อสอบที่ทำให้ผู้สอบกลุ่มหนึ่งมีโอกาสในการตอบข้อสอบถูกมากกว่าผู้สอบอีกกลุ่มหนึ่งอย่างสม่ำเสมอกัน ในทุกระดับความสามารถ เมื่อพิจารณาโค้งคุณลักษณะข้อสอบของผู้สอบ 2 กลุ่ม จะไม่พบว่ามีปฏิสัมพันธ์ระหว่างระดับความสามารถของผู้สอบกับการเป็นสมาชิกของกลุ่ม (Group Membership)

2. ข้อสอบทำหน้าที่ต่างกันแบบอเนกรูป (Nonuniform DIF) หมายถึง ข้อสอบที่ทำให้โอกาสในการตอบข้อสอบถูกของผู้สอบระหว่างกลุ่มแตกต่างกันอย่างไม่สม่ำเสมอกันในทุกระดับความสามารถ เมื่อพิจารณาโค้งคุณลักษณะข้อสอบของผู้สอบสองกลุ่ม พบว่า มีปฏิสัมพันธ์ร่วมกันระหว่างความสามารถของผู้สอบกับการเป็นสมาชิกของกลุ่ม เช่น ที่ระดับความสามารถหนึ่งกลุ่มผู้สอบกลุ่ม R มีโอกาสในการตอบข้อสอบถูกมากกว่ากลุ่มผู้สอบกลุ่ม F แต่ที่ระดับความสามารถอีกระดับหนึ่งกลุ่มผู้สอบกลุ่ม F มีโอกาสในการตอบข้อสอบถูกมากกว่ากลุ่มผู้สอบกลุ่ม R

ตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) สามารถพิจารณา “ปฏิสัมพันธ์” ดังกล่าวได้จากความแตกต่างค่าพารามิเตอร์ อำนาจจำแนกของข้อสอบระหว่างผู้สอบกลุ่มย่อยสองกลุ่ม กล่าวคือ ถ้าข้อสอบทำหน้าที่ต่างกันแบบเอกรูป แล้วโค้งลักษณะข้อสอบ (Item Characteristic Curves : ICCs) ระหว่างผู้สอบกลุ่มย่อยสองกลุ่มจะขนานกัน หรือมีฟังก์ชั่นการตอบสนองข้อสอบ (Item Response Functions : IRFs) เหมือนกัน แต่ถ้าข้อสอบมีหน้าที่ต่างกันแบบอเนกรูป แล้วโค้งลักษณะข้อสอบระหว่างผู้สอบกลุ่มย่อยสองกลุ่มจะไม่ขนานกัน หรือมีฟังก์ชันการตอบสนองข้อสอบต่างกัน ดังนั้นความแตกต่างระหว่างโค้งลักษณะข้อสอบทั้งสองแบบจะบ่งบอกถึงขนาดและทิศทางของข้อสอบที่ทำหน้าที่ต่างกัน ซึ่งสามารถคำนวณได้โดยใช้สูตรการคำนวณพื้นที่ของ Raju (1990)

ข้อสอบที่ทำหน้าที่ต่างกันแบบอเนกรูป สามารถจำแนกได้เป็น 2 ลักษณะ (Swaminathan และ Rogers, 1990) ดังนี้

1. ข้อสอบทำหน้าที่ต่างกันแบบอเนกรูป โดยมีปฏิสัมพันธ์ไม่เป็นลำดับ (Disordinal Interaction) เป็นการทำหน้าที่ต่างกันสำหรับกลุ่มผู้สอบ ซึ่งเกิดขึ้นเมื่อโค้งลักษณะข้อสอบตัดกันระหว่างช่วงความสามารถของผู้สอบหรือเรียกว่า ข้อสอบทำหน้าที่ต่างกันแบบไม่มีทิศทาง (Non-Unidirectional DIF)

2. ข้อสอบทำหน้าที่ต่างกันแบบอเนกรูป โดยมีปฏิสัมพันธ์เป็นลำดับ (Ordinal Interaction) เป็นการทำหน้าที่ต่างกันสำหรับกลุ่มผู้สอบ ซึ่งเกิดขึ้นเมื่อโค้งลักษณะข้อสอบต่างกันอย่างไม่สม่ำเสมอ แต่ไม่ตัดกัน หรืออาจตัดกันนอกช่วง ความสามารถของผู้สอบตรงปลายสุดของช่วงความสามารถต่ำ หรือสูง อาจเรียกข้อสอบลักษณะนี้ว่า ข้อสอบทำหน้าที่ต่างกันแบบมีทิศทางเดียว (Unidirectional DIF)

โดยทั่วไปในแบบสอบมาตรฐานมักจะมีข้อสอบที่ต่างกันแบบเอกรูปมากกว่าข้อสอบที่ทำหน้าที่ต่างกันแบบอเนกรูป แต่ในข้อมูลจริงจะมีข้อสอบที่ทำหน้าที่ต่างกันแบบอเนกรูปได้มากกว่า จะเห็นได้ว่าประเภทของการทำหน้าที่ต่างกันของข้อสอบแบ่งเป็น 2 ได้แก่ การทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป (Uniform DIF) เกิดขึ้นเมื่อผู้สอบกลุ่มหนึ่งมีโอกาสในการตอบข้อสอบถูกมากกว่าผู้สอบอีก กลุ่มหนึ่งในทุกระดับความสามารถ และการทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูป (Nonuniform DIF) เกิดขึ้นเมื่อโอกาสในการตอบข้อสอบถูกของผู้สอบระหว่างกลุ่มย่อย 2 กลุ่มไม่สม่ำเสมอ

3. การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

แนวคิดเกี่ยวกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

การศึกษาเรื่องความยุติธรรมของข้อสอบ ในกรณีที่ข้อสอบทำให้ผู้สอบระหว่างกลุ่มย่อยเกิดการได้เปรียบเสียเปรียบกัน เดิมใช้คำว่า “ความลำเอียงของข้อสอบ” (Item Bias) ซึ่งเป็นภาษาที่ใช้กันในทางสังคมและมีความหมายในทางลบ แต่ระยะหลังนักวิจัยได้เปลี่ยนไปใช้คำใหม่ว่า “การทำหน้าที่ต่างกันของข้อสอบ” (Differential item Functioning : DIF) เนื่องจากเห็นว่าเป็นคำที่มีความหมายเป็นกลาง จึงมีความเหมาะสมเชิงวิชาการมากกว่า คำสองคำนี้มีจุดเน้นที่แตกต่างกันที่คำว่า “การทำหน้าที่ต่างกันของข้อสอบ” เน้นที่คุณลักษณะทางสถิติของข้อสอบที่ตรวจสอบได้ด้วยวิธีการวิเคราะห์ทางสถิติ ซึ่งเป็นส่วนประกอบหนึ่งของสิ่งที่แสดงถึงความลำเอียงของข้อสอบ วิธีการทางสถิติที่ใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเป็นเงื่อนไขที่จำเป็นในการตัดสินความลำเอียงของข้อสอบ เนื่องจากถ้าใช้วิธีการทางสถิติตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเพียงอย่างเดียวแล้ว ผลการตรวจสอบพบว่าข้อสอบทำหน้าที่ต่างกันนั้นยังสรุปไม่ได้ว่าข้อสอบมีความลำเอียงหรือไม่ ต้องให้ผู้เชี่ยวชาญพิจารณาเนื้อหาของข้อสอบและจุดมุ่งหมายในการ

วัดข้อสอบที่เรียกว่า “วิธีการตัดสินข้อสอบ” (Judgemental Method) (Camilli และ Shepard, 1994)

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF Detection) เป็นการเปรียบเทียบผลการตอบข้อสอบเป็นรายข้อระหว่างกลุ่มผู้สอบอย่างน้อย 2 กลุ่ม มีความสามารถหลัก (Primary Ability) ที่มุ่งวัดเท่ากัน แต่คาดว่าจะมีความได้เปรียบเสียเปรียบกัน โดยกลุ่มหนึ่งถือเป็นกลุ่มอ้างอิง (Reference Group) ซึ่งคาดว่าน่าจะได้เปรียบในการตอบข้อสอบข้อนั้น หรือมีโอกาสตอบข้อสอบได้ถูกต้องมากกว่า ส่วนอีกกลุ่มคือ กลุ่มเปรียบเทียบ (Focal Group) ซึ่งเป็นกลุ่มที่สนใจศึกษา และคาดว่าน่าจะเป็นกลุ่มที่เสียเปรียบ (ศิริชัย กาญจนวาสี, 2555)

ในการเปรียบเทียบผลการตอบข้อสอบระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบจำเป็นต้องจับคู่ (Matching) ผู้สอบตามความสามารถ ซึ่งเป็นเงื่อนไขสำคัญของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เกณฑ์การจับคู่ (Matching Criteria) ที่นิยมใช้กันมี 2 วิธี ดังนี้

1. เกณฑ์ภายนอก (External Criterion)

การวิเคราะห์การทำหน้าที่ต่างกัน โดยใช้เกณฑ์ภายนอกนี้ สามารถนำไปใช้ได้ทั้งข้อสอบรายข้อและข้อสอบทั้งฉบับ โดยการใช้คะแนนจากแบบสอบอื่นเป็นเกณฑ์ภายนอก แล้วใช้เทคนิคการวิเคราะห์การถดถอย (Regression Analysis) เพื่อทำการเปรียบเทียบเส้นกราฟความสัมพันธ์ระหว่างตัวแปรเกณฑ์กับตัวแปรทำนายระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ

หลักการนี้มีจุดมุ่งหมาย เพื่อสร้างสมการทำนายตัวแปรเกณฑ์ ซึ่งเป็นคะแนนของแบบสอบอื่นจากตัวแปรทำนายซึ่งเป็นคะแนนรายข้อ หรือคะแนนแบบสอบระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบในการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ จะใช้คะแนนรายข้อเป็นตัวแปรทำนาย แต่ถ้าเป็นการวิเคราะห์การทำหน้าที่ต่างกันของแบบสอบจะใช้คะแนนรวมของแบบสอบทั้งฉบับเป็นตัวแปรทำนาย สำหรับตัวแปรเกณฑ์ที่ใช้เป็นเกณฑ์ภายนอกอาจใช้คะแนนรวมทั้งฉบับ หรือเกรดเฉลี่ย หรือผลสัมฤทธิ์ในงานที่เกี่ยวข้องของผู้สอบ (Cronbach, 1970)

การใช้เกณฑ์ภายนอกมีข้อดี คือเกณฑ์ที่ใช้มีความเป็นอิสระจากข้อสอบและแบบสอบที่ต้องการตรวจสอบ แต่มีจุดอ่อนตรงที่ความเหมาะสมของเกณฑ์ที่จะนำมาใช้ในทางปฏิบัติ เป็นการยากที่จะหาตัวแปรเกณฑ์ภายนอกจากแบบสอบฉบับอื่นที่มีความตรงเชิงทำนายและมีความยุติธรรมสำหรับกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ ถ้าตัวแปรเกณฑ์ภายนอกขาดคุณสมบัติดังกล่าว จะทำให้ผลการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบหรือแบบสอบขาดความแม่นยำและความสมบูรณ์

2. เกณฑ์ภายใน (Internal Criterion)

การวิเคราะห์การทำหน้าที่ต่างกัน โดยใช้เกณฑ์ภายในเป็นการนำวิธีการทางสถิติมาตรวจสอบการทำหน้าที่ต่างกันของข้อสอบหรือแบบสอบ โดยเน้นการพิจารณาจากโครงสร้าง

ภายในของแบบสอบเป็นหลัก ด้วยการวิเคราะห์ผลจากการตอบข้อสอบและความสามารถหรือคะแนนจริงของผู้สอบที่ได้จากแบบทดสอบฉบับนั้น เพื่อนำมาเปรียบเทียบระหว่างผู้สอบจากกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่มีความสามารถหรือคะแนนจริงเท่ากันว่า จะมีผลการตอบหรือโอกาสในการตอบข้อสอบได้ถูกต้องแตกต่างกันหรือไม่ เพื่อบ่งชี้การทำหน้าที่ต่างกันของข้อสอบ การวิเคราะห์ในลักษณะนี้นิยมใช้ค่าสถิติต่าง ๆ เป็นตัวบ่งชี้ถึงการทำหน้าที่ต่างกันของข้อสอบ ค่าสถิติทดสอบที่นิยมนามาใช้ พอสรุปได้ดังนี้

2.1 การทดสอบปฏิสัมพันธ์ (Interaction)

ในระยะเริ่มแรกของการศึกษาความลำเอียงของข้อสอบ มีการใช้สถิติทดสอบเอฟ (F-test) จากการวิเคราะห์ความแปรปรวน (ANOVA) เพื่อทดสอบปฏิสัมพันธ์ระหว่างกลุ่มผู้สอบกับข้อสอบ ถ้าการทดสอบมีนัยสำคัญเป็นสัญญาณการทำหน้าที่ต่างกันของข้อสอบ (Cleary และ Hilton, 1968 ; Jensen, 1974) จากนั้นจึงทำการวิเคราะห์ต่อด้วยวิธีการ Post Hoc เพื่อระบุข้อสอบที่มีผลต่อการเกิดปฏิสัมพันธ์ซึ่งเป็นข้อที่ทำหน้าที่ต่างกัน วิธีการนี้มีข้อดีที่สามารถศึกษาผู้สอบหลาย ๆ กลุ่มได้สะดวก แต่มีจุดอ่อนในเรื่องการควบคุมกลุ่มต่าง ๆ ให้มีความสามารถที่ทัดเทียมกันขนาดกลุ่มตัวอย่างของกลุ่มต่าง ๆ และอัตราความคลาดเคลื่อนประเภทที่ 1 จะสูงขึ้นถ้าจำนวนข้อสอบเพิ่มมากขึ้น

2.2 การวัดความเบี่ยงเบนสัมพัทธ์ (Relative Deviation)

การคำนวณค่าความยากของข้อสอบ เช่น p, b เป็นต้น เมื่อคำนวณแยกระหว่างกลุ่ม และแปลงให้เป็นค่าความยากมาตรฐาน สามารถนำมาพล็อตเปรียบเทียบเป็นรายข้อ ถ้าข้อใดเบี่ยงเบนไปจากแกนหลักที่คาดหมายหรือเบี่ยงเบนเกินจากความคลาดเคลื่อนมาตรฐานของค่าความยากที่กำหนด ย่อมแสดงถึงการทำหน้าที่ต่างกันของข้อสอบ (Cleary และ Hilton, 1968 ; Angoff และ Ford, 1973) รวมทั้งสามารถคำนวณค่าสหสัมพันธ์ระหว่างค่าความยากรายข้อระหว่างกลุ่ม เพื่อแสดงถึงการทำหน้าที่ต่างกันของแบบสอบค่าสหสัมพันธ์เข้าใกล้ 1.00 แสดงว่าค่าความยากสัมพัทธ์ของข้อสอบมีค่าใกล้เคียงกันระหว่างกลุ่ม ดังนั้นแบบสอบวัดคุณลักษณะคล้ายกันระหว่างกลุ่ม

วิธีการนี้มีข้อดีและข้อเสียคล้ายการทดสอบปฏิสัมพันธ์ นอกจากนี้ค่าความยากของข้อสอบ (p) มิใช่ตัวแทนของค่าความยากจริงของข้อสอบ และได้รับอิทธิพลจากค่าแทรกซ้อนอื่น ได้แก่ ค่าอำนาจจำแนก และความสามารถของผู้สอบ

2.3 การเปรียบเทียบน้ำหนักตัวประกอบ (Factor Loading)

การวิเคราะห์ตัวประกอบ (Factor Analysis) เป็นเทคนิคทางสถิติที่นิยมใช้ในการตรวจสอบความตรงเชิงทฤษฎีหรือโครงสร้าง (Construct Validity) เมื่อนำการวิเคราะห์ตัวประกอบมาใช้ในการวิเคราะห์โครงสร้างของแบบสอบแยกตามกลุ่มสอบ

ความไม่สอดคล้องกันระหว่างน้ำหนักตัวประกอบบนคุณลักษณะสำคัญที่มุ่งวัด หรือความแตกต่างของค่าเฉลี่ยคะแนน ตัวประกอบ (Factor Scores) ระหว่างกลุ่มผู้สอบ ย่อมสะท้อนการทำหน้าที่ต่างกันของข้อสอบและแบบสอบ

การใช้เทคนิคการวิเคราะห์ตัวประกอบเชิงสำรวจ (Exploratory Factor Analysis : EFA) สำหรับศึกษาการทำหน้าที่ต่างกันจะมีจุดอ่อนในเรื่องความไม่สอดคล้องระหว่างน้ำหนักตัวประกอบ อาจเกิดความแตกต่างของความสามารถระหว่างกลุ่มก็ได้ แนวทางที่เหมาะสมจึงควรใช้เทคนิค การวิเคราะห์ตัวประกอบเชิงยืนยัน (Confirmatory Factor Analysis : CFA) นอกจากนี้ยังสามารถใช้ CFA สำหรับตรวจสอบความแตกต่างระหว่างกลุ่ม ในด้านคุณลักษณะหรือความสามารถหลักและความสามารถรองได้อีกด้วย (Camilli และ Shepard, 1994)

2.4 การเปรียบเทียบโอกาสการตอบข้อสอบถูก

การวิเคราะห์โอกาสตอบข้อสอบถูกของผู้สอบจากกลุ่มอ้างอิงและกลุ่มเปรียบเทียบที่มีความสามารถเท่ากัน เป็นแนวทางสำคัญที่นิยมใช้กันและเป็นที่ยอมรับในปัจจุบันสำหรับบ่งชี้การทำหน้าที่ต่างกันของข้อสอบ มีการคำนวณค่าสถิติ 2 แนวทาง ดังนี้

1. เปรียบเทียบค่าสัดส่วนและความน่าจะเป็นในการตอบข้อสอบถูกของผู้สอบ ต่างกลุ่มที่มีความสามารถเท่ากัน เช่น วิธีแมนเทล-แฮนส์เซล (MH) เป็นต้น

2. เปรียบเทียบค่าฟังก์ชันการตอบสนองข้อสอบ หรือโค้งลักษณะข้อสอบระหว่างกลุ่มที่มีระดับความสามารถเท่ากัน เป็นวิธีที่อยู่บนพื้นฐานของทฤษฎี IRT เช่น วิธีวัดความแตกต่างของพื้นที่ วิธีวัดความแตกต่างค่าพารามิเตอร์ความยาก วิธีทดสอบไค-สแควร์ของลอร์ดเป็นต้น

วิธีการนี้มีข้อดีที่สำคัญได้แก่ การคำนวณค่าสถิติของข้อสอบมีความน่าเชื่อถือ มีกลไกควบคุมความสามารถของผู้สอบโดยการจับคู่กลุ่มความสามารถ เพื่อทำการเปรียบเทียบ ณ ตำแหน่งต่าง ๆ ที่มีความสามารถเท่ากัน จึงเป็นวิธีการที่ยอมรับกันทั่วไป แต่มีความจำกัดในด้านความสลับซับซ้อนของแนวคิดพื้นฐาน และการวิเคราะห์มีความจำเป็นต้องใช้โปรแกรมคอมพิวเตอร์โดยเฉพาะ

หลักการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบจะเป็นการเปรียบเทียบผลการตอบข้อสอบระหว่างผู้สอบสองกลุ่มที่มีระดับความสามารถเดียวกัน โดยกำหนดให้ผู้สอบกลุ่มหนึ่งเป็นกลุ่มอ้างอิงและผู้สอบอีกกลุ่มหนึ่งเป็นกลุ่มเปรียบเทียบ ถ้าข้อสอบทำหน้าที่ต่างกันแล้ว โอกาสในการตอบข้อสอบถูกของผู้สอบแต่ละกลุ่มจะไม่เท่ากัน ต่อไปจะให้ความสำคัญในเรื่องขั้นตอนทางสถิติ นั่นคือ วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

การแบ่งกลุ่มวิธีการทางสถิติที่ใช้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ แบ่งได้หลายวิธี ซึ่ง Hambleton และคณะ (1991) จำแนกวิธีการตรวจสอบการทำหน้าที่ต่างกันออกเป็น 3 กลุ่มใหญ่ ๆ ดังนี้

1. กลุ่มวิธีที่ใช้ทฤษฎีการทดสอบแบบดั้งเดิม (Methods Using Classical Test Theory : CTT) วิธีในกลุ่มนี้พัฒนามาจากหลักการของทฤษฎีการทดสอบแบบดั้งเดิม ใช้คะแนนที่สังเกตได้ของผู้เข้าสอบแต่ละคนเป็นเกณฑ์การจับคู่กลุ่มผู้เข้าสอบย่อย และเปรียบเทียบค่าความยากของข้อสอบแต่ละข้อระหว่างกลุ่มผู้เข้าสอบย่อยที่สนใจศึกษา วิธีการในกลุ่มนี้ ได้แก่ การวิเคราะห์ความแปรปรวน (Analysis of Variance) วิธีสหสัมพันธ์ (Correlational Method) วิธีแปลงค่าความยากของข้อสอบ (Transformed Item Difficulty Method : TID) หรือวิธีการกำหนดจุดเดลต้า (Delta Plot Method) (Angoff, 1982) การวิเคราะห์ตัวลวง (Distractor Analysis) (Scheuneman, 1979) วิธีสหสัมพันธ์บางส่วน (Partial Correlation Methods) (Stricker, 1982) และวิธีการทำให้เป็นมาตรฐาน (Standardization Method) (Dorans และ Kulick, 1986)

ข้อดีของวิธีการกลุ่มนี้ คือ กระบวนการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบไม่ยุ่งยาก เสียค่าใช้จ่ายไม่สูงนัก ใช้ตรวจสอบกับกลุ่มตัวอย่างขนาดเล็กได้ และสามารถอธิบายให้คนทั่วไปเข้าใจได้ง่าย ส่วนข้อเสียก็คือ ค่าสถิติของข้อสอบเปลี่ยนไปตามกลุ่มตัวอย่าง เมื่อกลุ่มตัวอย่างเปลี่ยนไปผลการตรวจพบข้อสอบทำหน้าที่ต่างกันก็เปลี่ยนไป ทำให้การสรุปอ้างอิงผลการศึกษาไปยังกลุ่มประชากรอาจมีความน่าเชื่อถือได้น้อยลง

2. กลุ่มวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ (Methods Using Item Response Theory : IRT) วิธีการในกลุ่มนี้ตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ตามกรอบแนวคิดของทฤษฎีการตอบสนองข้อสอบ โดยปกติแล้วจะใช้ในการเปรียบเทียบเส้นโค้งลักษณะข้อสอบ (Item Characteristic Curves : ICCs) ของกลุ่มผู้เข้าสอบย่อยตามระดับความสามารถของผู้เข้าสอบ ถ้าเส้นโค้งลักษณะข้อสอบของกลุ่มผู้เข้าสอบย่อยสองกลุ่มมีรูปร่างเหมือนกัน แสดงว่าข้อสอบข้อนั้นทำหน้าที่ไม่ต่างกัน แต่ถ้าเส้นโค้งลักษณะข้อสอบของกลุ่มผู้เข้าสอบย่อยสองกลุ่มมีรูปร่างแตกต่างกัน แสดงว่า ข้อสอบข้อนั้นทำหน้าที่ต่างกัน วิธีการในกลุ่มนี้ ได้แก่ วิธี Analysis of fit (Durovic, 1975 cited in Hambleton and others, 1993) วิธี Difficulty Shift (Wright, Mead and Draba, 1976 cited in Hambleton and others, 1993) ซึ่งใช้โมเดล IRT แบบหนึ่งพารามิเตอร์ วิธี IRT Area (Ironson and Subkoviak, 1979 : Raju, 1990) วิธี Two-Stage (Lord, 1980) ซึ่งใช้โมเดล IRT แบบสองหรือสามพารามิเตอร์ วิธี Plot (Hambleton and Rogers, 1991 cited in Hambleton and others, 1993) และวิธี SIBTEST (Shealy and Stout, 1993)

ข้อดีของวิธีการในกลุ่มนี้ คือ การแก้ไขข้อบกพร่องของทฤษฎีการทดสอบแบบดั้งเดิมทำให้ค่าสถิติของข้อสอบไม่เปลี่ยนไปตามกลุ่มตัวอย่างที่สุ่มมาจากประชากรเดียวกัน

การประมาณค่าความสามารถของผู้เข้าสอบเป็นอิสระจากค่าความยากของแบบทดสอบ โมเดลทางคณิตศาสตร์ง่ายต่อการจับคู่เส้นโค้งลักษณะข้อสอบตามระดับความสามารถของผู้เข้าสอบ ทำให้สามารถศึกษาความแตกต่างของผลการตอบข้อสอบตามระดับความสามารถของกลุ่มผู้เข้าสอบย่อยได้ ไม่ต้องมีข้อจำกัดเบื้องต้นเรื่องแบบทดสอบคู่ขนาน ในการหาค่าสัมประสิทธิ์ความเที่ยงของแบบทดสอบ และถ้าผลการตอบข้อสอบของกลุ่มผู้เข้าสอบสอดคล้องกับข้อตกลงเบื้องต้นของโมเดล IRT แล้ว วิธีการในกลุ่มนี้ก็น่าจะเป็นวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่ได้ผลดี เนื่องจากเป็นวิธีที่มีทฤษฎีการตอบสนองข้อสอบสนับสนุนและใช้ค่าประมาณค่าความสามารถที่แท้จริงของผู้เข้าสอบแทน ข้อสอบสลับซับซ้อน เสียค่าใช้จ่ายในการดำเนินการสูง และต้องการกลุ่มตัวอย่างขนาดใหญ่

3. กลุ่มวิธีที่ใช้วิธีไค-สแควร์ (Methods Using Chi-Square Methods) วิธีในกลุ่มนี้บางครั้งก็เรียกว่า กลุ่มวิธีไค-สแควร์ เนื่องจากใช้ค่าสถิติไค-สแควร์ แสดงการทำหน้าที่ต่างกันของข้อสอบ และใช้คะแนนของแบบทดสอบหรือคะแนนของแบบทดสอบที่ทำให้บริสุทธิ์ เป็นเกณฑ์การจับคู่กลุ่มผู้เข้าสอบย่อยสองกลุ่มที่ทำการศึกษา ก่อนการเปรียบเทียบผลการตอบข้อสอบ วิธีการในกลุ่มนี้ ได้แก่ วิธีตารางการณ์จร (Contingency Table Method) (Scheuneman, 1975) วิธีตารางการณ์ปรับใหม่ (Modified Contingency Table Method) (Veale, 1977 cited in Hambleton and others, 1993) วิธีลอก-ลิเนียร์ (Log-Linear Methods) (Mellenbergh, 1982) วิธีแมนเทล-แฮนส์เซล (Mantel-Haenszal Method : MH) (Holland and Thayer, 1986 ; 1988) และวิธีถดถอยโลจิสติก (Logistic Regression Methods : LR) (Swaminathan and Rogers, 1990) และวิธีการวิเคราะห์องค์ประกอบจากัด (Restricted Factor Analysis Methods : RFA) (Oort, 1998)

ข้อดีของวิธีการในกลุ่มนี้ คือ กระบวนการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบไม่ยุ่งยาก เสียค่าใช้จ่ายในการวิเคราะห์ข้อมูลไม่สูง ใช้ขนาดกลุ่มตัวอย่างไม่ใหญ่นัก และบางวิธีมีหลักการที่ดีในการจับคู่กลุ่มผู้เข้าสอบย่อยตามความสามารถของผู้เข้าสอบ และมีการทดสอบนัยสำคัญ ส่วนข้อเสียของวิธีการในกลุ่มนี้ก็คล้าย ๆ กับวิธีที่ใช้ทฤษฎีการทดสอบแบบดั้งเดิม

4. หลักการทำหน้าที่ต่างกันของข้อสอบ (DIF)

ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบดำเนินการโดยเปรียบเทียบผลการตอบของข้อสอบระหว่างผู้สอบ 2 กลุ่ม ที่มีความสามารถระดับเดียวกัน โดยกำหนดให้ผู้สอบกลุ่มหนึ่งเป็น “กลุ่มอ้างอิง” (Reference Group : R) ซึ่งเป็นกลุ่มที่คาดว่าจะได้ผลประโยชน์ในการตอบข้อสอบ คือ มีโอกาสในการตอบข้อสอบถูกมากกว่าอีกกลุ่ม ส่วนอีกกลุ่มเป็น “กลุ่มเปรียบเทียบหรือกลุ่มสนใจ” (Focal Group : F) ซึ่งเป็นกลุ่มที่คาดว่าจะเสียประโยชน์ในการตอบข้อสอบ คือ

มีโอกาสตอบข้อสอบได้ถูกต้องน้อยกว่าผู้สอบอีกกลุ่มหนึ่ง สำหรับเกณฑ์ที่ใช้ในการจำแนกผู้สอบเป็นกลุ่มอ้างอิง มีหลายลักษณะ เช่น เพศ สีผิว เชื้อชาติ ภาษา วัฒนธรรม และภูมิลำเนา เป็นต้น

ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ จะเริ่มต้นด้วยการกำหนดประชากรให้ชัดเจน และแบ่งประชากรนั้นออกเป็น 2 กลุ่ม ดังที่กล่าวข้างต้นมาคือ กลุ่มอ้างอิง (R) ซึ่งคาดว่าจะได้ประโยชน์จากการที่ข้อสอบทำหน้าที่ต่างกัน หรืออาจกล่าวได้ว่าเป็นกลุ่มที่คาดว่าจะได้คะแนนมากกว่าอีกกลุ่มหนึ่ง ทั้ง ๆ ที่มีความสามารถที่แท้จริงเท่ากัน และกลุ่มเปรียบเทียบ (F) เป็นกลุ่มที่คาดว่าจะเสียประโยชน์จากการที่ข้อสอบทำหน้าที่ต่างกัน หรือเป็นกลุ่มที่คาดว่าจะได้คะแนนน้อยกว่ากลุ่มอ้างอิงนั่นเอง หลังจากที่มีการสอบแล้วนำคำตอบที่ได้ไปหาค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าอำนาจจำแนกของข้อสอบ (a) ค่าความยากของข้อสอบ (b) และค่าโอกาสในการเดาของข้อสอบ (c) ดังได้กล่าวมาแล้วว่ามีการแบ่งกลุ่มผู้สอบเป็น 2 กลุ่ม ดังนั้นแต่ละกลุ่มก็มีชุดพารามิเตอร์ของข้อสอบเฉพาะกลุ่มของตน

การทำหน้าที่ต่างกันของข้อสอบดังกล่าวเป็นการพิจารณาจากค่าสถิติซึ่งสามารถคำนวณได้ จากหลายวิธีโดยใช้คะแนนที่สังเกตได้ (Observed Score) และคะแนนที่สังเกตไม่ได้ (Latent Variable)

ลักษณะของข้อสอบโดยทั่วไปที่แสดงการทำหน้าที่ต่างกัน

1. มีเนื้อหาหรือภาษาที่ใช้ในข้อสอบยั่วยุให้ผู้สอบสนใจ โกรธเกิดการโต้แย้งหรือเกิดอารมณ์ ไม่พอใจ

2. เนื้อหาหรือภาษาที่ใช้ในข้อสอบมีความหมายไปทางลบ ดูถูกเหยียดหยามหรือก้าวร้าวต่อผู้ตอบข้อสอบกลุ่มสนใจ

3. เนื้อหาหรือภาษาในข้อสอบแสดงว่าผู้ตอบข้อสอบกลุ่มสนใจมีปมด้อยเกี่ยวกับอำนาจ หรือความเป็นผู้นำ

4. เนื้อหาหรือภาษาในข้อสอบหลาย ๆ ข้อให้ความสนใจเน้นความสำคัญและยกย่องผู้ตอบข้อสอบกลุ่มอ้างอิง

5. เนื้อหาหรือภาษาในข้อสอบมีสารสนเทศเป็นประโยชน์กับกลุ่มอ้างอิงมากกว่ากลุ่มสนใจ

ลักษณะข้อสอบที่แสดงการทำหน้าที่ต่างกันต่อเพศ

1. รูปแบบหรือโครงสร้างของข้อสอบเป็นปัญหาต่อผู้ตอบข้อสอบเพศใดเพศหนึ่งมากกว่าผู้ตอบข้อสอบอีกเพศหนึ่ง

2. เนื้อหาในข้อสอบมีสรรพนามเฉพาะเพศใดเพศหนึ่ง

3. เนื้อหาในข้อสอบกำหนดสถานการณ์ที่ผู้ตอบข้อสอบเพศใดเพศหนึ่งได้รับการฝึกฝนเฉพาะทาง มีความสนใจและมีโอกาสพบเห็นในชีวิตประจำวันมากกว่า

วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

วิธีการในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบมีหลายวิธี สามารถจำแนกได้หลายลักษณะ ขึ้นอยู่กับเกณฑ์ที่ใช้จำแนก เช่น การใช้เกณฑ์การให้คะแนน แบ่งได้เป็น 2 กลุ่มวิธี คือ

1. กลุ่มวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่มีการให้คะแนนเป็นแบบ 2 ค่า (Dichotomous DIF Procedures) กลุ่มนี้ข้อสอบที่ตรวจสอบการทำหน้าที่ต่างกันมีการให้คะแนนเป็นแบบ 0-1 เช่น แบบทดสอบเลือกตอบที่ให้คะแนนตอบถูกเป็น 1 คะแนน และตอบผิดเป็น 0 คะแนน

2. กลุ่มวิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่มีการให้คะแนนแบบหลายค่า (Polychromous DIF Procedures) เช่น ข้อสอบวัดการปฏิบัติ (Performance Test) ข้อสอบที่ให้สร้างคำตอบเอง (Constructed-response Item) ไม่ว่าจะเป็นข้อสอบวัดการอ่าน (Reading Item) หรือการเขียน (Writing Lethem) หรือแบบทดสอบเลือกตอบที่มีการให้คะแนนความรู้บางส่วน เช่น แบบทดสอบเลือกตอบแบบถูกผิด เป็นต้น การใช้เกณฑ์ที่ยึดทฤษฎีของการวิเคราะห์ข้อมูลแบ่งเป็น 2 กลุ่มวิธี คือ กลุ่มวิธีที่ยึดทฤษฎี IRT ที่วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ โดยใช้คะแนนที่สังเกตไม่ได้หรือ ตัวแปรแฝงภายใต้ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) และกลุ่มวิธีที่ไม่ใช้ IRT (Non IRT) กลุ่มนี้จะวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบโดยใช้คะแนนสังเกตได้ภายใต้ทฤษฎี การทดสอบแบบดั้งเดิม (Classical Test Theory) การใช้เกณฑ์ข้อสอบเบื้องต้นของแบบจำลองแบ่ง เป็น 2 กลุ่มวิธี คือ กลุ่มวิธีที่ยึดรูปแบบพาราเมตริก (Parametric Form) การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบมีข้อตกลงเบื้องต้นของแบบจำลองสำหรับอธิบายความสัมพันธ์ระหว่างคะแนนของข้อสอบและการจับคู่ตัวแปร และกลุ่มวิธีที่ยึดรูปแบบนันพาราเมตริก (Nonparametric Form) ซึ่งกลุ่มนี้จะไม่มีข้อตกลงเบื้องต้น

5. การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี MIMIC

แบบจำลองกลุ่มนี้ประกอบด้วยแบบจำลองความสัมพันธ์ทั้งแบบที่มีและไม่มีความคลาดเคลื่อน ในการวัดจะประกอบขึ้นด้วยตัวแปรสังเกตได้ทั้งหมด โดยไม่มีตัวแปรแฝง เขียนรูปสมการได้ดังนี้

$$Y = \beta Y + \Gamma X + \zeta \qquad (4)$$

หรือเขียนในรูปเมทริกซ์ได้ดังนี้

$$[Y] = [BE][Y] + [GA][X] + [Z] \quad (5)$$

เมทริกซ์พารามิเตอร์ LY, LX, TD และ TE จึงมีค่าเป็นศูนย์ทั้งหมด การกำหนดข้อมูลจำเพาะของโมเดลกำหนดรูปแบบและสถานะของเมทริกซ์ GA, BE, PH, c]t PH เท่านั้น โมเดลความสัมพันธ์โครงสร้างเชิงสาเหตุที่มีความคลาดเคลื่อนในการวัดมีตัวแปรครบทุกประเภทได้ตามโมเดลใหญ่ในโปรแกรมลิสเรล เมื่อเขียนในรูปสมการจะประกอบด้วยสมการการวัดสองสมการ และสมการโมเดลโครงสร้างหนึ่งสมการ ดังนี้

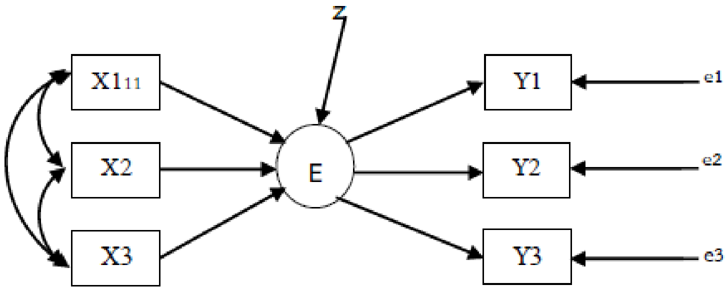
$$[X] = [LX][K] + [d] \quad (6)$$
$$[Y] = [LY][E] + [e] \quad (7)$$
$$[E] = [BE][E] + [GA][K] + [Z] \quad (8)$$

แบบจำลองกลุ่มนี้ยังแบ่งออกได้เป็น 3 แบบ ได้แก่

1. Regression Models and ANOVA Models
2. Path Analysis
3. Multiple Indicators and Multiple Causes Models หรือ MIMIC Models

โมเดลมิมิค (MIMIC Model)

MIMIC เป็นค่าที่ย่อมาจาก Multiple Indicators and Multiple Causes ซึ่งหมายถึง โมเดลลิสเรลที่มีตัวแปรแฝงเพียงตัวแปรเดียว โดยที่ตัวแปรแฝงนั้นได้รับอิทธิพลจากตัวแปรภายนอกสังเกตได้หลายตัวแปร และส่งอิทธิพลไปยังตัวแปรภายในสังเกตได้หลายตัวแปร กล่าวอีกอย่างหนึ่ง คือ เป็นโมเดลลิสเรลของคุณลักษณะแฝงที่มีหลายสาเหตุและวัดได้จากตัวบ่งชี้หลายตัว ดังภาพประกอบ 7 ในที่นี้มีตัวบ่งชี้ 3 ตัวแปร และมีตัวแปรสาเหตุ 3 ตัวแปรตามลักษณะโมเดลจะเห็นว่าการวัดตัวแปรภายนอกสังเกตได้ต้องมีข้อตกลงข้างต้นว่า ไม่มีความคลาดเคลื่อนในการวัดและในการวิเคราะห์ข้อมูลจะกำหนดข้อมูลจำเพาะ เฉพาะรูปแบบและสถานะของเมทริกซ์ PH, BE, GA, PS, LY และ TE เท่านั้น ส่วนเมทริกซ์ TD และ LX มีค่าเป็นศูนย์ทั้งหมด โมเดลมิมิคนี้เป็นประโยชน์มากในการตรวจสอบความเป็นเอกมิติ (Unidimensionality) ในการวิจัยสาขาในการวัดผลการศึกษา



ภาพประกอบ 7 โมเดลย่อยของ MIMIC

การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) สามารถแบ่งออกเป็นการวัดองค์ประกอบและโครงสร้างองค์ประกอบ ในองค์ประกอบการวัด *y*I* ของข้อที่/ลักษณะของตัวแปรแฝง y ที่ทดสอบเป็นการออกแบบการวัด และกลุ่มของตัวแปร z (ในที่นี้เป็นการศึกษาเพียง 1 กลุ่มตัวแปร) ที่เกี่ยวข้องกับการทำหน้าที่ต่างกันของข้อสอบ (DIF) ในการวิเคราะห์องค์ประกอบของโมเดล ดังนี้

สูตรสำหรับ MIMIC ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) คือ

$$y_i^* = \lambda_i \theta + \beta_i' z + \varepsilon_i' \quad (9)$$

เมื่อ

y_i^* คือ ข้อที่ *i*

θ คือ องค์ประกอบ

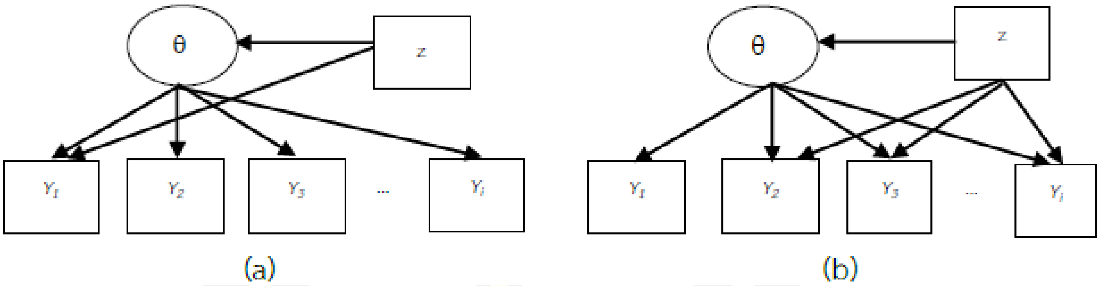
β_i' คือ สัมประสิทธิ์ของตัวแปรเพศและสถานที่ตั้งทางภูมิศาสตร์ของโรงเรียน

z คือ กลุ่มเพศและสถานที่ตั้งทางภูมิศาสตร์ของโรงเรียน

λ_i คือ น้ำหนักองค์ประกอบ

ε_i คือ ค่าความแปรปรวน

เมื่อ λ_i เป็นน้ำหนักองค์ประกอบและเกี่ยวข้องกับความชันของพารามิเตอร์ของข้อสอบข้อที่ *i* ในบริบทของทฤษฎีการตอบสนองข้อสอบ (IRT) แล้ว ε_i มีการแจกแจงแบบปกติสำหรับ Ordinal Probit และการแจกแจงแบบโลจิสติก สำหรับ Ordinal Logit และ β_i' คือ อิทธิพลของกลุ่มตัวแปร Z ต่อ y_i^* ถ้า β_i' = 0 แล้วข้อสอบข้อที่ *i* มีค่าเท่ากันในทุกๆกลุ่ม ตรงกันข้าม ถ้า β_i' ≠ 0 จะเกิดการทำหน้าที่ต่างกันของข้อสอบ (DIF) แบบอเนกรูป เนื่องจากสมการไม่มีเทอมปฏิสัมพันธ์ เป็นตัวทำนาย ดังนั้น สมการ MIMIC จึงใช้แบบเอกรูปได้เพียงอย่างเดียว



ภาพประกอบ 8 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี MIMIC แบบเอกรูป

ซึ่งวิธีหลายตัวชี้วัดหลายสาเหตุในรูปแบบองค์ประกอบเชิงยืนยัน (MIMIC) เป็นหลักการของ CFA กับตัวแปร แล้ววิธี MIMIC ยังสามารถนำไปใช้สาหรับการวิเคราะห์ DIF ได้ด้วย ซึ่งผลที่ได้ต้องมีค่าเป็น แบบ 2 ค่า (Dichotomous) ค่าพารามิเตอร์ของตัวชี้วัดไม่ต่อเนื่องเป็นสิ่งที่จำเป็น ในความเป็นจริงแล้ว มีหลายวิธีที่ตัวชี้วัดของค่าพารามิเตอร์ เป็นแบบ 2 ค่า (Dichotomous) โดยใช้ฟังก์ชันเชื่อมโยงที่เหมาะสม (เช่น การเชื่อมโยงแบบโลจิสหรือโพรบิต) ข้อตกลงเบื้องต้นคือ ตัวแปรแฝงเป็นตัวแปรต่อเนื่องและตัวแปรสังเกตได้เป็นการตอบแบบไบนารี (Binary) เมื่อ y_{ij}^* เป็นตัวแปรแฝงแบบ ต่อเนื่อง และตัวแปรสังเกตได้เป็นการตอบแบบ ไบนารี (Binary) ของข้อสอบ y_{ij} แล้วสามารถเขียนสมการได้ดังนี้

$$y_{ij} = \begin{cases} 1\ if\ \ y_{ij}^* > 0 \\ 0\ if\ \ y_{ij}^* \leq 0 \end{cases} \tag{10}$$

สูตรสำหรับวิธี MIMIC ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) คือ

$$y_{ij}^* = \lambda_i \theta_j + \beta_i G_j + \varepsilon_{ij}, \tag{11}$$

เมื่อ λ_i เป็นน้ำหนักองค์ประกอบของข้อที่ *i* และ θ_j เป็นลักษณะของตัวแปร ส่วน β_iเป็นสัมประสิทธิ์ความชันสำหรับความแปรปรวนร่วม G_i ซึ่งเป็นกลุ่มตัวชี้วัดของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) และ ε_{ij} เป็นเศษเหลือ นอกจากนี้โมเดลการถดถอยเป็นสิ่งจำเป็นสำหรับการพยากรณ์ตัวแปรแฝง θโดยกลุ่มของตัวชี้วัด G_j เพื่อควบคุมความแตกต่างในลักษณะตัวแปรแฝงข้ามกลุ่มย่อย

$$\eta_j = yz_j + \zeta_j, \tag{12}$$

เมื่อ y เป็นความชันของกลุ่มตัวแปร G_j และ ζ_j เป็นความคลาดเคลื่อนของสมการถดถอย β_i เป็นการทำหน้าที่ต่างกันของข้อสอบ (DIF) เป็นเอกรูป เมื่อ y เป็นผลต่างของค่าเฉลี่ยคุณลักษณะแฝงของกลุ่มเปรียบเทียบกับกลุ่มอ้างอิงและมีเกณฑ์การจับคู่ ตามตัวแปรคงที่ในสมการข้างต้นมีข้อตกลงกำหนดให้เป็น 0 ซึ่งจะไม่ปรากฏในสมการข้างต้น

$$\alpha_i = \frac{\lambda_i\sqrt{\sigma_n^2}}{\sqrt{1-\lambda_1^2\sigma_\zeta^2}'} \tag{13}$$

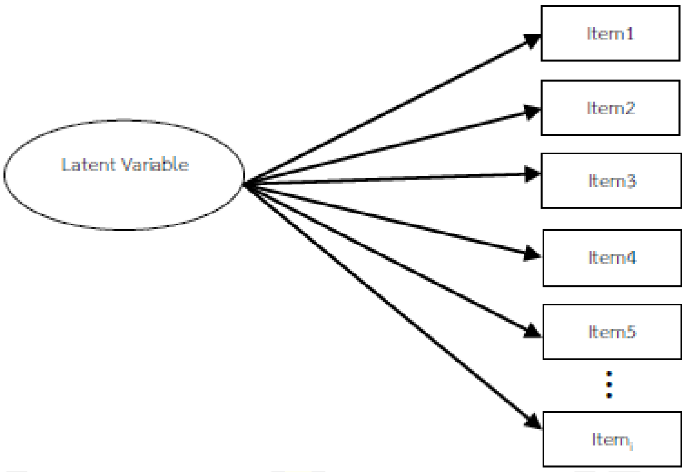
$$b_i = \frac{\lfloor(\tau_i - \beta_i z)\lambda_i{-}^1 - \mu_\eta\rfloor}{(\sigma_\eta^2)^{1/2}}, \tag{14}$$

เมื่อ σ_η^2 เป็นตัวแปรสำหรับองค์ประกอบ θ_j และ σ_ζ^2 เป็นตัวแปรของความคลาดเคลื่อนของสมการถดถอยเชิงเส้นตรง ζ_j สำหรับการทำนายองค์ประกอบทั่วไป τ_i เป็นความยากของข้อสอบ ข้อที่/และ μ_η เป็นค่าเฉลี่ยขององค์ประกอบทั่วไป θ_j

ข้อดีหลายประการของการใช้โมเดล MIMIC ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) ที่แสดงข้างต้น แสดงขนาดของการทำหน้าที่ต่างกันของข้อสอบ (DIF) โดยใช้หลักทฤษฎีการตอบสนองข้อสอบ (IRT) ประมาณค่าการทำหน้าที่ต่างกันของข้อสอบ (DIF) จากค่าพารามิเตอร์ตามทฤษฎีการตอบสนองข้อสอบ (IRT) ซึ่งเป็นประโยชน์ต่อผู้ปฏิบัติ

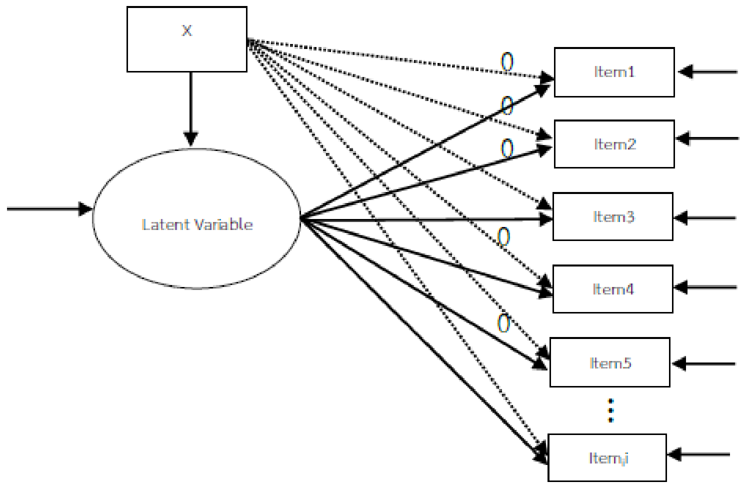
ในการวัดสิ่งต่าง ๆ (Measurement) สิ่งที่น่าสนใจคือความแตกต่างของกลุ่มในตัวแปรแฝง (Latent Variable) ในการศึกษาความแตกต่างของค่าเฉลี่ยของตัวแปรแฝงเป็นการศึกษาความไม่แปรเปลี่ยนของกลุ่ม (Invariant) ในขณะที่การศึกษาความแตกต่างของตัวแปรสังเกตได้ เช่น ค่าเฉลี่ยของข้อคำถามในแต่ละกลุ่ม ซึ่งการศึกษาการทำหน้าที่ต่างกันของข้อสอบ เป็นการศึกษาความแตกต่างของตัวแปรสังเกตได้หรือตัวชี้วัด

โดยปกติรูปแบบการวิเคราะห์ข้อมูลตามทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นการศึกษาคุณสมบัติบนพื้นฐานข้อตกลงความเป็นเอกมิติ (Unidimentsional) ของตัวแปรแฝงซึ่งสังเกตไม่ได้โดยตรง สำหรับตัวแปรแฝงในโมเดล IRT จะดูจากค่าเซต้า (θ) ซึ่งสามารถประมาณค่าได้โดยตรง ซึ่งมีอิทธิพลตรงต่อตัวชี้วัดหรือข้อคำถามที่สังเกตได้ ซึ่งเราสามารถอธิบายโมเดลการวิเคราะห์องค์ประกอบ



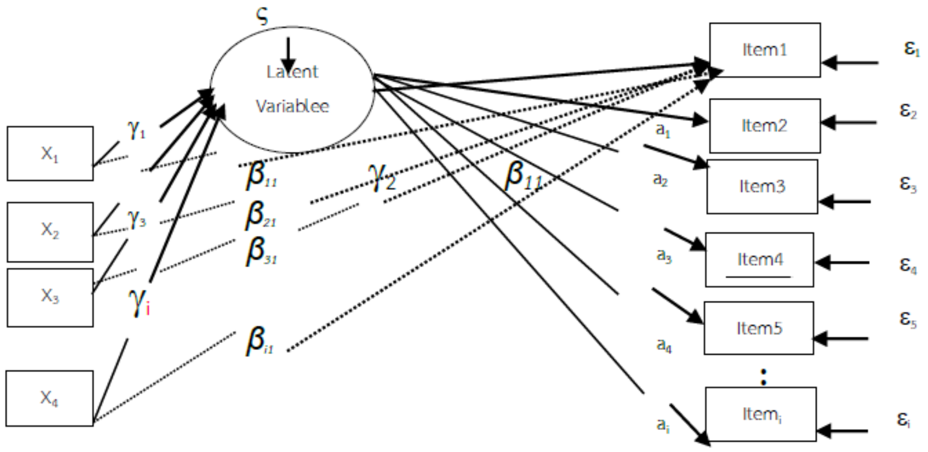
ภาพประกอบ 9 โมเดลการวิเคราะห์องค์ประกอบตามแนวคิด IRT

จากภาพถ้าข้อคำถามหรือตัวชี้วัดเป็นตัวแปรจัดกลุ่ม (Dichotomous) และตัวแปรแฝงมีการแจกแจงแบบปกติ (Normal Distribution) ซึ่งมีลักษณะเช่นเดียวกับการแจกแจงโค้งความถี่สะสมในโมเดล IRT น้ำหนักองค์ประกอบที่เกิดขึ้นบนตัวชี้วัดจะหมายถึง ค่าดัชนีประมาณค่าอำนาจจำแนกตามทฤษฎีการตอบสนองข้อสอบ ในขณะที่ค่าเฉลี่ย (Intercepts) ของแต่ละข้อคำถาม คือค่าประมาณความยาก (Difficulty) ตามทฤษฎีการตอบสนองข้อสอบในกรณีที่มีตัวแปรแฝงมากกว่าหนึ่งตัวโปรแกรมที่พัฒนามาใช้ตามทฤษฎี IRT โดยปกติจะอนุมาน (Assumes) ว่ามีข้อมูลเป็นลักษณะมีความเป็นเอกมิติ (Fleishman, 2003) ในการนำมาประยุกต์ใช้ในการวิจัยจึงทำได้กว้างยิ่งขึ้น จึงง่ายต่อการนำแนวคิดมาประยุกต์ใช้ในกรณีที่ต้องการนำแนวคิดของโมเดล MIMIC มาใช้ในกรณีที่ตัวแปรแฝงมีหลายมิติ (Multi–Dimensional) หลายองค์ประกอบ (Multi–Factor) การนำโปรแกรม MIMIC มาใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ แสดงได้โดยใช้ตัวแปรสาเหตุ (Causes) เพียงตัวเดียว ดังภาพ



ภาพประกอบ 10 โมเดลการวิเคราะห์การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ MIMIC Model โดยใช้ตัวแปรสาเหตุ 1 ตัว

จากภาพประกอบ 10 เป็นการวิเคราะห์องค์ประกอบโดยใช้ตัวแปรทำนาย (X) จำนวนหนึ่งตัวแปร ในการทำนายตัวแปรแฝงที่ประกอบด้วยตัวแปรสังเกตได้ที่เป็นข้อสอบหรือข้อคำถาม (Item) จำนวน I ตัว โดยการจำกัดความคลาดเคลื่อนจากการวัดความคลาดเคลื่อนจากการวัดของตัวแปรแฝง และให้อิสระกับความคลาดเคลื่อนของตัวแปรที่สอดคล้องกับโมเดลมากกว่าการประมาณค่าระหว่าง X กับตัวแปรแฝง (Latent Variable) อิทธิพลตรงของตัวแปร X ที่ทำนาย Item หลังจากที่มีอิทธิพลตรงไปยังตัวแปรแฝงแสดงทิศทางเดียว (Uniform) ในการทำหน้าที่ต่างกันของข้อสอบ (DIF) ซึ่งเป็นสิ่งที่แสดงความลำเอียง (Biased) ที่เกิดจากข้อสอบหรือข้อคำถามหรืออธิบายได้ว่าถ้าข้อสอบหรือข้อคำถาม (Item) ได้รับอิทธิพลอย่างมีนัยสำคัญจากตัวแปรสาเหตุ X แสดงให้เห็นว่าข้อสอบ หรือข้อคำถาม (Item) ขึ้นอยู่กับตัวแปรสาเหตุ X ไม่ได้อธิบายตัวแปรแฝงแสดงว่าข้อสอบหรือข้อคำถาม (Item) ข้อนั้นทำหน้าที่ต่างกันหรือ มีความลำเอียง (Biased) ดังภาพ



ภาพประกอบ 11 โมเดลการวิเคราะห์การตรวจสอบการทาหน้าที่ต่างกันของข้อสอบ MIMIC Model

โดยใช้ตัวแปรสาเหตุมากกว่า 1 ตัว (Riley และ Dennis, 2015)

จากภาพเป็นการวิเคราะห์องค์ประกอบโดยใช้ตัวแปรทำนาย X_1 - X_i ในการทำนายตัวแปรแฝงที่ประกอบด้วยตัวแปรสังเกตได้ที่เป็นข้อสอบหรือข้อคำถาม (Item) จำนวน/ตัวในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) สามารถดำเนินการตามขั้นตอนดังนี้

1. วิเคราะห์องค์ประกอบในโมเดลวัดที่ประกอบด้วยตัวแปรแฝงและข้อคำถาม
2. เพิ่มความแปรปรวนร่วมในการทดสอบโมเดล
3. เพิ่มอิทธิพลตรงกับตัวแปรแฝง (γ) อิทธิพลตรง (a) และกำหนดให้มีค่าเท่ากับศูนย์
4. ตรวจสอบโมเดลดัชนีปรับแก้ (Modification Indices)
5. เพิ่มอิทธิพลตรงจากความแปรปรวนร่วมกับข้อคำถามที่มีค่าดัชนีปรับแก้สูงสุด
6. ดำเนินการในขั้นตอนที่ 4-5 จนกว่าจะไม่พบค่าดัชนีปรับแก้ (M.I) ที่ไม่มีนัยสำคัญ

ประเมินความสอดคล้องของโมเดลและอิทธิพลตรง (β_i)

พหูน ปณ ฑิโต ชีเว

งานวิจัยที่เกี่ยวข้อง

งานวิจัยในประเทศ

รัชนก ยี่สุ่นศรี (2544) ได้ศึกษาการทำหน้าที่ต่างกันของข้อสอบและแบบสอบด้วยกระบวนการ ดี เอฟ ไอ ที สำหรับแบบสอบคัดเลือกบุคคลเข้าศึกษาในสถาบันอุดมศึกษา วิชาภาษาอังกฤษและวิชาคณิตศาสตร์ พบว่า แบบสอบวิชาภาษาอังกฤษทำหน้าที่ต่างกันตามเพศและสถานที่ ตั้งทางภูมิศาสตร์ของโรงเรียนที่จบการศึกษาของผู้สอบ ส่วนแบบสอบวิชาคณิตศาสตร์ทำหน้าต่างกันตามเพศของผู้สอบและผลการวิเคราะห์การทำหน้าที่ต่างกันที่ระดับข้อสอบ พบว่าข้อสอบทำหน้าที่ต่างกันตามเพศของผู้สอบมากที่สุดสอดคล้องกันทั้งสองวิชา นอกจากนี้ผลการเปรียบเทียบคุณภาพของแบบสอบฉบับก่อนและหลังตัดข้อสอบที่พบ DIF ทั้งในกรณีตัดทุกข้อและบางข้อทั้ง 2 วิชาได้ผลสอดคล้องกัน โดยที่แบบสอบฉบับก่อนและหลังตัดข้อสอบมีค่าความตรงเชิงโครงสร้างของแบบสอบไม่แตกต่างกัน แบบสอบฉบับหลังตัดข้อสอบที่พบ DIF ส่วนใหญ่มีค่าความเที่ยงลดลง และแบบสอบฉบับหลังตัดข้อสอบที่ DIF ส่วนใหญ่มีค่าฟังก์ชันสารสนเทศของแบบสอบมีมากขึ้น ส่วนการวิเคราะห์ความสัมพันธ์ระหว่างตำแหน่งของคะแนนรวมของผู้สอบก่อนและหลังตัดข้อสอบที่ พบ DIF ทั้งในกรณีตัดทุกข้อและบางข้อ พบว่า ในทุกกรณีมีความสัมพันธ์ในทางบวกซึ่งกันและกันอย่างมีนัยสำคัญ แบบสอบวิชาภาษาอังกฤษมีค่าสัมประสิทธิ์สหสัมพันธ์สเปียร์แมนแรงค์อยู่ระหว่าง 0.83-0.99 ส่วนในแบบสอบวิชาคณิตศาสตร์มีค่าสัมประสิทธิ์สหสัมพันธ์สเปียร์แมนแรงค์อยู่ระหว่าง 0.87-0.98

สุทธิพร ศุรธณ (2550) ได้ศึกษาความสามารถในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) ตามวิธีตัวแบบเชิงเส้นวางนัยทั่วไประดับลดหลั่น กำหนดให้ผลตอบสนองของข้อสอบแต่ละข้อเป็นแบบทวิภาค (มี ค่า 0 เมื่อตอบไม่ถูกต้อง หรือมีค่าเป็น 1 เมื่อตอบได้ถูกต้อง) ที่มีจำนวนผู้ตอบข้อสอบเท่ากับ 1,000 คน โดยแบ่งออกเป็น 2 กลุ่ม กลุ่มละเท่า ๆ กัน และจำนวนข้อสอบในแบบสอบ 20 ข้อ โดยพิจารณาความสามารถในการตรวจสอบ DIF กรณีที่ผู้สอบมีลักษณะการแจกแจงแบบปกติ และมีลักษณะการแจกแจงแบบฟิชเชอร์ทิปเปทท์ ภายใต้เงื่อนไขของปัจจัยความแตกต่างระหว่างค่าเฉลี่ยความสามารถของผู้สอบทั้งสองกลุ่ม, สัดส่วนของข้อสอบที่มี DIF กำหนดให้มี 3 ระดับ คื 5%, 15% และ 30% ดังนั้น จึงมีข้อสอบที่กำหนดให้ มี DIF จำนวน 1, 3 และ 6 ข้อตามลำดับ, ขนาด DIF ในมาตรวัดโลจิทมี 3 ขนาด คือ 0.3 (ขนาดเล็ก), 0.5 (ขนาดปานกลาง), 0.7 (ขนาดใหญ่) และวิธีการตรวจสอบ DIF 2 วิธี คือ วิธีการตรวจสอบทุกข้อพร้อมกัน และวิธีการตรวจสอบทีละข้อ การศึกษาจะใช้การจำลองข้อมูลด้วยเทคนิคมอนติ คาล์โล โดยใช้โปรแกรม R เวอร์ชัน 2.01 การวิเคราะห์จะใช้ขั้นตอนวิธี HLM2 โดยใช้โปรแกรมสำเร็จรูป HLM เวอร์ชัน 6.0 โดยในแต่ละกรณีจะทำซ้ำ จำนวน 100 ครั้ง และในแต่ละครั้งที่ดำเนินการ

จะคำนวณหาค่า ค่าประมาณของ DIF ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) กำลังการทดสอบ และอัตราความผิดพลาดประเภทที่ 1 สามารถสรุปได้ดังนี้

กรณีที่ความสามารถของผู้สอบทั้งสองกลุ่มมีการแจกแจงแบบปกติ พบว่า ความแตกต่างระหว่างค่าเฉลี่ยของความสามารถของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบไม่มีผลต่อค่า RMSE ค่าเฉลี่ยประมาณของ DIF กำลังการทดสอบ และอัตราความผิดพลาดประเภทที่ 1, วิธีการตรวจสอบทีละข้อจะขึ้นอยู่กับสัดส่วนของข้อสอบที่มี DIF โดยที่ค่า RMSE จะมีค่ามากขึ้นเมื่อสัดส่วนของข้อสอบที่มี DIF สูงขึ้น แต่ในทางกลับกันสัดส่วนของข้อสอบที่มี DIF ไม่มีผลต่อวิธีการตรวจสอบทุกข้อพร้อมกัน และวิธีการตรวจสอบทีละข้อยังมีค่า RMSE น้อยกว่าวิธีการตรวจสอบทุกข้อพร้อมกัน กรณีที่มีข้อสอบที่มี DIF จำนวน 1 ข้อ และ 3 ข้อ แต่กรณีที่มีข้อสอบมี DIF จำนวน 6 ข้อ วิธีการตรวจสอบทีละข้อยังมีค่า RMSE มากกว่าวิธีการตรวจสอบทุกข้อพร้อมกัน และค่าเฉลี่ยประมาณของมี DIF จะขึ้นอยู่กับวิธีการที่ใช้ในการตรวจสอบมี DIF โดยที่วิธีการตรวจสอบทีละข้อให้ค่าเฉลี่ยต่ำกว่าค่าจริง ขณะที่วิธีการตรวจสอบทุกข้อพร้อมกันให้ค่าเฉลี่ยทั้งต่ำกว่าและสูงกว่าค่าจริง ส่วนกำลังการทดสอบและอัตราความผิดพลาดประเภทที่ 1 ขึ้นอยู่กับที่อยู่กับวิธีที่ใช้ในการตรวจสอบและสัดส่วนของข้อสอบที่มี DIF ซึ่งวิธีการตรวจสอบทีละข้อมีกำลังการทดสอบสูงกว่าวิธีการตรวจสอบทุกข้อพร้อมกัน กรณีที่มีข้อสอบมี DIF จำนวน 1 ข้อ และ 3 ข้อ แต่กรณีที่มีข้อสอบมี DIF จำนวน 6 ข้อ วิธีการตรวจสอบทีละข้อยังมีกำลังการทดสอบน้อยกว่าวิธีการตรวจสอบทุกข้อพร้อมกัน กำลังการทดสอบของทั้งสองวิธีไม่ขึ้นอยู่กับสัดส่วนของข้อสอบที่มี DIF, อัตราความผิดพลาดประเภทที่ 1 ของวิธีการตรวจสอบทีละข้อส่วนใหญ่มีค่าอยู่นอกช่วงที่กำหนดและยังมีค่ามากกว่าวิธีการตรวจสอบทุกข้อพร้อมกัน

กรณีที่ความสามารถของผู้สอบทั้งสองกลุ่มมีการแจกแจงแบบฟิชเชอร์ทิปเปทท์ พบว่า ความแตกต่างระหว่างค่าเฉลี่ยของความสามารถของกลุ่มอ้างอิงและกลุ่มเปรียบเทียบไม่มีผลต่อค่าเฉลี่ยของค่าประมาณของ DIF และกำลังการทดสอบ แต่มีผลต่อค่า RMSE และอัตราความผิดพลาดประเภทที่ 1 โดยที่วิธีการตรวจสอบทีละข้อมีค่าเฉลี่ยของค่าประมาณ DIF ต่ำกว่าค่าจริง และมีค่ากำลังการทดสอบสูงกว่าวิธีการตรวจสอบทุกข้อพร้อมกัน เมื่อเปรียบเทียบทั้งกรณีที่ความสามารถของผู้สอบมีการแจกแจงแบบปกติและการแจกแจงแบบฟิชเชอร์ทิปเปทท์ พบว่า กำลังการทดสอบกรณีที่ความสามารถของผู้สอบมีการแจกแจงแบบปกติมีค่ามากกว่ากรณีที่ความสามารถของผู้สอบมีการแจกแจงแบบฟิชเชอร์ทิปเปทท์ ในกรณีที่มีข้อสอบมี DIF จำนวน 6 ข้อ สำหรับวิธีการตรวจสอบทีละข้อมีค่า RMSE มากกว่าวิธีการตรวจสอบทุกข้อพร้อมกันในกรณีที่มีความแตกต่างระหว่างค่าเฉลี่ยของความสามารถของผู้สอบทั้งสองกลุ่ม และกรณีที่มีข้อสอบมี DIF จำนวน 6 ข้อ อย่างไรก็ตามวิธีการตรวจสอบทุกข้อพร้อมกัน กรณีที่ความสามารถของผู้สอบมีการแจกแจงแบบฟิชเชอร์ทิปเปทท์ มีค่า RMSE มากกว่ากรณีที่ความสามารถของผู้สอบมีการแจกแจง

แบบปกติ สำหรับวิธีการตรวจสอบทุกข้อพร้อมกันในกรณีที่ไม่มีความแตกต่างระหว่างค่าเฉลี่ยของความสามารถของผู้สอบทั้งสองกลุ่มมีค่าอัตราความผิดพลาดประเภทที่ 1 น้อยกว่ากรณีที่มีความแตกต่างระหว่างค่าเฉลี่ยของความสามารถของผู้สอบทั้งสองกลุ่ม ในขณะที่วิธีการตรวจสอบทีละข้อมีค่าอัตราความผิดพลาดประเภทที่ 1 อยู่นอกช่วงที่กำหนดทั้งในกรณีที่มีความแตกต่างและไม่มีความแตกต่างระหว่างค่าเฉลี่ยของความสามารถของผู้สอบทั้งสองกลุ่ม

ธเกียรติกมล ทองงอก, โชติกา ภาษีผล และศิริชัย กาญจวาสี (2556) ได้ซึ่งศึกษาประสิทธิภาพการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในวิธีถดถอยโลจิสติกโดยใช้เกณฑ์ขนาดอิทธิพล 2 วิธี สำหรับข้อสอบที่มีรูปแบบการตรวจให้คะแนนแบบทวิวิภาค โดยใช้ข้อมูลจากการจำลองข้อมูลและข้อมูลเชิงประจักษ์ พบว่า วิธีถดถอยโลจิสติกโดยการวัดขนาดอิทธิพลตามเกณฑ์ Jodoin และ Gierl มีอัตราความถูกต้องในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบสูงกว่าเกณฑ์ Zumbo และ Thomas นอกจากนี้ข้อสอบที่ทำหน้าที่ต่างกันแบบอเนกรูป แบบสอบที่มีจำนวนข้อสอบทำหน้าที่ต่างกันทั้งฉบับคิดเป็นร้อยละ 20 มีอัตราความถูกต้องจากการวัดขนาดอิทธิพลทั้ง 2 เกณฑ์สูงกว่าในแบบสอบที่มีจำนวนข้อสอบที่ทำหน้าที่ต่างกัน ทั้งฉบับคิดเป็นร้อยละ 10 และเมื่อขนาดอิทธิพลของข้อสอบที่การทำหน้าที่ต่างกันเพิ่มขึ้นมีผลทำให้อัตราความถูกต้องจากการวัดขนาดอิทธิพลทั้ง 2 เกณฑ์เพิ่มขึ้นภายใต้เกือบทุกเงื่อนไข และผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในข้อมูลเชิงประจักษ์พบว่า ขนาดอิทธิพลตามเกณฑ์ Jodoin และ Gierl ให้อัตราความถูกต้องสูงกว่าและมีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าเกณฑ์ของ Zumbo และ Thomas เมื่อข้อมูลเชิงประจักษ์มีประชากรขนาดใหญ่สามารถตรวจพบข้อสอบที่ทำหน้าที่ต่างกันด้วยการทดสอบระดับนัยสำคัญอย่างมีนัยสำคัญ ส่งผลให้ความคลาดเคลื่อนประเภทที่ 1 มีแนวโน้มสูงขึ้น

พนัส จันทร์เปล่ง (2557) ได้ศึกษาเปรียบเทียบโมเดลประเมินคุณภาพการจัดการศึกษาวิทยาศาสตร์: การประยุกต์ใช้โมเดลมูลค่าเพิ่มพหุระดับที่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและแบบสอบ การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อ 1) เพื่อศึกษาผลการทำหน้าที่ต่างกันของข้อสอบ (DIF) ที่มีการตรวจให้คะแนนทวิวิภาคและพหุวิภาค และการทำหน้าที่ต่างกันของแบบทดสอบ (DTF) ในแบบทดสอบการประเมินการรู้เรื่องวิทยาศาสตร์ 2) เพื่อเปรียบเทียบประสิทธิภาพของโมเดลประเมินคุณภาพการจัดการศึกษาวิทยาศาสตร์เมื่อใช้โมเดลมูลค่าเพิ่มพหุระดับ 2 โมเดล คือ โมเดลที่ไม่ได้พิจารณาการทำหน้าที่ต่างกันของข้อสอบและแบบสอบ (โมเดลที่ 1) กับโมเดลที่มีการพิจารณาการทำหน้าที่ต่างกันของข้อสอบและแบบสอบโดยการตัดข้อสอบที่มีการทำหน้าที่ต่างกันของข้อสอบและการทำหน้าที่ต่างกันเป็นขั้นที่แตกต่างกัน (โมเดลที่ 2) ซึ่งในการเปรียบเทียบประสิทธิภาพของโมเดลประเมินคุณภาพการจัดการศึกษาวิทยาศาสตร์

มีวัตถุประสงค์ย่อยสองข้อ คือ 2.1) เพื่อศึกษาความสัมพันธ์ระหว่างคะแนนมูลค่าเพิ่มกับ 3 ตัวแปร คือ เศรษฐานะของครอบครัว แหล่งทรัพยากรการศึกษาที่บ้าน และความมั่งคั่งของครอบครัว ทั้งก่อนและหลังตัดข้อสอบที่ทำหน้าที่ต่างกัน และ 2.2) เพื่อเปรียบเทียบความสอดคล้องของการจัดอันดับคุณภาพและการจัดกลุ่มคุณภาพของโมเดลมูลค่าเพิ่มพหุระดับทั้ง 2 โมเดล ข้อมูลที่ใช้ในการวิจัยครั้งนี้เป็นข้อมูลทุติยภูมิจากโครงการประเมินผลนักเรียนนานาชาติเกี่ยวกับการรู้เรื่องวิทยาศาสตร์ (PISA) ปี ค.ศ. 2009 จำนวน 9 ฉบับ ของกลุ่มตัวอย่างในประเทศไทย ซึ่งประกอบด้วยนักเรียนจำนวน 4,292 คน จากสถานศึกษา 230 แห่ง ผลการวิจัยที่สำคัญพบว่า

1. การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ พบว่า มีข้อสอบที่เอนเอียงเข้าข้างนักเรียนหญิง (3 ข้อ), นักเรียนที่เรียนพิเศษวิทยาศาสตร์นอกสถานศึกษา (1 ข้อ), นักเรียนที่มาจากครอบครัวที่มีเศรษฐานะต่ำ (2 ข้อ), นักเรียนที่มาจากครอบครัวที่มีแหล่งทรัพยากรที่บ้านสูง (5 ข้อ), และนักเรียนที่มาจากครอบครัวที่มีความมั่งคั่งสูง (2 ข้อ) ค่าความเที่ยงของแบบทดสอบในโมเดลที่ 2 ลดลงเมื่อข้อสอบที่ตรวจพบว่าทำหน้าที่ต่างกันถูกตัดออก ส่วนผลการวิเคราะห์การทำหน้าที่ต่างกันของแบบสอบ พบว่า ขนาดอิทธิพลการทำหน้าที่ต่างกันของแบบสอบ 2 ฉบับ ลดลงหลังจากตัดข้อสอบที่ตรวจพบการทำหน้าที่ต่างกันออกไป ส่วนแบบสอบอีก 7 ฉบับ มีขนาดอิทธิพลการทำหน้าที่ต่างกันของแบบสอบไม่เปลี่ยนแปลง และหลังจากตัดข้อสอบบางข้อที่ตรวจพบการทำหน้าที่ต่างกันออกไปในโมเดลที่ 2 พบว่า ผลการวิเคราะห์องค์ประกอบเชิงยืนยันทั้งแบบสอบรวมและแยกรายฉบับมีความตรงเชิงโครงสร้างที่ยอมรับได้

2. ผลการศึกษาสัมประสิทธิ์การทำนาย (R2) พบว่า โมเดลที่ 1 มีประสิทธิภาพการทำนายมากกว่าโมเดลที่ 2 โดยที่สัมประสิทธิ์การทำนาย (R2) ของโมเดลที่ 1 เท่ากับ .49946 ซึ่งสูงกว่า โมเดลที่ 1 (R2 เท่ากับ .49507)

2.1 ผลการศึกษาความสัมพันธ์ระหว่างคะแนนมูลค่าเพิ่มกับ 3 ตัวแปรควบคุม พบว่า คะแนนมูลค่าเพิ่มกับทั้ง 3 ตัวแปร มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และคะแนนมูลค่าเพิ่มกับทั้ง 3 ตัวแปร ภายหลังตัดข้อสอบที่ตรวจพบการทำหน้าที่ต่างกันออกไปแล้วในการวิเคราะห์มูลค่าเพิ่มมีขนาดลดลงเมื่อเปรียบเทียบกับการวิเคราะห์ที่ไม่ได้ตัดข้อสอบที่ตรวจพบการทำหน้าที่ต่างกันออกไป

2.2 ผลการเปรียบเทียบความสอดคล้องของการจัดอันดับคุณภาพโรงเรียนจากคะแนนมูลค่าเพิ่มโดยสถิติทดสอบ Wilcoxon Signed Ranks พบว่า โมเดลที่ 1 และโมเดลที่ 2 มีการจัดอันดับคุณภาพคะแนนมูลค่าเพิ่มไม่สอดคล้องกัน ส่วนผลการเปรียบเทียบความสอดคล้องของการจัดกลุ่มคุณภาพโรงเรียน 5 กลุ่ม ตามคะแนนมูลค่าเพิ่ม พบว่า โมเดลที่ 1 กับโมเดลที่ 2 มีการจัดกลุ่มคุณภาพโรงเรียนที่สอดคล้องกันอย่างมีนัยสำคัญทางสถิติ (Cohen’s Kappa = .865, p<.05)

ภาวิณี ประทุมสาย (2559) ได้ศึกษาการพัฒนาแบบทดสอบการรู้เรื่องวิทยาศาสตร์ โดยประยุกต์ใช้โมเดลการตอบสนองข้อสอบแบบพหุมิต การวิจัยครั้งนี้มีความมุ่งหมายเพื่อ 1) เพื่อสร้างแบบทดสอบการรู้เรื่องวิทยาศาสตร์ของนักเรียน โดยประยุกต์ข้อสอบแบบพหุมิติ 2) เพื่อหาคุณภาพของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ของนักเรียน โดยประยุกต์ใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ กลุ่มตัวอย่างในงานวิจัยครั้งนี้ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2558 ในภาคตะวันออกเฉียงเหนือ จำนวน 3,000 คน จากการสุ่มแบบหลายขั้นตอน (Multi–Stage Random Sampling) เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบการรู้เรื่องวิทยาศาสตร์ (Scientific Literacy Test) จำนวน 85 ข้อ ในการวิเคราะห์ข้อมูลใช้วิธีการวิเคราะห์องค์ประกอบเชิงยืนยัน โดยประยุกต์ใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ ซึ่งใช้โมเดลปกติสะสมแบบพหุมิติ 2 พารามิเตอร์ ด้วยโปรแกรม NOHARM และ IRTPRO

ผลการวิจัยพบว่า แบบทดสอบการรู้เรื่องวิทยาศาสตร์ โดยประยุกต์โมเดลการตอบสนองข้อสอบแบบพหุมิติมีข้อสอบ จำนวน 76 ข้อ ที่ผ่านตามเกณฑ์ โดยมีจำนวนข้อสอบที่วัด 2 มิติ มากที่สุด จำนวน 73 ข้อ รองลงมา คือ ข้อสอบที่วัด 3 มิติ จำนวน 3 ข้อ ค่าอำนาจจำแนกในมิติที่ 1 มีค่าอยู่ระหว่าง 0.136 ถึง 1.95 มิติที่ 2 มีค่าอำนาจจำแนกอยู่ระหว่าง -0.126 ถึง 2.219 และมิติที่ 3 มีค่าอำนาจจำแนกอยู่ระหว่าง 0.038 ถึง 0.323 ค่าจุดตัดของความยากแบบพหุมิติ (d) มีค่าอยู่ระหว่าง -0.786 ถึง 0.49 ค่าอำนาจจำแนกแบบพหุมิติ (MDISC) อยู่ระหว่าง 0.146 ถึง 2.721 ค่าความยากแบบพหุมิติ (MDIFF) มีค่าอยู่ระหว่าง -3.958 ถึง 2.666 ค่าอำนาจจำแนกพหุมิติของข้อสอบมีค่าเฉลี่ย 0.965 พหุมิติต่ำสุดมีค่า 0.146 ค่าอำนาจจำแนกพหุมิติสูงสุดมีค่า 2.721 โดยค่าเฉลี่ยอำนาจจำแนกต่ำสุดในมิติที่ 1, 2, 3 มีค่า 0.743, 0.572 และ 0.006 ตามลำดับ และค่าอำนาจจำแนกต่ำสุดในมิติที่ 1, 2, 3 มีค่า 0.136, -0.126, และ .038 ตามลำดับ ค่าอำนาจจำแนกสูงสุดในมิติที่ 1, 2, 3 มีค่า 1.95, 2.219 และ 0.323 ตามลำดับ ค่าความยากพหุมิติของข้อสอบมีค่าเฉลี่ย -0.456 ค่าความยากพหุมิติต่ำสุดมีค่า -3.958 ค่าความยากพหุมิติสูงสุดมีค่า 2.666 และตัดข้อสอบที่มีค่าความยากแบบพหุติ (MDIFF) เกิน 4.00 ออก 2 ข้อ คือ ข้อที่ 69 และ 79 ค่าฟังก์ชันสารสนเทศของข้อสอบ (Item Information) มีค่าความสามารถสูงสุดอยู่ในช่วง -3 ถึง 3 ค่าฟังก์ชันสารสนเทศของข้อสอบ ณ ระดับความสามารถสูงสุดอยู่ในช่วง -1.2 ถึง -0.6 มีค่าสารสนเทศของแบบทดสอบ (Test Information) ข้อสอบ ณ ระดับความสามารถเท่ากับ 0 ได้ค่าสารสนเทศของแบบทดสอบคือ 51.2

งานวิจัยต่างประเทศ

Holmes, Finch และ French (2007) ได้ตรวจสอบวิธีการทำหน้าที่ต่างกันแบบตัดกันของข้อสอบ โดยเปรียบเทียบจากวิธี 4 วิธี ได้แก่ วิธีซิบเทสท์ (SIBTEST) วิธีการถดถอยโลจิสติก (LR) วิธีการทดสอบอัตราส่วนไลค์ลิฮู้ดตามทฤษฎีการตอบสนองข้อสอบ (IRTLR) และการวิเคราะห์

องค์ประกอบเชิงยืนยัน (CFA) โดยองค์ประกอบที่ศึกษาประกอบไปด้วย ขนาดกลุ่มตัวอย่าง (Sample Size) ความสามารถที่แตกต่างระหว่างกลุ่ม (Ability Differences Between Groups) ร้อยละของการเกิดการทำหน้าที่ต่างกันของข้อสอบ (Percentage of DIF) โดยศึกษาการจำลองข้อมูล ซึ่งผลการวิจัยพบว่า ทุกวิธีสามารถควบคุมความคลาดเคลื่อนประเภทที่หนึ่ง (Type I error) แต่วิธีซิบเทสท์ (SIBTEST) มีอำนาจในการตรวจสอบสูงสุด

Kim และคณะ (2007) ได้ศึกษาการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบและขนาดอิทธิพลในข้อสอบที่มีการตรวจให้คะแนนพหุวิภาค โดยใช้ข้อมูลจากการประเมินความสามารถระดับชาติ โดยใช้วิธีการตรวจจับการทำหน้าที่ต่างกันของข้อสอบ 4 วิธี โดยที่ 2 วิธี เป็นวิธีพื้นฐานตามโมเดลทฤษฎีการตอบสนองข้อสอบ (IRT Model-based Method) คือวิธี Likelihood Ratio Test for IRT และ วิธี Logistic Regression Likelihood Ratio Test และอีก 2 เป็นวิธีตามแนวคิดของแมนเทล–แฮนส์เซล คือ วิธี Mantel Test และวิธี Generalized Mantel-haenszel Test จากผลการวิจัยชี้ให้เห็นว่า เนื่องจากอำนาจการทดสอบทางสถิติเป็นการหน้าที่ของกลุ่มตัวอย่าง การตรวจจับการทำหน้าที่ต่างกันของข้อสอบมาจากกลุ่มข้อมูลที่มีขนาดใหญ่ไม่ได้ใช้ประโยชน์ในเชิงปฏิบัติ และ Indices of Standardized และ Observed-score Indices of Standardized มีผลต่อข้อสอบที่มีการตรวจให้คะแนนพหุวิภาคซึ่งได้มาจากเปรียบเทียบ ขนาดอิทธิพลจากค่า R^2 ตาม Logistic Regression

Penfield (2007) ได้ศึกษาวิธีการสำหรับการจัดกลุ่มการทำหน้าที่ต่างกันของข้อสอบในข้อสอบที่มีการตรวจให้คะแนนพหุวิภาคของหน่วยงานให้บริการการทดสอบทางการศึกษา (ETS) มีการจัดระดับการทำหน้าที่ต่างกันในข้อสอบซึ่งมีพื้นฐานมาจากการแปลงค่าจาก อัตราส่วนเลขออดตาวิธีแมนเทล-แฮนส์เซล (Mantel-haenzel Common Odds Ratio : $\hat{\lambda}$MH) โดยที่การวิจัยครั้งนี้มีจัดกลุ่มแบบแผนการทำหน้าที่ต่างของข้อสอบ สำหรับข้อสอบที่มีการตรวจให้คะแนนพหุวิภาคแบ่งเป็น 2 แบบ คือ แบบแผน P1 (ข้อสอบที่มีการเดา) และแบบแผน P2 (ข้อสอบที่ไม่มีการเดา) ผลจากการศึกษาทั้ง 2 แบบสอดคล้องกับเกณฑ์จากหน่วยงานให้บริการ การทดสอบทางการศึกษา ซึ่งใช้ $\hat{\lambda}$MH เป็นตัวประมาณค่าผลการทำหน้าที่ต่างกันของข้อสอบ (the DIF effect estimator) และทั้งนี้ทั้ง 2 รูปแบบที่ศึกษาสามารถนำไปประยุกต์ใช้กับข้อสอบที่มีการตรวจให้คะแนนพหุวิภาค

Penfield และ GÜler (2009) ได้เปรียบเทียบวิธีการถดถอยโลจิสติก (LR) และวิธีแบบตารางเลขจรณ์ คือ วิธีแมนเทล-แฮนส์เซล (MH) และวิธีเบรสโรล-เดย์ (BD) สำหรับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่มีลักษณะแบบเอกรูป (Uniform DIF) และอเนกรูป (Nonuniform DIF) โดยใช้การศึกษาจากการจำลองข้อมูลและเปรียบเทียบความคลาดเคลื่อนประเภทที่หนึ่ง (Type I error) และอำนาจของกฎผสานการตัดสินใจ (power of a combined decision rule)

ซึ่งเป็นกฎที่เกิดจากการรวมระหว่างวิธีเบรสโรล-เดย์ (BD) และวิธีแมนเทล-แฮนส์เซล (MH) และเปรียบเทียบกับวิธีการถดถอยโลจิสติก (LR) ซึ่งผลการวิจัยพบว่า อัตราความคลาดเคลื่อนประเภทที่หนึ่ง (Type I error rate) ตามกฎผสมการตัดสินใจ (a combined decision rule) อยู่ในระดับความคงที่ที่ต่ำในระดับแอลฟาเชิงกลุ่ม (the nominal alpha level) ส่วนในวิธีการถดถอยโลจิสติก (LR) อยู่ในระดับที่สูงสำหรับเงื่อนไขที่มีการแจกแจงความสามารถไม่เท่ากัน (unequal ability distribution) นอกจากนี้อำนาจในการทดสอบตามกฎผสานการตัดสินใจ (power of a combined decision rule) อยู่ในระดับความคงที่ที่สูงมากกว่าวิธีการถดถอยโลจิสติก (LR) ในทุก ๆ รูปแบบที่มีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ

Carvajal และ Skorupski (2010) ได้ศึกษาผลของขนาดกลุ่มตัวอย่างขนาดเล็กในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในข้อสอบที่มีการตรวจให้คะแนนพหุวิภาค โดยใช้ตัวประมาณค่า liu-agresti ของอัตราส่วนเลขออดสะสม (cumulative common odds ratio) ทั้งนี้ตัวประมาณค่า liu-agresti ถูกนำเสนอโดย Penfield และ Algina ซึ่งเป็นวิธีที่ใช้ศึกษาการทำหน้าที่ต่างกันของข้อสอบของข้อสอบที่มีการตรวจให้คะแนนพหุวิภาค โดยการวิจัยครั้งนี้ศึกษากับกลุ่มตัวอย่างขนาดเล็กเพื่อวิเคราะห์ตัวประมาณค่าดังกล่าว พิจารณาจากขนาดอิทธิพล (effect size) ความคลาดเคลื่อนประเภทที่หนึ่ง (Type I error) และอัตราอำนาจทดสอบ (power rates) ผลการวิจัยชี้ให้เห็นว่า กลุ่มตัวอย่างที่มีขนาดเล็กกว่า 200 มีอำนาจการทดสอบน้อยมาก ซึ่งสังเกตได้จากการทดสอบนัยสำคัญทางสถิติ อย่างไรก็ตาม อัตราความคลาดเคลื่อนประเภทที่หนึ่งเข้าใกล้ในระดับเชิงกลุ่ม (nominal level) และพบว่า log odd ratio เป็นขนาดอิทธิพลที่ไม่ผลในเชิงเปรียบเทียบสำหรับข้อสอบที่มีการจำแนกสูง

Penfield (2010 a) ได้ศึกษาโมเดลผลการทำหน้าที่ต่างกันของข้อสอบโดยใช้ผลความไม่แปรระดับของตัวลวง (Distractor-Level Invariance) ซึ่งเป็นนัยยะสำคัญสำหรับการทำความเข้าใจของสาเหตุการทำหน้าที่ต่างกันของข้อสอบ (DIF) ซึ่งแสดงว่าความไม่แปรเปลี่ยนการวัดข้ามทุก ๆ ตัวเลือกในข้อสอบประเภทหลายตัวเลือก (multiple-choice item) สามารถใช้ nominal response model (NRM) ที่ประกอบด้วย ผลการทำหน้าที่ต่างกันของตัวลวง สำหรับตัวลวงแต่ละตัวทำให้เข้าใจแนวคิดเพื่อที่จะพิจารณาผลการทำหน้าที่ต่างกันของตัวลวง (DDF) ที่ถูกกำหนดโดยผลของผลการทำหน้าที่ต่างกันของตัวลวง (DDF) ลำดับที่ j โดยมีเงื่อนไขการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบที่มีลักษณะแบบเอกรูป (uniform DIF) อเนกรูป (nonuniform DIF) และการตรวจสอบการทำหน้าที่ตัดกันในข้อสอบ ผลการวิจัยชี้ให้เห็นศักยภาพของลักษณะระดับข้อสอบ (item-level properties) ที่นำไปสู่ทั้ง 3 ลักษณะ อีกทั้งข้อค้นพบอาจจะให้มูลเหตุของรูปแบบการทำหน้าที่ต่างกันของข้อสอบที่จะช่วยวิเคราะห์สาเหตุการทำหน้าที่ต่างกันของข้อสอบในลักษณะเฉพาะได้

Penfield (2010 b) ได้ศึกษารูปแบบและขนาดของผลการทำหน้าที่ต่างกันของข้อสอบในข้อสอบประเภทหลายตัวเลือก ที่ถูกกำหนดโดยผลของความไม่แปรเปลี่ยนระดับตัวลวง ในโมเดล NRM จากผลการวิจัยชี้ให้เห็นว่า 1) การมีผลการทำหน้าที่ต่างกันของตัวลวง (a differential distractor functioning : DDF) หนึ่งตัว หรือมากกว่าศูนย์ (nonzero) ทำให้ทราบผลการทำหน้าที่ต่างกันของตัวลวงประเภทไม่เป็นศูนย์ (nonzero DIF effect) 2) ขนาดของผลการทำหน้าที่ต่างกันของตัวลวง (DDF) สร้างจากขอบเขตบนไปจนถึงขนาดของผลจากการทำหน้าที่ต่างกันของข้อสอบ 3) ผลการทำหน้าที่ต่างกันของตัวลวงขนาดใหญ่ (a large DDF effect) ไม่ได้เป็นตัวที่บอกถึงผลการทำหน้าที่ต่างกันของข้อสอบขนาดใหญ่ (a large DIF effect) 4) ภายใต้เงื่อนไขของผลการทำหน้าที่ต่างกันของตัวลวงที่คงที่เป็นผลจากการทำหน้าที่ต่างกันของข้อสอบที่เป็นอิสระกับคุณลักษณะข้อสอบ แต่ภายใต้เงื่อนไขของผลการทำหน้าที่ต่างกันของตัวลวงที่มีความ แปรเปลี่ยนเป็นผลจากขนาดผลการทำหน้าที่ต่างกันของข้อสอบที่ขึ้นอยู่กับคุณลักษณะข้อสอบ และ 5) ถึงแม้ว่าการตรวจสอบการทำหน้าที่ต่างกันแบบตัดกันในข้อสอบสามารถเกิดได้เพียงกรณีเดียว คือ ผลการทำหน้าที่ต่างกันของตัวลวงที่มีลักษณะลู่ออก (divergent DDF effects) (เครื่องหมายต่างกัน) แต่ว่าผลการทำหน้าที่ต่างกันของตัวลวงที่มีลักษณะลู่ออกก็ไม่ได้เกิดการตรวจสอบการทำหน้าที่ต่างกันแบบตัดกันในข้อสอบ (CDIF) เสมอไป

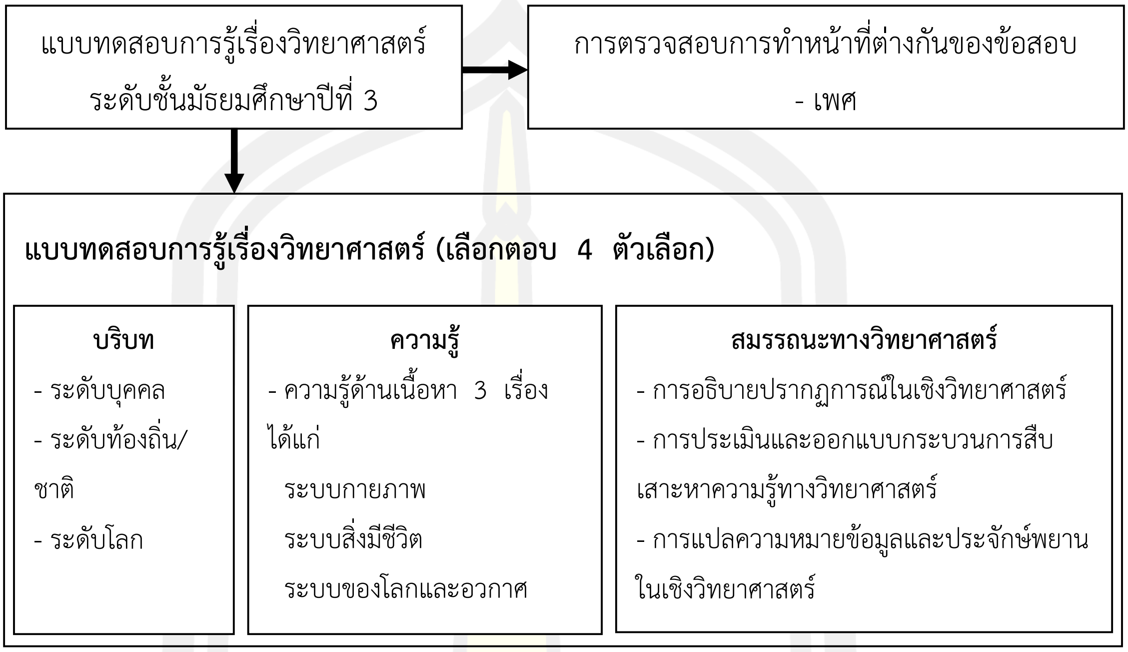
Taylor และ Lee (2012) ได้ศึกษาการทำหน้าที่ต่างกันของข้อสอบระหว่างเพศของนักเรียนเกรด 4, 7 และ 10 ในข้อสอบวิชาการอ่านและคณิตศาสตร์ ซึ่งการทำหน้าที่ต่างกันของข้อสอบระหว่างเพศ ใช้การทดสอบผ่าน POLYSIBTEST และ Rasch พบว่า ในข้อสอบการอ่านและคณิตศาสตร์ โดยทั่วไปในข้อสอบที่มีตัวเลือกหลายตัวเลือกเข้าข้างนักเรียนชาย ในขณะที่ข้อสอบอัตนัยที่ต้องตอบแบบมีโครงสร้างเข้าข้างนักเรียนหญิง การวิเคราะห์เนื้อหาแสดงให้เห็นว่าข้อสอบประเภทการอ่านวัดความเข้าใจในเนื้อหาหรือความหมายที่มีนัยยะ นักเรียนชายมักจะทำข้อสอบได้จากข้อสอบระบุการตีความหมายอย่างมีเหตุผลและการวิเคราะห์เนื้อหาของสิ่งที่เรียน ส่วนนักเรียนหญิงมักจะทำข้อสอบได้จากข้อสอบที่ต้องใช้การตีความเฉพาะบุคคล นอกจากนี้การวิเคราะห์เนื้อหาของข้อสอบวิชาคณิตศาสตร์แสดงให้เห็นว่า ข้อสอบเข้าข้างนักเรียนชายที่มีการทดสอบในเนื้อหาวิชาเรขาคณิต ความน่าจะเป็น และพีชคณิต ส่วนข้อสอบวิชาคณิตศาสตร์ที่เข้าข้างนักเรียนหญิงเป็นการตีความเชิงสถิติ การแก้ปัญหาหลากหลายขั้นตอน และการให้เหตุผลเชิงคณิตศาสตร์

การศึกษาการทำหน้าที่ต่างกันของข้อสอบ (DIF) เป็นประเด็นทางการศึกษาที่นักวัดผลทางการศึกษาควรให้ความสำคัญ ประวัติความเป็นมาในการศึกษาการทำหน้าที่ต่างกันของข้อสอบมีการศึกษามาเป็นระยะเวลานาน ซึ่งมีทั้งการศึกษาในเชิงทฤษฎีโดยอาศัยการจำลองข้อมูลและการศึกษาโดยอาศัยข้อมูลที่มีอยู่ในสภาพจริงตามบริบทต่าง ๆ ทั้งนี้เมื่อพิจารณาการทำหน้าที่ต่างกันทั้งในข้อสอบที่มีการตรวจให้คะแนน 2 ค่า หรือทวิวิภาค และข้อสอบที่มีการตรวจให้คะแนนหลายค่า

หรือพหุวิภาค โดยมีการศึกษาในหลายตัวแปร เช่น เชื้อชาติ ศาสนา วัฒนธรรม ภูมิลำเนา สังคม เพศ ภาษา อายุ ประสบการณ์ เป็นต้น สามารถแยกวิธีที่มีการศึกษาการทำหน้าที่ต่างกันของข้อสอบเป็นประเภท 2 ประเภทหลัก ๆ คือ ตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) และตามทฤษฎีการทดสอบแนวใหม่ (IRT) ทั้งที่อยู่ในรูปแบบพาราเมทริก และรูปแบบนันพาราเมทริก นอกจากนี้ประเภทการทำหน้าที่ต่างกันของข้อสอบยังสามารถแบ่งออกได้เป็นแบบเอกรูป (uniform DIF) และแบบอเนกรูป (nonuniform DIF) ซึ่งทั้งสองลักษณะของข้อสอบยังสามารถแยกประเภทการทำหน้าที่ต่างกันออกเป็น 2 ทิศทาง คือ แบบมีทิศทางเดียว (unidirectional DIF) เกิดขึ้นได้กับข้อสอบที่โค้งที่ต่างกันแต่ไม่ตัดกันในข้อสอบ (distinct but noncrossing DIF) และแบบไม่มีทิศทาง (nondirectional DIF) เกิดขึ้นได้กับคือ โค้งที่ตัดกันในข้อสอบ (crossing ICCs)

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง แม้ว่าจะมีการศึกษากับการทำหน้าที่ต่างกันของข้อสอบและแบบสอบมาเป็นระยะเวลานาน ส่วนใหญ่เป็นการศึกษาการทำหน้าที่ต่างกันของข้อสอบตามลักษณะการเกิดการทำหน้าที่ต่างกันของข้อสอบหรือการทำหน้าที่ต่างกันของแบบสอบโดยทั่วไป ภายหลังนี้ยังมีนักวิชาการบางท่านที่พัฒนาองค์ความรู้ทางด้านการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ โดยการศึกษาการทำหน้าที่ต่างกันของตัวลวงในข้อสอบเพื่อนำเสนอสารสนเทศที่เป็นประโยชน์ต่อวงการวิชาการทางการวัดผล ซึ่งนับว่าเป็นความพยายามที่จะพัฒนาองค์ความรู้ทางด้านการศึกษา การทำหน้าที่ต่างกันของข้อสอบให้มีความหลากหลายมากยิ่งขึ้น ดังนั้นจากเหตุผลและความสำคัญดังกล่าวทำให้ผู้วิจัยมีแนวคิดที่จะพัฒนาองค์ความรู้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบให้มีความกว้างขวางและเป็นประโยชน์ต่อวงการการศึกษา ผู้วิจัยจึงนำประเด็นข้อสอบที่มีการตรวจให้คะแนนแบบทวิวิภาคมาพัฒนาองค์ความรู้เพิ่มเติมให้เข้ากับบริบทของประเทศไทยมากยิ่งขึ้น

กรอบแนวคิดในการวิจัย



ภาพประกอบ 12 กรอบแนวคิดในการวิจัยการสร้างแบบทดสอบการรู้เรื่องวิทยาศาสตร์ที่มีการทำหน้าที่ต่างกันของข้อสอบ

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ เพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 3 ซึ่งผู้วิจัยจะกล่าวถึงวิธีดำเนินการวิจัยตามลำดับหัวข้อต่อไปนี้

1. ประชากรและกลุ่มตัวอย่าง
2. เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล
3. ขั้นตอนดำเนินการสร้างและหาคุณภาพของเครื่องมือ
4. การเก็บรวบรวมข้อมูล
5. การวิเคราะห์ข้อมูล
6. สถิติที่ใช้ในการวิเคราะห์ข้อมูล

ประชากรและกลุ่มตัวอย่าง

1. ประชากร

ประชากรที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 1 ปีการศึกษา 2561 ในจังหวัดชัยภูมิ จำนวน 6,890 คน จาก 66 โรงเรียน

2. กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 1 ปีการศึกษา 2561 ในจังหวัดชัยภูมิ จำนวน 689 คน จากโรงเรียน 13 โรงเรียน ที่ได้มาจากการสุ่มตัวอย่างแบบหลายขั้นตอน (Multi-Stage Random Sampling) มีขั้นตอนดังนี้

ขั้นที่ 1 กำหนดขนาดกลุ่มตัวอย่าง (Sample Size) โดยใช้เกณฑ์ ร้อยละ 10 ของจำนวนประชากรทั้งหมด (บุญชม ศรีสะอาด และคณะ, 2553) พบว่า ได้กลุ่มตัวอย่าง 689 คน

ขั้นที่ 2 ดำเนินการสุ่มตัวอย่างด้วยวิธีสุ่มแบบหลายขั้นตอน (Multi-Stage Random Sampling) ดังนี้

2.1 ใช้อำเภอหน่วยที่มีโรงเรียนในจังหวัดชัยภูมิ 50% จากทั้งหมด 16 อำเภอ สุ่มมา 8 อำเภอ โดยใช้วิธีการสุ่มแบบกลุ่ม (Cluster Random Sampling) ประกอบด้วยโรงเรียนระดับมัธยมศึกษา 45 โรง

2.2 สุ่มโรงเรียนตามขนาดของโรงเรียน โดยวิธีการสุ่มแบบแบ่งชั้น (Stratified Sampling) โดยใช้ขนาดโรงเรียนเป็นชั้น ใช้โรงเรียนเป็นหน่วยในการสุ่ม การกำหนดขนาดของโรงเรียนใช้เกณฑ์ของสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานกำหนดขนาดของโรงเรียนดังนี้

โรงเรียนขนาดใหญ่ 1,500 คนขึ้นไป

โรงเรียนขนาดกลาง 500–1,499 คน

โรงเรียนขนาดเล็กน้อยกว่า 500 คน

พบว่า เป็นโรงเรียนขนาดใหญ่ 8 โรง โรงเรียนขนาดกลาง 9 โรง และโรงเรียนขนาดเล็ก 28 โรง รวม 45 โรงเรียน

2.3 ใช้โรงเรียนในแต่ละขนาดของแต่ละอำเภอเป็นหน่วยการสุ่ม ทำการสุ่มอย่างง่าย (Simple Random Sampling) มาร้อยละ 30 พบว่า ได้โรงเรียนขนาดใหญ่ 3 โรง ขนาดกลาง 3 โรง และขนาดเล็ก 7 โรง ประกอบด้วย ห้องเรียน 38 ห้อง

2.4 ใช้ห้องเรียนแต่ละขนาดของแต่ละโรงเรียนที่สุ่มได้เป็นหน่วยในการสุ่มทำการสุ่มอย่างง่าย (Simple Random Sampling) มาร้อยละ 50 พบว่า โรงเรียนขนาดใหญ่ มี 6 ห้อง โรงเรียนขนาดกลาง มี 6 ห้อง และโรงเรียนขนาดเล็ก มี 7 ห้อง ได้กลุ่มตัวอย่างทั้งหมด 689 คน ปรากฏดังตาราง 2

ตาราง 2 จำนวนกลุ่มตัวอย่างที่ใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำแนกตามขนาดของโรงเรียน

ขนาดโรงเรียน	ลำดับที่	ชื่อโรงเรียน	จำนวน (ห้องเรียน)	จำนวน (นักเรียน)
ขนาดใหญ่	1	โรงเรียนชัยภูมิภักดีชุมพล	2	90
	2	โรงเรียนเกษตรสมบูรณ์วิทยาคม	2	86
	3	โรงเรียนแก้งคร้อวิทยา	2	82
ขนาดกลาง	1	โรงเรียนบ้านเป้าวิทยา	2	67
	2	โรงเรียนโนนกอกวิทยา	2	62
	3	โรงเรียนนาหนองทุ่มวิทยา	2	70
ขนาดเล็ก	1	โรงเรียนกาญจนาภิเษกวิทยาลัย ชัยภูมิ	1	26
	2	โรงเรียนพระธาตุหนองสามหมื่น	1	33

ตาราง 2 (ต่อ)

ขนาดโรงเรียน	ลำดับที่	ชื่อโรงเรียน	จำนวน (ห้องเรียน)	จำนวน (นักเรียน)
ขนาดเล็ก	3	โรงเรียนหนองสังข์วิทยายน	1	34
	4	โรงเรียนโนนครัอวิทยา	1	32
	5	โรงเรียนสระพังวิทยาคม	1	31
	6	โรงเรียนกุดตุ้มวิทยา	1	36
	7	โรงเรียนกวางโจนศึกษา	1	40
รวม			19	689

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลในงานวิจัยครั้งนี้ ได้แก่ แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 เป็นแบบทดสอบเชิงสถานการณ์ 4 ตัวเลือก จำนวน 1 ฉบับ 54 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 18 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ ซึ่งเป็นเนื้อหาที่ครอบคลุมตามกรอบเนื้อหาของ PISA 2015

ขั้นตอนดำเนินการสร้างและหาคุณภาพของเครื่องมือ

ในการสร้างแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ผู้วิจัยได้ดำเนินการสร้างตามลำดับขั้นตอน ปรากฏดังภาพประกอบ 13

1. ศึกษาเอกสาร ทฤษฎีที่เกี่ยวข้อง เนื้อหาวิชาที่ต้องการออกข้อสอบ

↓

2. กำหนดเนื้อหา พฤติกรรมที่ต้องการวัด

↓

3. สร้างแบบทดสอบตามองค์ประกอบของการรู้เรื่องวิทยาศาสตร์

4. นำแบบทดสอบการรู้เรื่องวิทยาศาสตร์ไปทดลองใช้ (Try out) จำนวน 100 คน

↓

5. วิเคราะห์คุณภาพข้อสอบรายข้อและทั้งฉบับ ตามแนวคิด CTT

↓

6. นำแบบทดสอบการรู้เรื่องวิทยาศาสตร์ไปทดลองใช้กับกลุ่มตัวอย่าง จำนวน 689 คน

↓

7. วิเคราะห์คุณภาพข้อสอบรายข้อ หาค่าพารามิเตอร์ a และ b ตามแนวคิด IRT

↓

8. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ
(Differential Item Functioning Analysis) ข้อมูลจากกลุ่มตัวอย่าง

↓

9. แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

ภาพประกอบ 13 ขั้นตอนการสร้างและพัฒนาเครื่องมือแบบทดสอบการรู้เรื่องวิทยาศาสตร์

ขั้นที่ 1 ศึกษาแนวคิดทฤษฎีเกี่ยวกับการรู้เรื่องวิทยาศาสตร์ตามแนว PISA 2015 จากเอกสาร และงานวิจัยที่เกี่ยวข้อง

ศึกษาเอกสารที่เกี่ยวข้อง วิเคราะห์ สังเคราะห์ แนวคิด ทฤษฎี เอกสาร ตาราง บทความ งานวิจัยทั้งในประเทศและต่างประเทศ รวมทั้งฐานข้อมูลต่าง ๆ ในอินเตอร์เน็ตที่เกี่ยวข้องกับการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 และสังเคราะห์เอกสารที่เกี่ยวข้องโดยสรุปได้ดังนี้

การรู้เรื่องวิทยาศาสตร์ (Scientific Literacy) หมายถึง สมรรถนะของการเข้าใจบทบาทของวิทยาศาสตร์ที่มีต่อโลก ตัดสินใจในประเด็นต่าง ๆ บนพื้นฐานของวิทยาศาสตร์ มีความผูกพันกับวิทยาศาสตร์ตามความต้องการหรือจำเป็นของตน ในอันที่จะทำให้บุคคลนั้นเป็นผู้มี

ส่วนร่วมในสังคม โดยการนำความรู้วิทยาศาสตร์ แนวคิดวิทยาศาสตร์ ทักษะกระบวนการทางวิทยาศาสตร์ไม่ว่าจะเป็นการให้เหตุผล การคิดวิเคราะห์ การสื่อสาร สื่อความหมาย การแก้ปัญหาเชิงวิทยาศาสตร์ การติดตามและประเมินผลข้อโต้แย้ง การนำเสนอข้อมูลมาใช้ในสถานการณ์ต่าง ๆ ในชีวิตจริง และเตรียมพร้อมสำหรับการเป็นพลเมืองที่มีวิจารณญาณ มีความมั่นใจในตนเอง ห่วงใยและสร้างสรรค์สังคมในอนาคต มิติเนื้อหาวิทยาศาสตร์ (Scientific Literacy) ที่ผู้เรียนต้องนำมาใช้ในการแก้ปัญหาซึ่งประกอบไปด้วย 3 องค์ประกอบ คือ

1. บริบทของวิทยาศาสตร์
2. สมรรถนะทางวิทยาศาสตร์
3. ความรู้ทางวิทยาศาสตร์ : เนื้อหาทางวิทยาศาสตร์

จากการศึกษาเอกสารเกี่ยวกับการรู้เรื่องวิทยาศาสตร์ประเด็นและสภาพปัญหาดังกล่าวจึงทำให้ผู้วิจัยสนใจที่จะทำวิจัยกับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 โดยทำการทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ที่ต้องการทดสอบ คือ บริบทของวิทยาศาสตร์ สมรรถนะทางวิทยาศาสตร์ ความรู้ทางวิทยาศาสตร์ (ความรู้ด้านเนื้อหา) ซึ่งประกอบไปด้วยรายละเอียด ดังนี้

1. บริบท ซึ่งแบ่งเป็น 3 ด้าน ได้แก่
   1.1 บริบทระดับบุคคล
   1.2 บริบทระดับท้องถิ่น/ชาติ
   1.3 บริบทระดับโลก
2. สมรรถนะทางวิทยาศาสตร์ มี 3 ด้าน ได้แก่
   2.1 อธิบายปรากฏการณ์อย่างมีหลักการทางวิทยาศาสตร์
   2.2 การประเมินและการออกแบบการสืบเสาะทางวิทยาศาสตร์
   2.3 การแปลความหมายข้อมูลและประจักษ์พยานอย่างมีหลักการทางวิทยาศาสตร์
3. ความรู้ทางวิทยาศาสตร์ : เนื้อหาทางวิทยาศาสตร์ ซึ่งประกอบไปด้วย 2 องค์ประกอบย่อย ได้แก่
   3.1 ความรู้ทางวิทยาศาสตร์ ประกอบด้วย 1 ด้าน ได้แก่ ความรู้ด้านเนื้อหา
   3.2 เนื้อหาทางวิทยาศาสตร์ ประกอบด้วย 3 ด้าน ได้แก่
      3.2.1 ระบบทางกายภาพ
      3.2.2 ระบบสิ่งมีชีวิต
      3.2.3 ระบบโลกและอวกาศ

ขั้นที่ 2 สังเคราะห์องค์ประกอบและนิยามของการรู้เรื่องวิทยาศาสตร์ที่ต้องการทดสอบว่าจะแสดงออกให้เห็นในลักษณะใดบ้าง เพื่อกำหนดเป็นองค์ประกอบ พฤติกรรมเป้าหมาย เลือกพฤติกรรมบ่งชี้ที่มีความสำคัญหรือจะเป็นตัวแทนที่ดีของการรู้เรื่องวิทยาศาสตร์ และกำหนดข้อคำถามและสถานการณ์ ดังต่อไปนี้

1. รูปแบบข้อสอบการรู้เรื่องด้านวิทยาศาสตร์

1.1 กรอบการประเมินวิทยาศาสตร์ครอบคลุมสมรรถนะทางวิทยาศาสตร์ 3สมรรถนะ ได้แก่ การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์ การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์ และการแปลความหมายข้อมูลและการใช้ประจักษ์พยานเชิงวิทยาศาสตร์ ครอบคลุมความรู้ทางวิทยาศาสตร์ 1 ด้าน ได้แก่ ความรู้ด้านเนื้อหา และคลอบคลุมเนื้อหาทางวิทยาศาสตร์ 3 ด้าน ได้แก่ ระบบทางกายภาพ ระบบการดำรงชีวิต ระบบโลกและอวกาศ ทั้งนี้ข้อสอบมีการทดสอบกับนักเรียนแตกต่างกันหลายกลุ่ม จึงมีการสร้างข้อสอบให้มีความหลายหลายและเลือกให้แต่ละแบบมีจำนวนข้อที่ใกล้เคียงกัน ได้แก่

1.2 รูปแบบการตอบคำถามที่ใช้วัดสมรรถนะและความรู้ทางวิทยาศาสตร์ ได้แก่ การเลือกตอบทั่วไป มีลักษณะการตอบคำถาม ดังนี้

1.2.1 การเลือกหนึ่งคำตอบจากสี่ตัวเลือก

1.2.2 การเลือกคำตอบที่เป็นองค์ประกอบที่อยู่ในภาพหรือข้อความ

1.3 ร่างโครงข้อสอบการรู้เรื่องวิทยาศาสตร์ ตามแนว PISA 2015 ผู้วิจัยเลือกรูปแบบการตอบ 1 รูปแบบ คือ แบบเลือกตอบชนิด 4 ตัวเลือก จำแนกตามกรอบโครงสร้างการประเมินการรู้เรื่องวิทยาศาสตร์ ได้แก่ สมรรถนะทางวิทยาศาสตร์ ความรู้ทางวิทยาศาสตร์ (ความรู้ด้านเนื้อหา) บริบทของวิทยาศาสตร์ และระดับบริบทของวิทยาศาสตร์ ปรากฏดังตาราง 3

ตาราง 3 จำนวนข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ที่ใช้จำแนกตามกรอบโครงสร้างการประเมินผลนักเรียนโครงการ PISA 2015

การรู้เรื่องวิทยาศาสตร์	ความรู้ทางวิทยาศาสตร์ (ด้านเนื้อหา)			จำนวนข้อสอบ (ข้อ)
	ระบบทางกายภาพ (ข้อ)	ระบบสิ่งมีชีวิต (ข้อ)	ระบบโลกและอวกาศ (ข้อ)	
สมรรถนะทางวิทยาศาสตร์				
การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์	6	6	6	18
การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์	6	6	6	18
การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์	6	6	6	18
รวม	**18**	**18**	**18**	**54**
บริบทของวิทยาศาสตร์				
สุขภาพและโรคภัย	4	2	1	7
ทรัพยากรธรรมชาติ	5	2	9	16
คุณภาพสิ่งแวดล้อม	3	4	1	8
อันตราย	4	2	2	8
ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี	2	8	5	15
รวม	**18**	**18**	**18**	**54**
ระดับบริบทของวิทยาศาสตร์				
ระดับส่วนตัว	6	7	3	16
ระดับท้องถิ่น/ชาติ	9	6	6	21
ระดับโลก	3	5	9	17
รวม	**18**	**18**	**18**	**54**

ขั้นที่ 3 สร้างข้อคำถามและตัวเลือกแบบทดสอบการรู้เรื่องวิทยาศาสตร์ โดยยกเหตุการณ์หรือเรื่องราวที่คล้ายคลึงสถานการณ์จริง และมีความเกี่ยวข้องกับพฤติกรรมการแสดงออกหรือต้องการวัดให้ผู้ตอบแบบวัดพิจารณามาสร้างสถานการณ์และข้อคำถามแบบหลายตัวเลือกที่มีการให้คะแนนแบบ 0-1 ดังนั้น แบบทดสอบการรู้เรื่องวิทยาศาสตร์ที่พัฒนาขึ้นจึงเป็นแบบวัดเชิงสถานการณ์แบบหลายตัวเลือกที่มีการให้คะแนนแบบ 0-1 ซึ่งมีส่วนที่สำคัญ 2 ส่วน คือ สถานการณ์และคำถาม

ขั้นที่ 4 ผู้เชี่ยวชาญตรวจสอบความตรงเชิงเนื้อหาของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 โดยให้ผู้เชี่ยวชาญทางด้านวิทยาศาสตร์ จำนวน 5 ท่าน พิจารณาว่าแบบทดสอบนั้นมีข้อสอบแต่ละข้อตรงตามนิยามที่จะวัดหรือไม่ แล้วนำผลการพิจารณาจากผู้เชี่ยวชาญมาหาค่าเฉลี่ยรายข้อ ถ้าได้ค่าเฉลี่ยตั้งแต่ 0.50 ถึง 1.00 ข้อสอบนั้นใช้ได้

โดยใช้การประเมินความคิดเห็นของผู้เชี่ยวชาญ จำนวน 5 ท่าน ดังนี้

1. ผู้ช่วยศาสตราจารย์ ดร.สมทรง สิทธิ ผู้เชี่ยวชาญด้านหลักสูตรและการสอน อาจารย์ประจำภาควิชาหลักสูตรและการสอน คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

2. อาจารย์รัตนา เพ็งเพราะ ผู้เชี่ยวชาญด้านวิทยาศาสตร์ อาจารย์ประจำภาคชีววิทยา คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ คุณวุฒิวิทยาศาสตรมหาบัณฑิต สาขาชีววิทยา มหาวิทยาลัยมหาสารคาม

3. อาจารย์วีรพล เลพล ผู้เชี่ยวชาญด้านวิทยาศาสตร์ อาจารย์ประจำภาควิศวกรรม เครื่องกลคณะเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏพิบูลสงคราม คุณวุฒิวิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมเครื่องกล มหาวิทยาลัยมหาสารคาม

4. ดร.ตวัญจลักษณ์ พวงนิล ผู้เชี่ยวชาญด้านวิทยาศาสตร์ ตำแหน่ง ครูวิทยฐานะครูชำนาญการพิเศษ โรงเรียนสารคามพิทยาคม อำเภอเมืองมหาสารคาม จังหวัดมหาสารคาม คุณวุฒิปรัชญาดุษฎีบัณฑิต สาขาบริหารการศึกษา มหาวิทยาลัยมหาสารคาม

5. อาจารย์ตันติกร ขุนาพรม ผู้เชี่ยวชาญด้านวิจัย ตำแหน่ง ศึกษานิเทศก์ สำนักงานเขตพื้นที่การศึกษาประถมศึกษามหาสารคาม เขต 1 อำเภอเมืองมหาสารคาม จังหวัดมหาสารคาม คุณวุฒิการศึกษามหาบัณฑิต สาขาวิจัยและประเมินผลการศึกษา มหาวิทยาลัยมหาสารคาม

เกณฑ์ในการให้คะแนน ดังนี้

ให้ +1 คะแนน เมื่อแน่ใจว่าข้อคำถามนั้นวัดได้ตรงตามนิยาม

ให้ 0 คะแนน เมื่อไม่แน่ใจว่าข้อคำถามนั้นวัดได้ตรงตามนิยาม

ให้ -1 คะแนน เมื่อแน่ใจว่าข้อคำถามนั้นวัดได้ไม่ตรงตามนิยาม

ขั้นที่ 5 นำแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ไปทดสอบครั้งที่ 1 กับกลุ่มทดลองใช้ (Try Out) ที่ไม่ใช่กลุ่มตัวอย่าง จำนวน 100 คน เพื่อหาคุณภาพของข้อสอบตามแนวทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory : CTT) เพื่อวิเคราะห์ค่าความยาก (p) ค่าอำนาจจำแนก (r) และค่าความเชื่อมั่น แล้วคัดเลือกข้อสอบที่มีคุณภาพตามที่กำหนด แล้วคัดเลือกข้อสอบที่มีคุณภาพตามที่กำหนดไว้ จำนวน 45 ข้อ

ขั้นที่ 6 นำแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 45 ข้อ ไปทดสอบครั้งที่ 2 กับกลุ่มตัวอย่าง จำนวน 689 คน เพื่อหาคุณภาพของข้อสอบตามแนวทฤษฎีการตอบสนองของข้อสอบ (Item Response Theory : IRT) โดยใช้โมเดลโลจีสติกแบบ 2 พารามิเตอร์ (Two–Parameter Model) เพื่อวิเคราะห์ค่าความยากของข้อสอบ (b) ค่าอำนาจจำแนกของข้อสอบ (a) ทำการวิเคราะห์ข้อสอบโดยผู้วิจัยได้เขียนคำสั่งจากโปรแกรม Mplus แล้วคัดเลือกข้อสอบที่มีคุณภาพตามที่กำหนดไว้ จำนวน 45 ข้อ

ขั้นที่ 7 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ผู้วิจัยได้นำข้อสอบที่ได้จากการหาคุณภาพตามแนวทฤษฎีการตอบสนองของข้อสอบ (Item Response Theory : IRT) จำนวน 45 ข้อ โดยใช้วิธีมิมิค (Multiple Indicators and Multiple Causes : MIMIC) เมื่อจำแนกตามเพศ โดยประยุกต์ใช้โปรแกรม Mplus

การเก็บรวบรวมข้อมูล

การวิจัยครั้งนี้ มีเครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล จำนวน 2 ฉบับ ได้แก่ แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 และแบบประเมินความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับความตรงเชิงเนื้อหาของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 และเก็บรวบรวมข้อมูลจากนักเรียนที่เป็นกลุ่มตัวอย่างผู้วิจัยดำเนินการเก็บข้อมูล ดังนี้

1. ทำหนังสือขอความอนุเคราะห์ในการรวบรวมข้อมูลในการวิจัยถึงผู้บริหารสถานศึกษาที่ได้รับการสุ่มเป็นกลุ่มตัวอย่างในการวิจัย
2. ประสานงานชี้แจงทำความเข้าใจกับผู้บริหารและครูรวบรวมข้อมูลเกี่ยวกับจุดมุ่งหมายการวิจัย เครื่องมือในการรวบรวมข้อมูล วิธีการเก็บรวบรวมข้อมูล ระยะเวลาการเก็บรวบรวมข้อมูล และกำหนดนัดหมายสำหรับการรับคืนแบบทดสอบ
3. สถานศึกษาทำหน้าที่สุ่มห้องเรียนและทำแบบทดสอบของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ตามระยะเวลาที่กำหนด

4. ประสานงานทางโทรศัพท์เพื่อยืนยันความพร้อมของสถานศึกษาที่จะให้เก็บรวบรวมข้อมูล และการยืนยันระยะเวลาการรวบรวมข้อมูลในแต่ละครั้ง รวมถึงการปรับระยะเวลาการเก็บรวบรวมข้อมูล ในกรณีการรวบรวมข้อมูลในสถานศึกษาบางแห่งไม่เป็นไปตามที่กำหนด

5. ข้อมูลที่ได้รับกลับ ผู้วิจัยทำการลงรหัสรายชื่อนักเรียน สถานศึกษาเป็นรหัสที่สามารถแยกแยะได้ว่าผู้ตอบเป็นนักเรียนในโรงเรียนใด ทั้งนี้เพื่อเป็นการเชื่อมโยงข้อมูลการวัดแต่ละครั้งเข้าด้วยกันได้ง่าย

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลผู้วิจัยแบ่งการวิเคราะห์ออกเป็น 5 ขั้นตอน รายละเอียดดังนี้

ขั้นตอนที่ 1 ตรวจสอบความเที่ยงตรงของข้อสอบที่พัฒนาขึ้นตามกรอบโครงสร้างตามแนว PISA 2015 โดยพิจารณาค่าเฉลี่ย ($\bar{X}$) ของความคิดเห็นของผู้เชี่ยวชาญ จำนวน 5 ท่าน

ขั้นตอนที่ 2 วิเคราะห์หาคุณภาพของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ตามแนวทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory : CTT) เพื่อวิเคราะห์ค่าความยาก (p) ค่าอำนาจจำแนก (r) และค่าความเชื่อมั่นกับกลุ่มทดลองใช้ (Try Out) ที่ไม่ใช่กลุ่มตัวอย่าง จำนวน 100 คน

ขั้นตอนที่ 3 ตรวจสอบการแจกแจงแบบปกติคะแนนแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำแนกตามตัวแปรเพศ และวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson’s Product Moment Correlation Coefficient) กับกลุ่มตัวอย่าง จำนวน 689 คน

ขั้นตอนที่ 4 วิเคราะห์ค่าพารามิเตอร์ของข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ตามแนวทฤษฎีการตอบสนองของข้อสอบ (Item Response Theory : IRT) เพื่อวิเคราะห์ค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ (a) ของข้อสอบกับกลุ่มตัวอย่าง จำนวน 689 คน โดยใช้โมเดลโลจีสติกแบบ 2 พารามิเตอร์ (Two–Parameter Model) โดยผู้วิจัยได้เขียนคำสั่งจากโปรแกรม Mplus

การประมาณค่าพารามิเตอร์ของข้อสอบ ซึ่งมีเกณฑ์การคัดเลือกข้อสอบ ดังนี้ (ศิริชัย กาญจนวาสี, 2555)

ค่าความยากของข้อสอบ (b-parameter) มีค่าอยู่ระหว่าง -2.50 ถึง 2.50 ค่าที่เป็นลบแสดงว่าข้อสอบง่าย และค่าที่เป็นบวกแสดงว่าข้อสอบยาก

ค่าอำนาจจำแนกข้อสอบ (a-parameter) มีค่าอยู่ระหว่าง 0.50 ถึง 2.50

ขั้นตอนที่ 5 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค (Multiple Indicators and Multiple Causes : MIMIC) เมื่อจำแนกตามเพศ โดยประยุกต์ใช้โปรแกรม Mplus มีขั้นตอนดังต่อไปนี้

1. วิเคราะห์โมเดล CFA โดยไม่มีตัวแปรทำนาย ในการวิเคราะห์ข้อมูลตามทฤษฎีการตอบสนองของข้อสอบ (IRT) เป็นการศึกษาคุณสมบัติอยู่บนพื้นฐานข้อตกลงเอกมิติ (Unidimensional) ของตัวแปรแฝง

2. การเพิ่มตัวแปรทำนาย แต่ไม่มีอิทธิพลตรงต่อข้อสอบ ซึ่งในการวิจัยครั้งนี้ผู้วิจัยได้มีตัวแปร คือ เพศ

3. เพิ่มอิทธิพลทางตรงต่อข้อสอบบังคับค่าให้เป็น 0 เพื่อกำหนดให้ตัวแปรต้นไม่มีผลทางตรงต่อข้อสอบ

4. ประเมินความกลมกลืนของโมเดล และตรวจสอบผลทางตรงที่มีนัยสำคัญทางสถิติ ถ้าตรวจพบว่ามีผลทางตรง และมีนัยสำคัญทางสถิติ $p < 0.05$ แสดงว่าข้อสอบข้อนั้น คือ ข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบ (DIF) ปรากฏดังตาราง 4

ตาราง 4 วิธีการพิจารณา Mplus Output และการตรวจสอบดัชนีความสอดคล้องของโมเดล

การพิจารณา Mplus Output	ดัชนีและคำอธิบายความสอดคล้องของโมเดล
Chi-Square Test of Model Fit	ค่า Chi-Square เพื่อตรวจสอบความสอดคล้องของโมเดล
Value	Chi-Square (χ^2) = ควรต้องพิจารณาดัชนีอื่นประกอบ เนื่องจากความลำเอียงที่เกิดจากขนาดกลุ่มตัวอย่าง
Value/Degrees of Freedom	χ^2/df = ควรมีค่าไม่เกิน 2
P-Value	P-Value = มีค่ามากกว่า 0.01 หรือมากกว่า 0.05
Estimate	RMSEA = ควรมีค่าน้อยกว่า 0.07 0.000-0.050 = โมเดลสอดคล้องกับข้อมูลเชิงประจักษ์ดีมาก 0.051-0.080 = โมเดลค่อนข้างสอดคล้องกับข้อมูลเชิงประจักษ์ดี 0.081-0.100 = โมเดลสอดคล้องกับข้อมูลเชิงประจักษ์เล็กน้อย มากกว่า 0.100 = โมเดลยังไม่สอดคล้องกับข้อมูลเชิงประจักษ์

ตาราง 4 (ต่อ)

การพิจารณา Mplus Output	ดัชนีและคำอธิบายความสอดคล้องของโมเดล
SRMR	ค่ารากที่สองของค่าเฉลี่ยกำลังสองของส่วนเหลือในรูปคะแนน
(Standardized Root Mean Square Residual)	มาตรฐาน (SRMR)
Value	SRMR = ควรมีค่าน้อยกว่า 0.08
CFI/TLI	ดัชนีวัดความกลมกลืนเปรียบเทียบ/ดัชนี Tucker-Lewis
CFI (Comparative Fit Index) TLI (Tucker-Lewis Index)	CFI = ควรมีค่ามากกว่า 0.95 TLI = ควรมีค่ามากกว่า 0.95
STDYX Standardization	เป็นส่วนแสดงค่าตัวเลขที่นำไปเขียนกำกับเส้นในภาพโมเดล
Estimate	ค่าสัมประสิทธิ์น้ำหนักองค์ประกอบมาตรฐาน
R-SQUARE	ค่าสัมประสิทธิ์ความเที่ยงหรือสัมประสิทธิ์การทำนาย
Estimate	R^2 ค่าต้องมีนัยสำคัญด้วย
MODEL RESULTS	ผลของโมเดล
Residual Variances	เป็นตัวเลขที่นำไปใช้ปรับโมเดล กรณีต้องการใช้คำสั่ง @ หรือ *
MODEL MODIFICATION INDICES	ดัชนีปรับโมเดล
WITH Statements	กรณีต้องการใช้คำสั่ง With ปรับโมเดล

ที่มา : โกศล จิตวิรัตน์ (2556)

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. สถิติที่ใช้ในการวิเคราะห์หาคุณภาพของเครื่องมือ

1.1 ความเที่ยงตรงเชิงเนื้อหา (Content Validity) ใช้สูตรดังนี้ (สมบัติ ท้ายเรือคำ, 2553)

$$\text{ความคิดเห็นของความเห็น} = \frac{\sum R}{N}$$

เมื่อ $\sum R$ แทน ผลรวมความคิดเห็นของผู้เชี่ยวชาญ

N แทน จำนวนผู้เชี่ยวชาญ

1.2 ค่าความยากของข้อสอบ (Difficulty) ใช้สูตรดังนี้ (สมบัติ ท้ายเรือคำ, 2553)

$$P = \frac{R}{N}$$

เมื่อ P แทน ค่าความยาก

R แทน จำนวนผู้ตอบถูกทั้ง

N แทน จำนวนผู้เข้าสอบทั้งหมด

1.3 ค่าอำนาจจำแนกของข้อสอบ (Discriminating Power) ใช้สูตรดังนี้ (สมบัติ ท้ายเรือคำ, 2553)

$$r = \frac{P_H - P_L}{n}$$

เมื่อ r แทน ดัชนีอำนาจจำแนก

P_H แทน จำนวนผู้ตอบถูกในกลุ่มสูง

P_L แทน จำนวนผู้ตอบถูกในกลุ่มต่ำ

n แทน จำนวนผู้ตอบทั้งหมดของกลุ่มสูงหรือกลุ่มต่ำ

1.4 การวิเคราะห์ข้อสอบที่ใช้ค่าพารามิเตอร์แบบ 2 พารามิเตอร์ คือ ค่าความยากของข้อสอบ (b) และค่าอำนาจจำแนกของข้อสอบ (a) เขียนเป็นฟังก์ชันโลจิสติก (Hambleton, Swaminathan and Rogers, 1991)

$$P_i(\theta) = \frac{1}{1+e^{-Da_i(\theta-b_i)}} \quad i = 1, 2, 3,\ldots n$$

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นของผู้เข้าสอบที่มีความสามารถ (θ) จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

b_i คือ ค่าความยากของข้อสอบข้อที่ i

a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i

θ คือ ความสามารถของผู้เข้าสอบ

D คือ 1.70

1.5 ค่าความเชื่อมั่น โดยใช้สูตรของคูเดอร์-ริชาร์ดสัน สูตร KR-20 (สมนึก ภัททิยธนี, 2553)

$$r_{tt} = \frac{n}{n-1}\left\{1 \frac{\sum pq}{s_x^2}\right\}$$

เมื่อ r แทน จำนวนข้อสอบ

n แทน จำนวนข้อสอบของแบบทดสอบทั้งฉบับ

p แทน อัตราส่วนของผู้ตอบถูกในข้อนั้น

q แทน อัตราส่วนของผู้ตอบผิดในข้อนั้น

S^2 แทน ความแปรปรวนของคะแนนทั้งฉบับ

2. สถิติที่ใช้ในการวิเคราะห์ข้อมูลพื้นฐาน

2.1 ค่าเฉลี่ย (Arithmetic Mean) กรณีคำนวณค่าเฉลี่ยเลขคณิตสำหรับข้อมูลที่แจกแจงความถี่ (สมบัติ ท้ายเรือคำ, 2551)

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

เมื่อ $\bar{X}$ แทน ค่าเฉลี่ย

$\sum X$ แทน ผลรวมของคะแนนทั้งหมด

N แทน จำนวนคนในกลุ่ม

2.2 ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) (สมนึก ภัททิยธนี, 2555) ใช้สูตร ดังนี้

$$S.D. = \sqrt{\frac{N\sum X^2-(\sum X)^2}{N(N-1)}}$$

เมื่อ $S.D.$ แทน ส่วนเบี่ยงเบนมาตรฐาน

X แทน ข้อมูลแต่ละตัว

X^2 แทน ข้อมูลแต่ละตัวยกกำลังสอง

N แทน จำนวนข้อมูลทั้งหมด

3. สถิติที่ใช้ในการทดสอบสมมติฐาน

3.1 การทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC แบบเอกรูป (Uniform)

สูตรสำหรับวิธี MIMIC ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) แบบเอกรูป (Woods and Grimm, 2011) คือ

$$y_i^* = a_i\theta + \beta_i z + \varepsilon_i$$

เมื่อ y_i^* แทน การตอบสนองการทำหน้าที่ต่างกันของข้อสอบที่ *i*

a แทน อำนาจจำแนกของพารามิเตอร์

θ แทน ตัวแปรแฝง

β แทน สัมประสิทธิ์ของการถดถอยของการทดสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC แบบเอกรูป (Uniform)

z แทน ตัวแปรแฝง

ε แทน ความผิดพลาดในการวัดของข้อสอบที่ *i*

3.2 การทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC แบบอเนกรูป (Nonuniform)

สูตรสำหรับวิธี MIMIC ของการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) แบบอเนกรูป (Woods and Grimm, 2011) คือ

$$y_i^* = a_i\theta + a_i z + \omega_i\theta z + \varepsilon_i$$

เมื่อ y_i^* แทน การตอบสนองการทำหน้าที่ต่างกันของข้อสอบที่ *i*

a แทน อำนาจจำแนกของพารามิเตอร์

θ แทน ตัวแปรแฝง

ω แทน สัมประสิทธิ์ของการถดถอยของการทดสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) ด้วยวิธี MIMIC แบบอเนกรูป (Uniform)

z แทน ตัวแปรแฝง

ε แทน ความผิดพลาดในการวัดของข้อสอบที่ *i*

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยเรื่อง การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ผู้วิจัยได้นำเสนอผลการวิเคราะห์ข้อมูลตามลำดับดังต่อไปนี้

1. สัญลักษณ์ที่ใช้ในการเสนอผลการวิเคราะห์ข้อมูล
2. ลำดับขั้นตอนในการเสนอผลการวิเคราะห์ข้อมูล
3. ผลการวิเคราะห์ข้อมูล

สัญลักษณ์ที่ใช้ในการเสนอผลการวิเคราะห์ข้อมูล

การนำเสนอผลการวิเคราะห์ข้อมูลในการวิจัยครั้งนี้ ผู้วิจัยได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล ดังนี้

MIMIC	หมายถึง	วิธีการวิเคราะห์ข้อมูลคุณลักษณะแฝงที่มีหลายสาเหตุและวัดได้จากตัวบ่งชี้หลายตัว (Multiple Indicators and Multiple Causes Model)
p	หมายถึง	ค่าความยาก
r	หมายถึง	ค่าอำนาจจำแนก
r_{tt}	หมายถึง	ค่าความเชื่อมั่น
a	หมายถึง	ค่าอำนาจจำแนกของข้อสอบ
b	หมายถึง	ค่าความยากของข้อสอบ
Sci	หมายถึง	แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3
CFA	หมายถึง	การวิเคราะห์องค์ประกอบยืนยัน (Confirmatory Factor)
X^2	หมายถึง	ค่าสถิติไค-สแควร์ (Chi-Square)
df	หมายถึง	องศาอิสระ (Degree of Freedom)
CFI	หมายถึง	ดัชนีวัดความกลมกลืนเปรียบเทียบ (Comparative Fit Index)
TLI	หมายถึง	ดัชนี Tucker-Lewis
RMSEA	หมายถึง	ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของการประมาณค่า (Root Mean Square Error of Approximation)

SRMR หมายถึง ดัชนี Standardized Root Mean Square Residual

Estimate หมายถึง ค่าสัมประสิทธิ์น้ำหนักองค์ประกอบมาตรฐาน

ลำดับขั้นตอนในการเสนอผลการวิเคราะห์ข้อมูล

ในการวิจัยครั้งนี้ผู้วิจัยมีจุดประสงค์สำคัญ 1) เพื่อสร้างและหาคุณภาพแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 2) เพื่อตรวจสอบการทำหน้าที่ต่างกัน (DIF) ของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ซึ่งมีขั้นตอนในการดำเนินการดังนี้

ตอนที่ 1 ผลการสร้างและหาคุณภาพแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

ตอนที่ 2 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 3

ผลการวิเคราะห์ข้อมูล

ตอนที่ 1 ผลการสร้างและหาคุณภาพแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

1. ผลการวิเคราะห์คุณภาพข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ตามแนวทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory : CTT)

1.1 ความเที่ยงตรงเชิง (Validity)

ผู้วิจัยได้สร้างแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 54 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 18 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ แล้วนำไปให้ผู้เชี่ยวชาญจำนวน 5 คน พิจารณาความสอดคล้อง ปรากฏดังตาราง 5

พหุนํ ปณฺฑิโต ชีเว

ตาราง 5 ดัชนีความสอดคล้องของข้อสอบกับเนื้อหา แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

ความรู้ด้านเนื้อหา	ข้อที่	การพิจารณาของผู้เชี่ยวชาญ					รวม	IOC	แปลผล
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5			
ระบบทางกายภาพ	1	+1	-1	+1	+1	+1	3	0.6	คัดเลือก
	2	+1	+1	0	+1	+1	4	0.8	คัดเลือก
	3	+1	0	+1	+1	+1	4	0.8	คัดเลือก
	4	+1	-1	+1	+1	+1	3	0.6	คัดเลือก
	5	+1	+1	+1	+1	+1	5	1	คัดเลือก
	6	+1	+1	+1	+1	+1	5	1	คัดเลือก
	7	+1	+1	+1	+1	+1	5	1	คัดเลือก
	8	+1	-1	+1	+1	+1	3	0.6	คัดเลือก
	9	+1	+1	+1	+1	+1	5	1	คัดเลือก
	10	+1	+1	+1	+1	+1	5	1	คัดเลือก
	11	+1	+1	+1	+1	+1	5	1	คัดเลือก
	12	+1	+1	+1	+1	+1	5	1	คัดเลือก
	13	0	+1	+1	+1	+1	4	0.8	คัดเลือก
	14	+1	+1	+1	+1	+1	5	1	คัดเลือก
	15	+1	+1	+1	+1	+1	5	1	คัดเลือก
	16	+1	-1	+1	+1	+1	3	0.6	คัดเลือก
	17	+1	+1	+1	0	+1	4	0.8	คัดเลือก
	18	+1	+1	+1	+1	+1	5	1	คัดเลือก
ระบบสิ่งมีชีวิต	19	+1	0	0	+1	+1	3	0.6	คัดเลือก
	20	+1	+1	+1	+1	+1	5	1	คัดเลือก
	21	+1	+1	+1	+1	+1	5	1	คัดเลือก
	22	+1	+1	0	+1	+1	4	0.8	คัดเลือก
	23	+1	+1	+1	+1	+1	5	1	คัดเลือก
	24	+1	+1	+1	+1	+1	5	1	คัดเลือก

ตาราง 5 (ต่อ)

ความรู้ด้านเนื้อหา	ข้อที่	การพิจารณาของผู้เชี่ยวชาญ					รวม	IOC	แปลผล
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5			
ระบบสิ่งมีชีวิต	25	+1	-1	+1	+1	+1	3	0.6	คัดเลือก
	26	+1	+1	+1	+1	+1	5	1	คัดเลือก
	27	+1	+1	+1	+1	+1	5	1	คัดเลือก
	28	+1	+1	+1	0	+1	4	0.8	คัดเลือก
	29	+1	+1	+1	+1	+1	5	1	คัดเลือก
	30	+1	+1	+1	+1	+1	5	1	คัดเลือก
	31	+1	+1	+1	+1	+1	5	1	คัดเลือก
	32	+1	+1	+1	+1	+1	5	1	คัดเลือก
	33	+1	+1	+1	+1	+1	5	1	คัดเลือก
	34	+1	+1	+1	+1	+1	5	1	คัดเลือก
	35	+1	+1	+1	+1	+1	5	1	คัดเลือก
	36	+1	+1	+1	+1	+1	5	1	คัดเลือก
ระบบของโลก และอวกาศ	37	+1	+1	+1	+1	+1	5	1	คัดเลือก
	38	+1	+1	+1	+1	+1	5	1	คัดเลือก
	39	+1	+1	+1	+1	+1	5	1	คัดเลือก
	40	+1	+1	+1	+1	+1	5	1	คัดเลือก
	41	+1	+1	+1	+1	+1	5	1	คัดเลือก
	42	+1	+1	+1	+1	+1	5	1	คัดเลือก
	43	+1	+1	+1	+1	+1	5	1	คัดเลือก
	44	+1	+1	+1	+1	+1	5	1	คัดเลือก
	45	+1	+1	+1	+1	+1	5	1	คัดเลือก
	46	+1	+1	+1	+1	+1	5	1	คัดเลือก
	47	+1	+1	+1	+1	+1	5	1	คัดเลือก
	48	+1	+1	+1	+1	+1	5	1	คัดเลือก
	49	+1	+1	+1	+1	+1	5	1	คัดเลือก

ตาราง 5 (ต่อ)

ความรู้ด้านเนื้อหา	ข้อที่	การพิจารณาของผู้เชี่ยวชาญ					รวม	IOC	แปลผล
		คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5			
ระบบของโลกและอวกาศ	50	+1	+1	+1	+1	+1	5	1	คัดเลือก
	51	+1	+1	0	+1	+1	4	0.8	คัดเลือก
	52	+1	0	0	+1	+1	3	0.6	คัดเลือก
	53	+1	+1	+1	+1	+1	5	1	คัดเลือก
	54	+1	+1	+1	+1	+1	5	1	คัดเลือก

จากตาราง 5 ผลการพิจารณาความสอดคล้องของข้อสอบกับเนื้อหาของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 54 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 18 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ โดยผู้เชี่ยวชาญ จำนวน 5 คน พบว่า ข้อสอบผ่านเกณฑ์ทั้งหมด จำนวน 54 ข้อ และมีค่าตั้งแต่ 0.60 ถึง 1.00

1.2 ค่าความยาก Difficulty) ค่าอำนาจจำแนก (Discrimination) และค่าความเชื่อมั่น (Reliability)

ผู้วิจัยนำแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 54 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 18 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ โดยผ่านการพิจารณาจากผู้เชี่ยวชาญและแก้ไขแล้ว ไปทดสอบกับกลุ่มตัวอย่าง จำนวน 100 คน เพื่อวิเคราะห์ค่าความยาก (p) ค่าอำนาจจำแนก (r) และค่าความเชื่อมั่นด้วยโปรแกรมคอมพิวเตอร์ โดยพิจารณาคัดเลือกข้อสอบที่ค่าความยาก ตั้งแต่ 0.20 ถึง 0.80 ค่าอำนาจจำแนก ตั้งแต่ 0.20 ถึง 1.00 และค่าความเชื่อมั่นของแบบทดสอบ ปรากฏดังตาราง 6

ตาราง 6 การวิเคราะห์ค่าความยาก ค่าอำนาจจำแนก และค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

ความรู้ด้านเนื้อหา	ข้อที่	ความยาก (p)	อำนาจจำแนก (r)	ผลการพิจารณา	สรุปผล
ระบบทางกายภาพ	1	0.61	0.45	ผ่านเกณฑ์	คัดเลือก
	2	0.40	0.76	ผ่านเกณฑ์	คัดเลือก
	3	0.66	0.37	ผ่านเกณฑ์	คัดเลือก
	4	0.53	0.48	ผ่านเกณฑ์	คัดเลือก
	5	0.59	0.43	ผ่านเกณฑ์	คัดเลือก
	6	0.64	0.45	ผ่านเกณฑ์	คัดเลือก
	7	0.54	0.49	ผ่านเกณฑ์	คัดเลือก
	8	0.53	0.18	ไม่ผ่านเกณฑ์	ไม่คัดเลือก
	9	0.34	0.41	ผ่านเกณฑ์	คัดเลือก
	10	0.56	0.49	ผ่านเกณฑ์	คัดเลือก
	11	0.37	0.57	ผ่านเกณฑ์	คัดเลือก
	12	0.31	0.39	ผ่านเกณฑ์	คัดเลือก
	13	0.35	0.56	ผ่านเกณฑ์	คัดเลือก
	14	0.48	0.52	ผ่านเกณฑ์	คัดเลือก
	15	0.46	0.38	ผ่านเกณฑ์	คัดเลือก
	16	0.24	0.19	ไม่ผ่านเกณฑ์	ไม่คัดเลือก
	17	0.46	0.67	ผ่านเกณฑ์	คัดเลือก
	18	0.54	0.17	ไม่ผ่านเกณฑ์	ไม่คัดเลือก
ระบบสิ่งมีชีวิต	19	0.46	0.77	ผ่านเกณฑ์	คัดเลือก
	20	0.48	0.77	ผ่านเกณฑ์	คัดเลือก
	21	0.64	0.45	ผ่านเกณฑ์	คัดเลือก
	22	0.48	0.40	ผ่านเกณฑ์	คัดเลือก
	23	0.44	0.60	ผ่านเกณฑ์	คัดเลือก
	24	0.29	0.31	ผ่านเกณฑ์	คัดเลือก
	25	0.70	0.42	ผ่านเกณฑ์	คัดเลือก
	26	0.50	0.18	ไม่ผ่านเกณฑ์	ไม่คัดเลือก
	27	0.63	0.21	ผ่านเกณฑ์	คัดเลือก

ตาราง 6 (ต่อ)

ความรู้ด้านเนื้อหา	ข้อที่	ความยาก (p)	อำนาจจำแนก (r)	ผลการพิจารณา	สรุปผล
ระบบสิ่งมีชีวิต	28	0.79	0.19	ไม่ผ่านเกณฑ์	ไม่คัดเลือก
	29	0.46	0.73	ผ่านเกณฑ์	คัดเลือก
	30	0.36	0.30	ผ่านเกณฑ์	คัดเลือก
	31	0.43	0.45	ผ่านเกณฑ์	คัดเลือก
	32	0.40	0.18	ไม่ผ่านเกณฑ์	ไม่คัดเลือก
	33	0.42	0.28	ผ่านเกณฑ์	คัดเลือก
	34	0.39	0.29	ผ่านเกณฑ์	คัดเลือก
	35	0.51	0.47	ผ่านเกณฑ์	คัดเลือก
	36	0.54	0.61	ผ่านเกณฑ์	คัดเลือก
ระบบของโลกและอวกาศ	37	0.61	0.22	ผ่านเกณฑ์	คัดเลือก
	38	0.43	0.34	ผ่านเกณฑ์	คัดเลือก
	39	0.35	0.54	ผ่านเกณฑ์	คัดเลือก
	40	0.49	0.35	ผ่านเกณฑ์	คัดเลือก
	41	0.23	0.10	ไม่ผ่านเกณฑ์	ไม่คัดเลือก
	42	0.48	0.36	ผ่านเกณฑ์	คัดเลือก
	43	0.22	0.19	ไม่ผ่านเกณฑ์	ไม่คัดเลือก
	44	0.58	0.47	ผ่านเกณฑ์	คัดเลือก
	45	0.39	0.25	ผ่านเกณฑ์	คัดเลือก
	46	0.50	0.42	ผ่านเกณฑ์	คัดเลือก
	47	0.60	0.32	ผ่านเกณฑ์	คัดเลือก
	48	0.35	0.29	ผ่านเกณฑ์	คัดเลือก
	49	0.50	0.36	ผ่านเกณฑ์	คัดเลือก
	50	0.56	0.21	ผ่านเกณฑ์	คัดเลือก
	51	0.40	0.42	ผ่านเกณฑ์	คัดเลือก
	52	0.57	0.18	ไม่ผ่านเกณฑ์	ไม่คัดเลือก
	53	0.35	0.40	ผ่านเกณฑ์	คัดเลือก
	54	0.45	0.29	ผ่านเกณฑ์	คัดเลือก

ค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ = 0.92

จากตาราง 6 ผลการวิเคราะห์แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 54 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง พบว่า ระบบทางกายภาพ จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ รวมข้อสอบที่ผ่านเกณฑ์ทั้งหมด จำนวน 45 ข้อ ค่าความยาก (p) ตั้งแต่ 0.22 ถึง 0.79 ค่าอำนาจจำแนก (r) ตั้งแต่ 0.21 ถึง 0.77 และค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ 0.92

2. ผลการวิเคราะห์คุณภาพข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ตามแนวทฤษฎีการตอบสนองของข้อสอบ (Item Response Theory : IRT)

2.1 การตรวจสอบการแจกแจงแบบปกติ ปรากฏดังตาราง 7

ตาราง 7 การตรวจสอบการแจกแจงแบบปกติคะแนนแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำแนกตามตัวแปรเพศ

เพศ	Kolmogorov-Smirnov(a)			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
1	0.070	336	0.000	0.969	336	0.000
2	0.100	353	0.000	0.962	353	0.000

จากตาราง 7 ผลการวิเคราะห์การแจกแจงแบบปกติของคะแนนแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำแนกตามตัวแปรเพศ ได้แก่ เพศชาย (1) และเพศหญิง (2) พบว่า ทั้ง 2 กลุ่มไม่มีการแจกแจงแบบปกติ ซึ่งทดสอบด้วยสถิติ Kolmogorov-Smirnov Test แล้วพบว่ามีค่า Sig. (0.000 และ 0.000) น้อยกว่าระดับนัยสำคัญ (.05) แสดงว่าข้อมูลมีการกระจายเป็นแบบปกติไม่ใช่โค้งปกติ

2.2 การวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson’s Product Moment Correlation Coefficient) ปรากฏดังตาราง 8

ตาราง 8 การวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson’s Product Moment Correlation Coefficient) รายข้อ

ข้อ	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Item 12	Item 13	Item 14	Item 15
Item 1	1														
Item 2	.178(**)	1													
Item 3	.200(**)	.271(**)	1												
Item 4	.277(**)	.124(**)	.187(**)	1											
Item 5	.205(**)	.227(**)	.231(**)	.278(**)	1										
Item 6	.127(**)	.179(**)	.247(**)	.298(**)	.284(**)	1									
Item 7	.146(**)	.132(**)	.118(**)	.348(**)	.344(**)	.197(**)	1								
Item 8	.149(**)	.200(**)	.154(**)	.165(**)	.204(**)	.201(**)	.196(**)	1							
Item 9	.133(**)	.172(**)	.137(**)	.292(**)	.296(**)	.221(**)	.233(**)	.348(**)	1						
Item 10	.183(**)	.181(**)	.156(**)	.299(**)	.225(**)	.282(**)	0.07	.128(**)	.200(**)	1					
Item 11	.168(**)	.109(**)	.099(**)	.406(**)	.326(**)	.251(**)	.331(**)	.110(**)	.248(**)	.272(**)	1				
Item 12	.225(**)	.127(**)	.141(**)	.282(**)	.197(**)	.117(**)	.108(**)	.162(**)	.138(**)	.193(**)	.169(**)	1			
Item 13	.195(**)	.141(**)	.130(**)	.444(**)	.363(**)	.307(**)	.371(**)	.206(**)	.331(**)	.203(**)	.337(**)	.196(**)	1		
Item 14	.110(**)	.370(**)	.178(**)	.145(**)	.129(**)	.196(**)	.127(**)	.192(**)	.167(**)	.078(*)	.075(*)	.127(**)	.233(**)	1	
Item 15	.277(**)	.274(**)	.262(**)	.260(**)	.460(**)	.470(**)	.357(**)	.351(**)	.374(**)	.390(**)	.427(**)	.147(**)	.374(**)	.200(**)	1
Item 16	.122(**)	0.016	0.036	.210(**)	-0.011	0.015	-.098(*)	.096(*)	.186(**)	.159(**)	.114(**)	.172(**)	0.024	0.067	-.084(*)
Item 17	.156(**)	0.046	.075(*)	.169(**)	0.055	0.025	-.106(**)	.150(**)	.256(**)	.155(**)	.087(*)	.140(**)	0.048	0.021	-0.015
Item 18	.135(**)	0.048	.080(*)	.154(**)	0.017	0.058	-.081(*)	.181(**)	.260(**)	.134(**)	.105(**)	.155(**)	0.057	0.06	-0.006
Item 19	.134(**)	0.033	0.045	.160(**)	0.041	0.064	-.102(**)	.147(**)	.243(**)	.168(**)	0.073	.152(**)	0.042	0.026	-0.018
Item 20	.112(**)	0.002	0.032	.162(**)	0.018	0.006	-.080(*)	.172(**)	.220(**)	.128(**)	.094(*)	.182(**)	0.036	0.022	-0.011
Item 21	.107(**)	0.043	0.036	.174(**)	-0.005	0.027	-0.068	.121(**)	.234(**)	.167(**)	.079(*)	.168(**)	0.03	0.036	-0.033
Item 22	.121(**)	0.008	0.057	.177(**)	0.021	0.017	-.078(*)	.148(**)	.218(**)	.147(**)	.090(*)	.177(**)	0.004	0.047	-0.038
Item 23	.138(**)	0.036	0.032	.117(**)	-0.044	0.028	-.153(**)	0.003	0.013	.178(**)	0.052	.087(*)	-0.049	0.053	-.095(*)
Item 24	.118(**)	-.084(*)	-0.004	-0.035	-.094(*)	-.082(*)	-.167(**)	-0.04	-0.006	0.05	-0.012	.119(**)	-.115(**)	-0.012	-.138(**)
Item 25	.112(**)	0.014	-0.009	-0.011	-0.032	-0.018	-.100(**)	-0.029	-0.024	.092(*)	-0.012	.110(**)	-0.055	0.064	-.101(**)
Item 26	0.042	.100(**)	0.065	.089(*)	-0.029	-0.02	-.104(**)	0.039	0.065	.126(**)	0.052	.082(*)	-0.054	.076(*)	-0.074
Item 27	0.067	-0.046	0.053	0.041	-0.039	-0.062	-0.065	0.029	.077(*)	.100(**)	0.047	.101(**)	-0.05	-0.013	-0.05
Item 28	.129(**)	-0.018	-0.012	.229(**)	0.033	0.061	-0.046	.081(*)	.161(**)	.093(*)	.114(**)	.136(**)	0.055	0.003	-0.048
Item 29	.116(**)	-0.023	-0.015	.237(**)	-0.002	0.019	-0.034	.098(*)	.194(**)	.142(**)	.116(**)	.167(**)	0.07	0.003	-0.043
Item 30	.140(**)	-0.016	-0.01	.234(**)	0.046	0.066	-0.04	0.073	.177(**)	.124(**)	.106(**)	.162(**)	0.061	-0.01	-0.041
Item 31	.101(**)	.088(*)	0.004	.135(**)	0.01	0.044	-0.018	.086(*)	.130(**)	.098(*)	0.036	.105(**)	0.061	0.045	-0.031
Item 32	.077(*)	-0.027	-0.001	.171(**)	-0.013	0.002	-0.014	.133(**)	.178(**)	.093(*)	.077(*)	.126(**)	0.045	-0.009	-0.046
Item 33	0.04	0.029	0.032	0.066	-0.041	-0.003	-0.056	-0.047	0.041	.114(**)	.076(*)	0.061	0.023	-0.018	-.084(*)
Item 34	.143(**)	0.025	0.051	.137(**)	-0.037	0.018	-0.068	-0.004	.117(**)	0.065	0.064	.094(*)	-0.018	0.01	-.078(*)
Item 35	0.072	0.014	-0.019	.102(**)	0.01	0.01	-.106(**)	0.022	0.027	.084(*)	0.004	.110(**)	-0.065	0.025	-.112(**)

ตาราง 8 (ต่อ)

ข้อ	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Item 12	Item 13	Item 14	Item 15
Item 36	0.05	0.003	0.005	.114(**)	-0.061	0.061	-0.025	0.048	.094(*)	0.016	0.029	.101(**)	0.012	0.002	0.017
Item 37	.076(*)	0.02	0.061	.141(**)	-0.032	-0.012	-.084(*)	0.018	0.065	0.068	0.042	.120(**)	-0.07	0.005	-.079(*)
Item 38	0.074	0.067	0.006	.121(**)	0.03	0.063	-0.026	.110(**)	.161(**)	0.072	.088(*)	.088(*)	0.037	.121(**)	0.016
Item 39	0.065	-0.049	-0.006	.102(**)	-.094(*)	0.024	-.106(**)	0.06	.100(**)	.111(**)	0.018	.089(*)	0.013	0.011	-0.074
Item 40	0.074	-0.04	0.041	.168(**)	-0.069	0.002	-.115(**)	-0.034	0.001	0.053	-0.015	.165(**)	-0.036	0.022	-.188(**)
Item 41	.084(*)	-0.008	-0.045	.170(**)	-0.042	-0.033	-.090(*)	-0.035	0.061	.082(*)	0.066	.144(**)	0.011	0.015	-.166(**)
Item 42	.085(*)	-0.039	0.023	.230(**)	-.088(*)	-.105(**)	-.176(**)	-.158(**)	-.180(**)	.143(**)	0.058	.190(**)	-.076(*)	-.078(*)	-.267(**)
Item 43	0.021	0.032	0.006	.140(**)	0.039	-0.016	-0.02	0.059	0.074	.081(*)	0.053	.089(*)	0.057	0.041	-0.068
Item 44	0.012	0.039	.101(**)	-0.053	0.056	0.055	0.003	.161(**)	.131(**)	0.01	0.034	0.032	0.018	.080(*)	.140(**)
Item 45	-0.012	0.031	0.033	0.056	0.067	0.016	-0.02	.112(**)	.124(**)	0.02	0.02	0.041	0.046	0.02	-0.02

ตาราง 8 (ต่อ)

ข้อ	Item 16	Item 17	Item 18	Item 19	Item 20	Item 21	Item 22	Item 23	Item 24	Item 25	Item 26	Item 27	Item 28	Item 29	Item 30
Item 16	1														
Item 17	.782(**)	1													
Item 18	.789(**)	.830(**)	1												
Item 19	.753(**)	.780(**)	.783(**)	1											
Item 20	.782(**)	.814(**)	.778(**)	.716(**)	1										
Item 21	.787(**)	.822(**)	.799(**)	.763(**)	.745(**)	1									
Item 22	.794(**)	.783(**)	.773(**)	.724(**)	.759(**)	.767(**)	1								
Item 23	.487(**)	.459(**)	.448(**)	.453(**)	.448(**)	.460(**)	.433(**)	1							
Item 24	.487(**)	.400(**)	.398(**)	.404(**)	.407(**)	.402(**)	.411(**)	.307(**)	1						
Item 25	.259(**)	.235(**)	.222(**)	.212(**)	.252(**)	.212(**)	.217(**)	.165(**)	.209(**)	1					
Item 26	.537(**)	.459(**)	.476(**)	.454(**)	.458(**)	.471(**)	.452(**)	.326(**)	.323(**)	.175(**)	1				
Item 27	.420(**)	.410(**)	.391(**)	.396(**)	.399(**)	.394(**)	.394(**)	.203(**)	.296(**)	.100(**)	.163(**)	1			
Item 28	.737(**)	.614(**)	.599(**)	.598(**)	.635(**)	.619(**)	.593(**)	.376(**)	.394(**)	.224(**)	.412(**)	.279(**)	1		
Item 29	.726(**)	.604(**)	.600(**)	.598(**)	.613(**)	.598(**)	.605(**)	.349(**)	.367(**)	.213(**)	.408(**)	.330(**)	.692(**)	1	
Item 30	.747(**)	.633(**)	.618(**)	.628(**)	.644(**)	.649(**)	.655(**)	.372(**)	.405(**)	.206(**)	.393(**)	.304(**)	.886(**)	.727(**)	1
Item 31	.448(**)	.379(**)	.381(**)	.386(**)	.360(**)	.411(**)	.365(**)	.187(**)	.232(**)	.131(**)	.233(**)	.205(**)	.373(**)	.388(**)	.392(**)
Item 32	.737(**)	.666(**)	.655(**)	.596(**)	.655(**)	.676(**)	.650(**)	.363(**)	.376(**)	.198(**)	.418(**)	.359(**)	.594(**)	.582(**)	.592(**)
Item 33	.444(**)	.416(**)	.422(**)	.430(**)	.415(**)	.457(**)	.397(**)	.313(**)	.260(**)	.183(**)	.309(**)	.234(**)	.348(**)	.326(**)	.319(**)
Item 34	.470(**)	.434(**)	.459(**)	.430(**)	.446(**)	.411(**)	.390(**)	.325(**)	.294(**)	.210(**)	.360(**)	.265(**)	.381(**)	.366(**)	.423(**)
Item 35	.528(**)	.425(**)	.439(**)	.420(**)	.436(**)	.432(**)	.413(**)	.352(**)	.321(**)	.216(**)	.338(**)	.270(**)	.415(**)	.373(**)	.399(**)
Item 36	.332(**)	.311(**)	.370(**)	.341(**)	.338(**)	.368(**)	.316(**)	.252(**)	.174(**)	.160(**)	.260(**)	.208(**)	.295(**)	.327(**)	.340(**)
Item 37	.515(**)	.452(**)	.455(**)	.447(**)	.443(**)	.438(**)	.451(**)	.270(**)	.333(**)	.206(**)	.370(**)	.247(**)	.415(**)	.415(**)	.440(**)
Item 38	.629(**)	.543(**)	.547(**)	.505(**)	.489(**)	.518(**)	.531(**)	.302(**)	.279(**)	.127(**)	.366(**)	.267(**)	.470(**)	.473(**)	.488(**)
Item 39	.456(**)	.404(**)	.418(**)	.420(**)	.405(**)	.452(**)	.413(**)	.295(**)	.271(**)	.118(**)	.288(**)	.193(**)	.398(**)	.373(**)	.407(**)
Item 40	.542(**)	.421(**)	.407(**)	.387(**)	.421(**)	.425(**)	.451(**)	.380(**)	.457(**)	.217(**)	.324(**)	.249(**)	.452(**)	.436(**)	.462(**)
Item 41	.635(**)	.468(**)	.494(**)	.504(**)	.509(**)	.504(**)	.478(**)	.359(**)	.457(**)	.263(**)	.383(**)	.310(**)	.564(**)	.557(**)	.567(**)
Item 42	.393(**)	.281(**)	.278(**)	.285(**)	.298(**)	.302(**)	.310(**)	.377(**)	.285(**)	.240(**)	.316(**)	.260(**)	.311(**)	.305(**)	.329(**)
Item 43	.484(**)	.401(**)	.397(**)	.396(**)	.428(**)	.404(**)	.421(**)	.243(**)	.236(**)	.171(**)	.253(**)	.197(**)	.411(**)	.365(**)	.398(**)
Item 44	.178(**)	.231(**)	.241(**)	.241(**)	.201(**)	.210(**)	.187(**)	.113(**)	.101(**)	.090(*)	.137(**)	.152(**)	.083(*)	.122(**)	.108(**)
Item 45	.391(**)	.382(**)	.396(**)	.357(**)	.363(**)	.358(**)	.350(**)	.171(**)	.153(**)	.132(**)	.199(**)	.131(**)	.310(**)	.247(**)	.294(**)

ตาราง 8 (ต่อ)

ข้อ	Item 31	Item 32	Item 33	Item 34	Item 35	Item 36	Item 37	Item 38	Item 39	Item 40	Item 41	Item 42	Item 43	Item 44	Item 45
Item 31	1														
Item 32	.298(**)	1													
Item 33	.147(**)	.332(**)	1												
Item 34	.208(**)	.379(**)	.182(**)	1											
Item 35	.256(**)	.390(**)	.283(**)	.218(**)	1										
Item 36	.134(**)	.386(**)	.200(**)	.229(**)	.229(**)	1									
Item 37	.224(**)	.397(**)	.243(**)	.390(**)	.374(**)	.221(**)	1								
Item 38	.259(**)	.485(**)	.292(**)	.294(**)	.351(**)	.180(**)	.248(**)	1							
Item 39	.204(**)	.350(**)	.246(**)	.160(**)	.212(**)	.242(**)	.217(**)	.273(**)	1						
Item 40	.226(**)	.419(**)	.228(**)	.286(**)	.334(**)	.227(**)	.325(**)	.297(**)	.253(**)	1					
Item 41	.355(**)	.470(**)	.290(**)	.349(**)	.340(**)	.236(**)	.365(**)	.338(**)	.261(**)	.463(**)	1				
Item 42	.142(**)	.287(**)	.256(**)	.219(**)	.245(**)	.217(**)	.272(**)	.181(**)	.232(**)	.367(**)	.347(**)	1			
Item 43	.259(**)	.362(**)	.152(**)	.180(**)	.221(**)	.131(**)	.241(**)	.243(**)	.192(**)	.247(**)	.327(**)	.237(**)	1		
Item 44	0.021	.173(**)	0.059	.103(**)	.098(*)	.223(**)	0.068	.120(**)	.098(*)	0.001	0.03	-0.018	0.029	1	
Item 45	.207(**)	.318(**)	.117(**)	.143(**)	.184(**)	.106(**)	.142(**)	.239(**)	.153(**)	.131(**)	.222(**)	.093(*)	.828(**)	0.045	1

จากตาราง 8 ผลการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ พบว่า ความสัมพันธ์ระหว่างตัวแปรทั้งหมดแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรมีค่าอยู่ระหว่าง -0.188 ถึง 0.886 เมื่อพิจารณาค่าสถิติ Bartlett's test of Sphericity มีค่าเท่ากับ 16,018.970 (p = .000) แสดงว่าเมทริกซ์สหสัมพันธ์ดังกล่าวแตกต่างจากเมทริกซ์เอกลักษณ์อย่างมีนัยสำคัญ ผลการวิเคราะห์สอดคล้องกับค่าดัชนีไกเซอร์-เมเยอร์-ออลคิน ซึ่งมีค่าเท่ากับ 0.937 ซึ่งเป็นค่าที่เข้าใกล้ 1 แสดงว่าตัวแปรต่าง ๆ ในข้อมูลนี้มีความสัมพันธ์กันและเหมาะสมที่จะไปวิเคราะห์สมการโครงสร้างได้

2.3 การวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ

ผู้วิจัยนำแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 45 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 15 ข้อ ระบบสิ่งมีชีวิต จำนวน 15 ข้อ และระบบของโลกและอวกาศ จำนวน 15 ข้อ มาวิเคราะห์การประมาณค่าพารามิเตอร์ความยากของขอสอบ และค่าอำนาจจำแนกของข้อสอบ โดยประยุกต์ใช้โปรแกรม Mplus ซึ่งประมาณได้จากโมเดลตามทฤษฎีการตอสนองของข้อสอบ ปรากฏดังตาราง 9

การประมาณค่าพารามิเตอร์ของข้อสอบ ซึ่งมีเกณฑ์การคัดเลือกข้อสอบ ดังนี้ (ศิริชัย กาญจนวาสี, 2555)

ค่าอำนาจจำแนกข้อสอบ (a-parameter) มีค่าอยู่ระหว่าง 0.50 ถึง 2.50

ค่าความยากของข้อสอบ (b-parameter) มีค่าอยู่ระหว่าง -2.50 ถึง 2.50 ค่าที่เป็นลบแสดงว่าข้อสอบง่าย และค่าที่เป็นบวกแสดงว่าข้อสอบยาก

ตาราง 9 การวิเคราะห์ค่าความยากของข้อสอบ (b) ค่าอำนาจจำแนกของข้อสอบ (a) แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

ความรู้ด้านเนื้อหา	ข้อที่	ค่าความยากของข้อสอบ (b)	ค่าอำนาจจำแนกของข้อสอบ (a)	สรุปผล
	1	-0.537	0.744	คัดเลือก
	2	0.155	1.162	คัดเลือก
	3	1.568	0.937	คัดเลือก
	4	1.877	0.587	คัดเลือก
	5	1.686	0.672	คัดเลือก
	6	0.098	0.604	คัดเลือก
	7	1.827	1.170	คัดเลือก
ระบบทางกายภาพ	8	2.775	1.241	คัดเลือก
	9	0.784	0.865	คัดเลือก
	10	3.174	0.612	คัดเลือก
	11	1.665	0.313	คัดเลือก
	12	0.420	1.013	คัดเลือก
	13	0.824	0.370	คัดเลือก
	14	0.890	0.495	คัดเลือก
	15	0.241	0.543	คัดเลือก
	16	0.823	0.670	คัดเลือก
	17	3.002	0.902	คัดเลือก
ระบบสิ่งมีชีวิต	18	2.273	0.364	คัดเลือก
	19	1.843	0.944	คัดเลือก
	20	1.592	0.486	คัดเลือก

ตาราง 9 (ต่อ)

ความรู้ด้านเนื้อหา	ข้อที่	ค่าความยากของข้อสอบ (b)	ค่าอำนาจจำแนกของข้อสอบ (a)	สรุปผล
ระบบสิ่งมีชีวิต	21	1.702	0.953	คัดเลือก
	22	2.015	0.903	คัดเลือก
	23	1.586	0.429	คัดเลือก
	24	0.748	1.164	คัดเลือก
	25	2.045	0.846	คัดเลือก
	26	2.796	0.887	คัดเลือก
	27	1.847	0.966	คัดเลือก
	28	4.000	0.640	คัดเลือก
	29	0.958	1.119	คัดเลือก
	30	1.855	0.757	คัดเลือก
ระบบของโลกและอวกาศ	31	-0.266	1.381	คัดเลือก
	32	-1.094	0.686	คัดเลือก
	33	2.297	1.140	คัดเลือก
	34	2.714	1.197	คัดเลือก
	35	0.549	0.811	คัดเลือก
	36	2.265	1.037	คัดเลือก
	37	-0.278	0.741	คัดเลือก
	38	0.401	1.382	คัดเลือก
	39	0.561	0.372	คัดเลือก
	40	-0.328	0.539	คัดเลือก
	41	2.179	0.666	คัดเลือก
	42	3.939	0.901	คัดเลือก
	43	0.454	0.833	คัดเลือก
	44	1.040	0.833	คัดเลือก
	45	-0.016	1.056	คัดเลือก

จากตาราง 9 ผลการวิเคราะห์แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 45 ข้อ พบว่า ค่าความยากของข้อสอบ (b-parameter) มีค่าตั้งแต่ -0.537 ถึง 4.000 และค่าอำนาจจำแนกของข้อสอบ (a-parameter) มีค่าตั้งแต่ 0.313 ถึง 1.382

ตอนที่ 2 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำแนกตามเพศ

ผู้วิจัยนำแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 45 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 15 ข้อ ระบบสิ่งมีชีวิต จำนวน 15 ข้อ และระบบของโลกและอวกาศ จำนวน 15 ข้อ เพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ผู้วิจัยได้ใช้วิธี MIMIC จำแนกตามเพศ โดยประยุกต์ใช้โปรแกรม Mplus ปรากฏดังตาราง 10

ตาราง 10 ผลการตรวจสอบความสอดคล้องกลมกลืนของโมเดล CFA ของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC จำแนกตามเพศ

ความรู้ด้านเนื้อหา	การวิเคราะห์ความกลมกลืนของโมเดล CFA วิธี MIMIC						
	χ^2	df	P-Value	CFI	TLI	RMSEA	SRMR
ระบบทางกายภาพ	2.224	82	0.0000	0.976	0.965	0.042	0.051
ระบบสิ่งมีชีวิต	1.720	74	0.0001	0.999	0.999	0.032	0.024
ระบบของโลกและอวกาศ	2.086	75	0.0000	0.969	0.951	0.040	0.050

จากตาราง 10 ผลการตรวจสอบความสอดคล้องกลมกลืนของโมเดล CFA วิธี MIMIC ของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC จำแนกตามเพศ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ ระบบสิ่งมีชีวิต และระบบของโลกและอวกาศ พบว่า ระบบทางกายภาพ พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 2.224 องศาอิสระ (df) มีค่าเท่ากับ 82 P-Value มีค่าเท่ากับ 0.0000 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI) มีค่าเท่ากับ 0.976 ดัชนี The Tucker Lewis Index (TLI) มีค่าเท่ากับ 0.965 ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของการประมาณค่า (RMSEA) มีค่าเท่ากับ 0.042 ดัชนี Standardized Root Mean Square Residual มีค่าเท่ากับ 0.051

ระบบสิ่งมีชีวิต พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 1.720 องศาอิสระ (df) มีค่าเท่ากับ 74 P-Value มีค่าเท่ากับ 0.0001 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI)

มีค่าเท่ากับ 0.999 ดัชนี The Tucker Lewis Index (TLI) มีค่าเท่ากับ 0.999 ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของการประมาณค่า (RMSEA) มีค่าเท่ากับ 0.032 ดัชนี Standardized Root Mean Square Residual มีค่าเท่ากับ 0.024

ระบบของโลกและอวกาศ พิจารณาจากค่าไค-สแควร์ (Chi-square) มีค่าเท่ากับ 2.086 องศาอิสระ (df) มีค่าเท่ากับ 75 ดัชนีวัดระดับความกลมกลืนเปรียบเทียบ (CFI) มีค่าเท่ากับ 0.969 ดัชนี The Tucker Lewis Index (TLI) มีค่าเท่ากับ 0.951 ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของการประมาณค่า (RMSEA) มีค่าเท่ากับ 0.040 ดัชนี Standardized Root Mean Square Residual มีค่าเท่ากับ 0.050

ตาราง 11 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC

ความรู้ด้านเนื้อหา	ข้อที่	ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์		
		Estimate	P-Value	สรุปผล
ระบบทางกายภาพ	1	0.063	0.183	NO DIF
	2	0.002	0.974	NO DIF
	3	-0.009	0.852	NO DIF
	4	0.078	0.147	NO DIF
	5	0.128*	0.007	DIF
	6	-0.028	0.566	NO DIF
	7	0.085	0.079	NO DIF
	8	0.111*	0.024	DIF
	9	0.092	0.060	NO DIF
	10	0.076	0.109	NO DIF
	11	0.025	0.605	NO DIF
	12	0.060	0.204	NO DIF
	13	0.070	0.154	NO DIF
	14	0.059	0.217	NO DIF
	15	0.086	0.081	NO DIF

ตาราง 11 (ต่อ)

ความรู้ด้านเนื้อหา	ข้อที่	ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์		
		Estimate	P-Value	สรุปผล
ระบบสิ่งมีชีวิต	16	0.232*	0.000	DIF
	17	0.256*	0.000	DIF
	18	0.196*	0.001	DIF
	19	0.199*	0.000	DIF
	20	0.167*	0.003	DIF
	21	0.287*	0.000	DIF
	22	0.229*	0.000	DIF
	23	0.122*	0.010	DIF
	24	0.190*	0.000	DIF
	25	0.041	0.391	NO DIF
	26	0.149*	0.002	DIF
	27	0.241*	0.000	DIF
	28	0.228*	0.000	DIF
	29	0.413*	0.000	DIF
	30	0.248*	0.000	DIF
ระบบของโลกและอวกาศ	31	-0.394*	0.000	DIF
	32	-0.175*	0.000	DIF
	33	-0.181*	0.003	DIF
	34	0.380*	0.000	DIF
	35	-0.050	0.293	NO DIF
	36	-0.267*	0.000	DIF
	37	0.297*	0.000	DIF
	38	0.027	0.576	NO DIF
	39	-0.002	0.973	NO DIF

ตาราง 11 (ต่อ)

ความรู้ด้านเนื้อหา	ข้อที่	ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์		
		Estimate	P-Value	สรุปผล
ระบบของโลกและอวกาศ	40	-0.017	0.728	NO DIF
	41	-0.428*	0.000	DIF
	42	-0.008	0.864	NO DIF
	43	-0.133*	0.005	DIF
	44	-0.598*	0.000	DIF
	45	0.077	0.105	NO DIF

หมายเหตุ * นัยสำคัญทางสถิติที่ระดับ .05

จากตาราง 11 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC พบว่า ระบบทางกายภาพ จำนวน 15 ข้อ ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 2 ข้อ ได้แก่ 5 และ 8 คิดเป็นร้อยละ 13.33 ของข้อสอบทั้งหมด

ระบบสิ่งมีชีวิต จำนวน 15 ข้อ ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 14 ข้อ ได้แก่ 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29 และ 30 คิดเป็นร้อยละ 93.33 ของข้อสอบทั้งหมด

ระบบของโลกและอวกาศ จำนวน 15 ข้อ ตรวจพบข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 9 ข้อ ได้แก่ 31, 32, 33, 34, 36, 37, 41, 43 และ 44 คิดเป็นร้อยละ 60.00 ของข้อสอบทั้งหมด

พหุนํ ปณฺฑิโต ชีเว

ตาราง 12 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC จำแนกตามเพศ

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ด้วยวิธี MIMIC จำแนกตามเพศ	
ชาย (Reference)	หญิง (Focal)
5, 8, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29, 30, 34, 37	31, 32, 33, 36, 41, 43, 44

หมายเหตุ เพศชาย = 1 (Reference Group), เพศหญิง = 2 (Focal Group)

จากตาราง 12 ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 45 ข้อ ด้วยวิธี MIMIC จำแนกตามเพศ โดยใช้โปรแกรมประยุกต์ Mplus พบว่า เพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 18 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 7 ข้อ

บทที่ 5

สรุปผล อภิปรายผล และข้อเสนอแนะ

การวิจัยเรื่อง การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 นำเสนอสรุป อภิปรายผล และข้อเสนอแนะตามหัวข้อดังต่อไปนี้

1. ความมุ่งหมายของการวิจัย
2. สรุปผล
3. อภิปรายผล
4. ข้อเสนอแนะ

ความมุ่งหมายของการวิจัย

การวิจัยครั้งนี้มีจุดมุ่งหมาย เพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 โดยมีจุดมุ่งหมายเฉพาะ ดังนี้

1. เพื่อสร้างและหาคุณภาพแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

2. เพื่อตรวจสอบการทำหน้าที่ต่างกัน (DIF) ของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

สรุปผล

1. การสร้างและหาคุณภาพแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

1.1 ผลการวิเคราะห์คุณภาพข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ตามแนวทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory : CTT)

1.1.1 ความเที่ยงตรง (Validity)

ผู้วิจัยได้สร้างแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 1 ฉบับ 54 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 18 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ แล้วนำไปให้ผู้เชี่ยวชาญ จำนวน 5 คน พิจารณาความสอดคล้อง พบว่า ข้อสอบผ่านเกณฑ์ทุกข้อ มีค่าตั้งแต่ 0.60

ถึง 1.00 จึงสามารถนำข้อสอบทั้ง 54 ข้อ ไปทดลองกลุ่มตัวอย่าง จำนวน 100 คน เพื่อหาคุณภาพเป็นรายข้อและทั้งฉบับ

1.1.2 ค่าความยาก (Difficulty) ค่าอำนาจจำแนก (Discrimination) และค่าความเชื่อมั่น (Reliability)

การวิเคราะห์ค่าความยาก และค่าอำนาจจำแนก ของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง พบว่า ระบบทางกายภาพ จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ รวมข้อสอบที่ผ่านเกณฑ์ทั้งหมด จำนวน 45 ข้อ มีค่าความยาก (p) ตั้งแต่ 0.22 ถึง 0.79 ค่าอำนาจจำแนก (r) ตั้งแต่ 0.21 ถึง 0.77 และค่าความเชื่อมั่นทั้งฉบับมีค่าเท่ากับ 0.92

1.2 ผลการวิเคราะห์คุณภาพข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ตามแนวทฤษฎีการตอบสนองของข้อสอบ (Item Response Theory : IRT)

การวิเคราะห์แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 45 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 15 ข้อ ระบบสิ่งมีชีวิต จำนวน 15 ข้อ และระบบของโลกและอวกาศ จำนวน 15 ข้อ ที่เป็นกลุ่มตัวอย่าง จำนวน 689 คน พบว่า ค่าความยากของข้อสอบ (b-parameter) มีค่าตั้งแต่ -0.537 ถึง 4.000 และค่าอำนาจจำแนกของข้อสอบ (a-parameter) มีค่าตั้งแต่ 0.313 ถึง 1.382

2. การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

2.1 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC จำนวน 45 ข้อ กับกลุ่มตัวอย่าง จำนวน 689 คน พบว่า

ระบบทางกายภาพ จำนวน 15 ข้อ มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 2 ข้อ ได้แก่ 5 และ 8 คิดเป็นร้อยละ 13.33 ของข้อสอบทั้งหมด

ระบบสิ่งมีชีวิต จำนวน 15 ข้อ มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 14 ข้อ ได้แก่ 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29 และ 30 คิดเป็นร้อยละ 93.33 ของข้อสอบทั้งหมด

ระบบของโลกและอวกาศ จำนวน 15 ข้อ มีข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบ จำนวน 9 ข้อ ได้แก่ 31, 32, 33, 34, 36, 37, 41, 43 และ 44 คิดเป็นร้อยละ 60.00 ของข้อสอบทั้งหมด

2.2 การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำวน 45 ข้อ ด้วยวิธี MIMIC จำแนกตามเพศ โดยใช้โปรแกรมประยุกต์ Mplus พบว่า เพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 18 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 7 ข้อ

อภิปรายผล

การวิจัยครั้งนี้มีจุดมุ่งหมาย เพื่อตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 อภิปรายผลได้ดังนี้

1. ผลการสร้างและหาคุณภาพของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 1 ฉบับ 54 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง ได้แก่ ระบบทางกายภาพ จำนวน 18 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ ผู้วิจัยได้สร้าง วิเคราะห์ และตรวจสอบคุภาพของแบบทดสอบด้วยเทคนิคและวิธีการวิเคราะห์ข้อสอบตามแนวทฤษฎีการทดสอบแบบดั้งเดิม และทฤษฎีการตอบสนองของข้อสอบ ซึ่งทั้งสองแนวคิดต่างก็มีข้อดีและข้อเสียต่างกัน แต่เมื่อนำสองทฤษฎีมาใช้เป็นแนวทางในการวิเคราะห์ข้อสอบก็สามารถลดข้อจำกัดลงได้ จึงทำให้ได้แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ที่มีคุณภาพ รายละเอียดดังต่อไปนี้

1.1 ผลการวิเคราะห์คุณภาพข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ตามแนวทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory : CTT)

1.1.1 ค่าความเที่ยงตรง (Validity)

การหาความเที่ยงตรงเชิงเนื้อหาด้วยวิธีหาค่าเฉลี่ยความสอดคล้องความคิดเห็นของผู้เชี่ยวชาญ จำนวน 5 ท่าน พบว่า ดัชนีความสอดคล้องมีค่าตั้งแต่ 0.60 ถึง 1.00 แสดงให้เห็นว่า แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ข้อสอบแต่ละข้อมีความสอดคล้องกับนิยามที่ต้องการวัด โดยจุดประสงค์เชิงพฤติกรรมมีความครอบคลุมเนื้อหา จึงสรุปได้ว่าข้อสอบที่สร้างขึ้นมีความเที่ยงตรงเชิงเนื้อหา ซึ่งผู้วิจัยได้ดำเนินการสร้างข้อสอบให้สอดคล้องครอบคลุมตามกรอบเนื้อหาของ PISA 2015 และพฤติกรรมที่ต้องการวัด จึงทำให้ข้อสอบที่สร้างขึ้นมีความตรงเชิงเนื้อหา สอดคล้องกับแนวคิดของ อาร์แมนน์ และกล็อค (Ahmann and Glock, 1967 อ้างอิงมาจาก โชติ เพชรชื่น, 2544) กล่าวถึงลักษณะของแบบทดสอบวินิจฉัยไว้ว่าเน้นความเที่ยงตรงเชิงเนื้อหา (Content Validity) สอดคล้องกับแนวคิดของ ไพศาล วรคา (2558) กล่าวว่า ควรเลือกข้อคำถามที่มีค่าดัชนีความสอดคล้องตั้งแต่ 0.50 ขึ้นไป หากต่ำกว่า 0.50 ถือว่าใช้ไม่ได้ สอดคล้องกับงานวิจัยของ นิธิวดี วีระศิลป์ (2561) พบว่า แบบทดสอบความคิดสร้างสรรค์เชิง

ตัวเลขสำหรับนักเรียนระดับชั้นประกาศนียบัตรวิชาชีพชั้นปีที่ 1 มีค่าดัชนีความสอดคล้องตั้งแต่ 0.80 ถึง 1.00 สอดคล้องกับงานวิจัยของ ธิดารัตน์ ธนะขว้าง (2553) พบว่า แบบวัดความคิดสร้างสรรค์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 4 มีค่าดัชนีความสอดคล้องตั้งแต่ 0.71 ถึง 1.00

1.1.2 ค่าความยาก (Difficulty) และค่าอำนาจจำแนก (Discrimination)

ค่าความยากและค่าอำนาจจำแนก จากการทดสอบกับกลุ่มตัวอย่าง จำนวน 100 คน ของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 54 ข้อ ประกอบด้วย ความรู้ด้านเนื้อหา 3 เรื่อง พบว่า ระบบทางกายภาพ จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ ระบบสิ่งมีชีวิต จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ และระบบของโลกและอวกาศ จำนวน 18 ข้อ มีข้อสอบที่ผ่านเกณฑ์ จำนวน 15 ข้อ รวมข้อสอบที่ผ่านเกณฑ์ทั้งหมด จำนวน 45 ข้อ มีค่าความยากตั้งแต่ 0.22 ถึง 0.79 ค่าอำนาจจำแนก ตั้งแต่ 0.21 ถึง 0.77 ข้อสอบส่วนใหญ่มีค่าความยากและค่าอำนาจจำแนกรายข้ออยู่ในเกณฑ์ดี แต่มีบางข้อมีค่าความยากต่ำเกินไป แสดงว่าข้อสอบนั้นยาก ผู้วิจัยจึงคัดเลือกเฉพาะข้อสอบที่มีนัยสำคัญทางสถิติ ซึ่งมีค่าอำนาจจำแนกเข้าเกณฑ์ทุกข้อ จึงเชื่อได้ว่าแบบวัดทีละข้อมีอำนาจจำแนกอย่างแท้จริง และค่าอำนาจจำแนกนี้ได้แสดงให้เห็นว่าสามารถจำแนกนักเรียนกลุ่มสูงและกลุ่มต่ำออกจากกันได้ สอดคล้องกับ ล้วน สายยศ และอังคณา สายยศ (2543) ซึ่งกล่าวไว้ว่า ค่าความยากของข้อสอบควรอยู่ในเกณฑ์ 0.20 ถึง 0.80 สอดคล้องกับ สมนึก ภัททิยธนี (2551) ซึ่งกล่าวไว้ว่า ค่าความยากของข้อสอบควรอยู่ในเกณฑ์ 0.20 ถึง 0.80 และค่าอำนาจจำแนกของข้อสอบควรมีค่าตั้งแต่ 0.20 ถึง 1.00 และยังสอดคล้องกับงานวิจัยของ โชติกา จันทะวัน (2562) พบว่า แบบทดสอบความสามารถในการให้เหตุผลทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5 ในจังหวัดกาฬสินธุ์ มีค่าความยากรายข้อ ตั้งแต่ 0.34 ถึง 0.80 และมีค่าอำนาจจำแนกรายข้อตั้งแต่ 0.22 ถึง 0.65 แสดงว่า มีค่าตามเกณฑ์ทุกข้อ

1.1.3 ความเชื่อมั่นของแบบทดสอบ (Reliability)

การหาค่าความเชื่อมั่นของแบบทดสอบ ผู้วิจัยใช้วิธีของคูเดอร์-ริชาร์ดสัน โดยใช้แบบทดสอบฉบับเดียวดำเนินการสอบเพียงครั้งเดียว ใช้สูตร KR-20 (สมนึก ภัททิยธนี, 2553) ค่าความเชื่อมั่นของแบบทดสอบ เท่ากับ 0.92 จะเห็นว่า ค่าความเชื่อมั่นของแบบทดสอบวัดที่ผู้วิจัยสร้างขึ้นมีค่าสูงพอที่จะเชื่อถือได้ว่าเป็นแบบทดสอบวัดที่มีคุณภาพเหมาะสม สาเหตุที่เป็นเช่นนี้ เนื่องมาจากแบบทดสอบวัดได้ผ่านการคัดเลือกข้อที่เข้าเกณฑ์ปรับปรุงข้อคำถามให้เหมาะสมยิ่งขึ้น สอดคล้องกับแนวคิดของ ล้วน สายยศ และอังคณา สายยศ (2543) ซึ่งกล่าวว่า ค่าความเชื่อมั่นจะมีค่าอยู่ระหว่าง -1 ถึง +1 และจะพิจารณาเฉพาะค่าที่เป็นบวกเท่านั้น ซึ่งควรมีค่ามากกว่า 0.70 จึงจะเป็นแบบทดสอบที่มีความเชื่อมั่น แต่หากพิจารณาเป็นรายด้านจะพบว่า ค่าความเชื่อมั่นไม่ค่อยสูงนัก เนื่องจากจำนวนข้อของแต่ละด้านมีจำนวนน้อย ทำให้ค่าความเชื่อมั่นต่ำกว่า ความเชื่อมั่น

รวมทั้งฉบับ เนื่องจากจำนวนข้อสอบมีผลต่อค่าความเชื่อมั่นสอดคล้องกับแนวคิดของ สมนึก ภัททิยธนี (2546) กล่าวว่า ลักษณะของกลุ่มผู้สอบ ถ้ากลุ่มผู้สอบมีความสามารถไม่แตกต่างกันมาก จะทำให้ค่าความเชื่อมั่นต่ำ แต่ถ้ากลุ่มผู้สอบมีความสามารถแตกต่างกันมาก ค่าความเชื่อมั่นของแบบทดสอบจะสูง สอดคล้องกับงานวิจัยของ ยุพา สำเลิศรัมย์ (2558) พบว่า แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนรายวิชาวิทยาศาสตร์ สำหรับชั้นมัธยมศึกษาปีที่ 1 ตามแนวข้อสอบในโครงการประเมินผลนักเรียนร่วมกับนานาชาติ (PISA) จากการทดลองใช้ทั้ง 3 ครั้ง มีค่าความเชื่อมั่น 0.82, 0.85 และ 0.83 ตามลำดับ และสอดคล้องกับงานวิจัยของ จุฬาพรรณ ชุมพล (2555) พบว่า แบบทดสอบวินิจฉัยทางการเรียนวิทยาศาสตร์ เรื่อง แรงและพลังงาน สำหรับชั้นมัธยมศึกษาปีที่ 1 จำนวน 5 ฉบับ มีค่าความเชื่อมั่น 0.81, 0.82, 0.83, 0.8 และ 0.83 ตามลำดับ

1.2 ผลการวิเคราะห์คุณภาพข้อสอบของแบบทดสอบการรู้เรื่องวิทยาศาสตร์ตามแนวทฤษฎีการตอบสนองของข้อสอบ (Item Response Theory : IRT)

1.2.1 ค่าความยาก (b-parameter) ของข้อสอบ มีค่าระหว่าง -0.537 ถึง 4.000 หมายความว่า ค่าความยากของข้อสอบทุกข้ออยู่ในเกณฑ์ตามทฤษฎีกำหนด แสดงว่าแบบทดสอบนี้มีคุณภาพอยู่ในเกณฑ์ที่ใช้ได้ ซึ่ง ศิริชัย กาญจนวาสี (2555) กล่าวว่า ตามทฤษฎีการตอบสนองของข้อสอบ ค่าความยากมีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัตินิยมใช้ข้อสอบที่มีค่าความยากตั้งแต่ -2.50 ถึง +2.50 ค่า b ที่อยู่ใกล้ -2.50 แสดงว่าเป็นข้อสอบที่ง่าย ส่วนค่า b ที่อยู่ใกล้ +2.50 เป็นข้อสอบที่ยาก สอดคล้องกับแนวคิดของ Hambleton และ Cook (1977) Hambleton และ Swaminathan (1985) กล่าวว่า ค่าความยาก (b) มีพิสัยอยู่ระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติจะมีค่าตั้งแต่ -3.00 ถึง +3.00 โดยค่าตั้งแต่ -2.50 ถึง +2.50 ค่า b ที่อยู่ใกล้ -2.50 แสดงว่าเป็นข้อสอบที่ง่าย ส่วนค่า b ที่อยู่ใกล้ +2.50 เป็นข้อสอบที่ยาก

1.2.2 ค่าอำนาจจำแนก (a-parameter) ของข้อสอบ มีค่าอยู่ระหว่าง 0.307 ถึง 0.864 แสดงให้เห็นว่าข้อสอบมีค่าอยู่เกณฑ์คุณภาพที่ใช้ได้ ทั้งนี้อาจเนื่องมาจากข้อสอบได้ผ่านการวิเคราะห์และปรับปรุงคุณภาพมาแล้วครั้งหนึ่ง จึงทำให้ข้อสอบที่ได้มีค่าพารามิเตอร์อยู่ในเกณฑ์ที่ดี ซึ่ง ศิริชัย กาญจนวาสี (2555) กล่าวว่า ตามตามทฤษฎีการตอบสนองของข้อสอบ ในทางทฤษฎีมีค่าระหว่าง ($-\infty$ ถึง $+\infty$) ควรมีค่าเป็นบวก แต่ในทางปฏิบัตินิยมใช้ข้อสอบที่มีค่าอำนาจจำแนกอยู่ระหว่าง +0.5 ถึง +2.5 สอดคล้องกับแนวคิดของ Hambleton และ Cook (1977) Hambleton และ Swaminathan (1985) ค่าอำนาจจำแนก (a) มีพิสัยอยู่ระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติมีค่าตั้งแต่ 0 ถึง +2.00 เพราะค่า a ที่เป็นลบแสดงว่าข้อสอบไม่ดี ใช้ไม่ได้ต้องตัดทิ้ง และค่า a ที่เป็น 0 แสดงว่าข้อสอบไม่มีอำนาจจำแนก ค่า a เป็น +2.00 แสดงว่าข้อสอบมีค่าอำนาจจำแนกสูง ในการคัดเลือกข้อสอบควรตัดข้อที่มีค่า a ตั้งแต่ 0.30 ขึ้นไป

2. การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ (DIF) การรู้เรื่องวิทยาศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 45 ข้อ ด้วยวิธี MIMIC โดยใช้โปรแกรมประยุกต์ Mplus จำแนกตามเพศ พบว่า เพศชายมีโอกาสตอบถูกมากกว่าเพศหญิง จำนวน 18 ข้อ และเพศหญิงมีโอกาสตอบถูกมากกว่าเพศชาย จำนวน 7 ข้อ ซึ่งสอดคล้องกับแนวคิดของ ศิริชัย กาญจนวาสี (2555) กล่าวว่า การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เป็นการเปรียบเทียบผลการตอบข้อสอบเป็นรายข้อระหว่างกลุ่มผู้สอบอย่างน้อย 2 กลุ่ม มีความสามารถหลัก (Primary Ability) ที่มุ่งวัดเท่ากัน แต่คาดว่าจะมีความได้เปรียบเสียเปรียบกัน โดยกลุ่มหนึ่งถือเป็นกลุ่มอ้างอิง (Reference Group) ซึ่งคาดว่าน่าจะได้เปรียบในการตอบข้อสอบข้อนั้น หรือมีโอกาสตอบข้อสอบได้ถูกต้องมากกว่า ส่วนอีกกลุ่มคือ กลุ่มเปรียบเทียบ (Focal Group) ซึ่งเป็นกลุ่มที่สนใจศึกษาและคาดว่าน่าจะเป็นกลุ่มที่เสียเปรียบ ซึ่งสอดคล้องกับ พิชิตชัย สุขเอก (2560) กล่าวว่า ลักษณะข้อสอบที่เกิดจากการทำหน้าที่ต่างกัน โดยจำแนกตามเพศ เนื่องจากข้อสอบสามารถตอบได้หลายคำตอบตามความคิดเห็นของแต่ละบุคคล และข้อสอบที่เกิดการทำหน้าที่ต่างกันของข้อสอบอาจมีคำศัพท์ที่เกี่ยวข้องกับเพศชายหรือเพศหญิง จึงทำให้ข้อสอบเข้าข้างกลุ่มนั้น ๆ และทั้งนี้อาจเป็นเพราะความสามารถที่แตกต่างกันระหว่างเพศชายหรือเพศหญิงที่มีความสามารถ ความถนัดและความสนใจในเรื่องนั้น ๆ ต่างกัน สอดคล้องกับผลการวิจัยของ Le (2006 อ้างอิงมาจาก ประกฤติยา ทักษิโณ, 2552) ที่วิเคราะห์หาข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบโดยใช้ข้อมูลจาก PISA รอบทดลองใช้ ปี 2006 ในมิติที่ต่างกัน เมื่อวิเคราะห์ในมิติของสมรรถนะทางวิทยาศาสตร์ พบข้อสอบที่ทำหน้าที่ต่างกันที่โอนเอียงให้เพศชาย ร้อยละ 13.5 และโอนเอียงให้กับเพศหญิงร้อยละ 9.5 และเมื่อวิเคราะห์ในมิติของเป้าหมายของคำถามพบข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบที่โอนเอียงให้เพศชายร้อยละ 18.1 และโอนเอียงให้เพศหญิงร้อยละ 8.2 สอดคล้องกับผลการวิจัยของ Maier และ Casselmen (1970 อ้างอิงมาจาก สุพัฒนา หอมบุปผา, 2556) กล่าวว่า เพศชายจะมีความสามารถในการแก้ปัญหาทางคณิตศาสตร์และวิทยาศาสตร์เก่งกว่าเพศหญิง ส่วนเพศหญิงจะมีความสามารถในทางภาษาและสังคมศาสตร์เก่งกว่าเพศชาย และ PISA 2015 ยังพบความแตกต่างของคะแนนเฉลี่ยวิทยาศาสตร์ระหว่างนักเรียนชายและนักเรียนหญิง โดยเฉลี่ยในประเทศสมาชิก OECD นักเรียนชายมีคะแนนวิทยาศาสตร์สูงกว่านักเรียนหญิง 4 คะแนน แม้ความแตกต่างจะมีช่องว่างไม่มากนักแต่ก็มีนัยสำคัญทางสถิติ (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี กระทรวงศึกษาธิการ, 2561)

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1.1 หน่วยงานที่เกี่ยวข้องในการประเมินการรู้เรื่องวิทยาศาสตร์ของนักเรียนสามารถนำสารสนเทศที่เกิดจากการตรวจพบไปศึกษาลักษณะข้อสอบที่ทำหน้าที่ต่างกัน เพื่อเป็นแนวทางในการพัฒนาคุณภาพข้อสอบต่อไป

1.2 ผู้ที่เกี่ยวข้องในการออกข้อสอบ ควรนำผลการวิจัยนี้ไปใช้เป็นข้อมูลในการพิจารณาทบทวนข้อสอบ กำจัดความเป็นได้ของสาเหตุที่จะทำให้เกิดการทำหน้าที่ต่างกันของข้อสอบ เช่น ภาษาที่ใช้ในการเขียนข้อสอบ ลักษณะเนื้อหาที่นำมาสร้างข้อสอบ ลักษณะเฉพาะของกลุ่มผู้สอบ เป็นต้น

1.3 นักวิจัยและนักวัดผลทางการศึกษาที่มีความสนใจเกี่ยวกับการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ สามารถเลือกใช้ขนาดกลุ่มตัวอย่างที่มีความเหมาะสมที่จะทำให้มีประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบดีที่สุด กลุ่มตัวอย่างขนาดใหญ่จะทำให้สามารถตรวจพบข้อสอบที่ทำหน้าที่ต่างกันได้ดีกว่ากลุ่มตัวอย่างขนาดเล็ก

2. ข้อเสนอแนะในการศึกษาค้นคว้าต่อไป

2.1 วิธี MIMIC มีประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบที่มีการให้คะแนนแบบ 2 ค่า จึงควรมีการนำไปเปรียบเทียบเพิ่มเติมกับวิธีการตรวจสอบอื่น ๆ และตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบทดสอบที่มีการให้คะแนนมากกว่า 2 ค่า

2.2 ควรมีการศึกษาตัวแปรกลุ่มอื่น ๆ ในการเปรียบเทียบการทำหน้าที่ต่างกันของข้อสอบ เช่น ภูมิลำเนา ศาสนา เชื้อชาติ ภาษา อายุ เป็นต้น ว่าจะส่งผลให้ข้อสอบเกิดการทำหน้าที่ต่างกันหรือไม่

พหุนํ ปณฺฑิโต ชีเว

บรรณานุกรม

บรรณานุกรม

กระทรวงศึกษาธิการ. (2551). *ตัวชี้วัดและสาระการเรียนรู้แกนกลางกลุ่มสาระการเรียนรู้วิทยาศาสตร์ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551*. กรุงเทพฯ : ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย.

กระทรวงศึกษาธิการ. (2552). *หลักสูตรแกนกลางการศึกษาขั้นพื้นฐานพุทธศักราช 2551 สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ*. กรุงเทพฯ : ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย.

โกศล จิตวิรัตน์. (2556). การวิเคราะห์และนำเสนอผลการวิเคราะห์โมเดลสมการโครงสร้างพหุระดับด้วยโปรแกรม Mplus. *วารสารสมาคมนักวิจัย*, 18(2), 70-83, พฤษภาคม-สิงหาคม.

จุฬาพรรณ ชุมพล. (2555). *การพัฒนาแบบทดสอบวินิจฉัยทางการเรียนวิทยาศาสตร์ เรื่อง แรงและพลังงาน สำหรับชั้นมัธยมศึกษาปีที่ 1*. วิทยานิพนธ์ปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม.

โชติ เพชรชื่น. (2544). แบบทดสอบวินิจฉัย. *สารานุกรมศึกษาศาสตร์*, 23, 7-11.

โชติกา จันทะวัน. (2562). *การพัฒนาแบบทดสอบความสามารถในการให้เหตุผลทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5 ในจังหวัดกาฬสินธุ์*. วิทยานิพนธ์ปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลและประกันคุณภาพการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม.

ธเกียรติกมล ทองงอก, โชติกา ภาษีผล และศิริชัย กาญจนวาสี. (2556). ประสิทธิภาพการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในวิธีถดถอยโลจิสติก โดยใช้เกณฑ์ขนาดอิทธิพล 2 วิธี สำหรับข้อสอบที่มีรูปแบบการตรวจให้คะแนนแบบทวิภาค : ข้อมูลจำลองและข้อมูลเชิงประจักษ์. *วารสารวิจัย มสด. สาขามนุษยศาสตร์และสังคมศาสตร์*, 9(2), 31-50.

ธิดารัตน์ ธนะขว้าง. (2553). *การสร้างแบบวัดความคิดสร้างสรรค์ทางคณิตศาสตร์สำหรับนักเรียนชั้นประถมศึกษาปีที่ 4*. วิทยานิพนธ์ปริญญาศึกษาศาสตรมหาบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา บัณฑิตวิทยาลัย มหาวิทยาลัยเชียงใหม่.

นิธิวดี วีระศิลป์. (2561). *การสร้างแบบทดสอบความคิดสร้างสรรค์เชิงตัวเลขของนักเรียนระดับประกาศนียบัตรวิชาชีพชั้นปีที่ 1 วิทยาลัยพณิชยการอินทราชัย สังกัดสำนักงานคณะกรรมการการอาชีวศึกษา*. วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิจัยวัดผล และสถิติการศึกษา บัณฑิตวิทยาลัย มหาวิทยาลัยบูรพา.

บุญชม ศรีสะอาด และคณะ. (2553). *พื้นฐานการวิจัยการศึกษา*. มหาสารคาม : ภาควิชาการวิจัยและพัฒนาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม.

ประกฤติยา ทักษิโณ. (2552). *การประเมินคุณภาพการจัดการศึกษาวิชาวิทยาศาสตร์ ของสถานศึกษาขั้นพื้นฐาน : การประยุกต์ใช้การทำหน้าที่ต่างกันของข้อสอบและโมเดลมูลค่าเพิ่ม*. วิทยานิพนธ์ปริญญาครุศาสตรดุษฎีบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

พนัส จันทร์เปล่ง. (2557). *การวิเคราะห์เปรียบเทียบโมเดลประเมินคุณภาพการจัดการศึกษาวิทยาศาสตร์ : การประยุกต์ใช้โมเดลมูลค่าเพิ่มพหุระดับที่มีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและแบบสอบ*. วิทยานิพนธ์ปริญญาครุศาสตรดุษฎีบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา ครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

พิชิต ฤทธิ์จรูญ. (2550). *หลักการวัดและประเมินผลการศึกษา*. พิมพ์ครั้งที่ 4. กรุงเทพฯ : เฮ้าส์ออฟเคอร์มิสท์.

พิชิตชัย สุขเอก. (2560). *การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบโดยใช้วิธีมิมิค*. วิทยานิพนธ์ปริญญาศึกษาศาสตรมหาบัณฑิต สาขาวิชาการวิจัยและประเมินผลการศึกษา บัณฑิตวิทยาลัย มหาวิทยาลัยสงขลานครินทร์.

ไพศาล วรคา. (2558). *การวิจัยทางการศึกษา*. พิมพ์ครั้งที่ 7. มหาสารคาม : ตักสิลาการพิมพ์.

ภาวิณี ประทุมสาย. (2559). *การพัฒนาแบบทดสอบการรู้เรื่องวิทยาศาสตร์ โดยประยุกต์ใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ*. วิทยานิพนธ์ปริญญาการศึกษามหาบัณฑิต สาขาวิชาวิจัยและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม.

ยุพา สำเลิศรัมย์. (2558). การสร้างแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนรายวิชาวิทยาศาสตร์สำหรับชั้นมัธยมศึกษาปีที่ 1 ตามแนวข้อสอบในโครงการประเมินผลนักเรียนร่วมกับนานาชาติ (PISA). *วารสารมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยราชภัฏสุรินทร์*, 19(2), 21-34.

รักชนก ยี่สุ่นศรี. (2544). *การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและแบบสอบด้วยกระบวนการ ดี เอฟ ไอ ที สำหรับแบบสอบคัดเลือกบุคคลเข้าศึกษาในสถาบันอุดมศึกษาวิชาภาษาอังกฤษและวิชาคณิตศาสตร์*. วิทยานิพนธ์ปริญญาครุศาสตรมหาบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ล้วน สายยศ และอังคณา สายยศ. (2539). *เทคนิคการวัดผลการเรียนรู้*. กรุงเทพฯ : สุวีริยาสาส์น.

ล้วน สายยศ และอังคณา สายยศ. (2543). *การวัดด้านจิตพิสัย*. กรุงเทพฯ : สุวีริยาสาส์น.

ศิริชัย กาญจนวาสี. (2552). *ทฤษฎีการทดสอบแบบดั้งเดิม*. พิมพ์ครั้งที่ 6. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2555). *ทฤษฎีการทดสอบแนวใหม่*. พิมพ์ครั้งที่ 4. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย.

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2560 ก). *กรอบโครงสร้างการประเมินผลนักเรียนนานาชาติ PISA 2015*. กรุงเทพฯ : อรุณการพิมพ์.

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2560 ข). *สรุปผลการวิจัย PISA 2015*. กรุงเทพฯ : อรุณการพิมพ์.

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2561). *ผลการประเมิน PISA 2015 วิทยาศาสตร์ การอ่าน และคณิตศาสตร์ ความเป็นเลิศและความเท่าเทียมทางการศึกษา*. กรุงเทพฯ : ซัคเซสพับลิเคชั่น.

สมนึก ภัททิยธนี. (2546). *การวัดผลการศึกษา*. พิมพ์ครั้งที่ 4. กาฬสินธุ์ : ประสานการพิมพ์.

สมนึก ภัททิยธนี. (2549). *การวัดผลการศึกษา*. พิมพ์ครั้งที่ 5. กาฬสินธุ์ : ประสานการพิมพ์.

สมนึก ภัททิยธนี. (2553). *การวัดผลการศึกษา*. พิมพ์ครั้งที่ 7. กรุงเทพฯ : ประสานการพิมพ์.

สมนึก ภัททิยธนี. (2555). *การวัดผลการศึกษา*. พิมพ์ครั้งที่ 8. กาฬสินธุ์ : ประสานการพิมพ์.

สมบัติ ท้ายเรือคำ. (2551). *ระเบียบวิธีวิจัยสำหรับมนุษยศาสตร์และสังคมศาสตร์*. กาฬสินธุ์ : ประสานการพิมพ์.

สมบัติ ท้ายเรือคำ. (2553). *วิธีการทางสถิติสำหรับการวิจัย*. มหาสารคาม : ภาควิชาวิจัยและพัฒนาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม.

สุทธิพร ศุรธณ. (2550). *การศึกษาความสามารถในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบตามตัวแบบเชิงเส้นวางนัยทั่วไประดับลดหลั่น*. วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาคณิตศาสตร์ บัณฑิตวิทยาลัย มหาวิทยาลัยธรรมศาสตร์.

สุพัฒนา หอมบุปผา. (2556). การเปรียบเทียบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธี HGLM วิธี MIMIC และวิธี BAYESIAN. *วารสารวิจัยราชภัฏพระนคร*, 8(2), 10-24.

Ahman, S.J. and Glock, M.D. (1967). *Evaluating Pupil Growth Principle of Tests and Measurement*. 3rd ed. Boston : Allyn and Bacon.

Angoff, W.H. (1982). Use of difficulty and discrimination indices for detecting item bias. In Berk, R.A., editor, *Handbook of methods for detecting test bias* (pp. 96–116). Baltimore, MD : Johns Hopkins University Press,

Angoff, W.H. (1993). Perspective on differential item functioning methodology. In P.W. Holland H. Wainer (Eds.), *Differential item functioning* (pp. 3–24). Hillsdale, N.J. : Lawrence Erlbaum.

Angoff, W.H. and Ford, S.F. (1973). Item-race interaction on a test of scholastic aptitude. *Journal of Educational Measurement*, 10, 95-106.

Camilli, G. and Shepard, L.A. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA : Sage.

Carvajal, J. and Skorupski, W.P. (2010). The Effects of Small Sample Size on Identifying Polytomous DIF Using the Liu-Agresti Estimator of the Cumulative Common Odds Ratio. *Educational and Psychological Measurement*, 70(6), 914-925.

Cleary, T.A. and Hilton, T.L. (1968). An investigation of item bias. *Educational and Psychological Measurement,* 28(1), 61–75.

Cronbach, L.J. (1970). *Essentials of Psychological Testing*. 3rd ed. New York : Harper and Row.

Dorans, N.J. and Kulick, E. (1986). Demonstrating the unity of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement*, 23(4), 355-368.

Finch, H.W. and French, B.F. (2007). Detection of Crossing Differential Item Functioning : A Comparison of Four Methods. *Educational and Psychological Measurement*, 67(4), 565-582.

Fleishman, M. (2003). Economic grand rounds: Psychopharmacosocioeconomics and the global burden of disease. *Psychiatric Services*, 54, 1424144.

Gronlund, N.E. and Linn, R.L. (2009). *Measurement and Evaluation in Teaching*. New York : Macmillan.

Hambleton, R.K. and Cook, L.L. (1977). Talent Trait Model and Their use in The Analyze of Educational Test Data. *Journal of Educational Measurement*, 14, 75-79.

Hambleton, R.K. and others. (1991). *Fundamentals of item response theory.* Newbury Park, CA : Sage.

Hambleton, R.K. and others. (1993). Advances in the detection of differentially functioning test items. *European Journal of Psychological Assessment*, 9(1), 1-18.

Hambleton, R.K. and Swaminathan, H. (1985). *Item Response Theory : Principles and Application*. Boston : Kluwer–Nijhoff Publishing.

Hambleton, R.K., Swaminathan, H. and Rogers, H.J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, Calif : Sage.

Holland, P.W. and Thayer, D.T. (1986). Differential Item Performance and the Mantel Haenszel Pro-cedure. In H. Wainer & H.I. Braun (eds.), *Test Validity* (pp.129-145), Hillsdale, N.J. : Lawren Erlbaum Associate.

Holland, P.W. and Thayer, D.T. (1988). Differential item performance and thee Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 129-145). Hillsdale, NJ : Lawrence Erlbaum.

Holland, P.W. and Wainer, H. (1993). *Differential Item Functioning*. Hillsdale, NJ. : Lawrence Erlbaum Associates.

Jensen, A. (1974). Morphological studies in Cystoseiraceae and Sargasaceae (Phaeophyceae) with special reference to apical organization. *University of California Publications in Botany*, 68, 1-61.

Kim, S. and others. (2007). DIF Detection and Effect Size Measures for Polytomously Scored Items. *Journal of Educational Measurement*, 44(2), 93-116.

Lord, F.M. (1980). *Application of item Response Theory to practical Testing Problems*. New Jersey : Addision-Wesley.

Mellenbergh, G.J. (1982). Contingency table models for assessing item bias. *Journal of Educational Statistics*, 7, 105-118.

Narayanan, P. and Swaminathan, H. (1996). Identification of item that show nonuniform DIF. *Applied Psychological Measurement*, 20(3), 257-274.

Oort, F.J. (1992). Using restricted factor analysis to detect item bias. *Methodika, Methodika*, 6, 150-166.

Penfield, R.D. (2007). An approach for categorizing DIF in polytomous items. *Applied Measurement in Education*, 20(3), 335-355.

Penfield, R.D. (2010 a). How Are the Form and Magnitude of DIF Effects in MultipleChoice Items Determined by Distractor-Level Invariance Effects?. *Educational and Psychological Measurement*, 71(1), 54-67.

Penfield, R.D. (2010 b). Modeling DIF Effects Using Distractor-Level Invariance Effects : Implications for Understanding the Causes of DIF. *Applied Psychological Measurement*, 34(3), 151-165.

Penfield, R.D. and GÜler, N. (2009). A Comparison of the Logistic regression and Contigency Talbe Methods for Simutaneous Detection of Uniform and Nonuniform DIF. *Journal of Educational Measurement*, 46(3), 314-329.

Raju, N.S. (1990). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement*, 14, 197-207.

Scheuneman, J.D. (1975). *A New Method of Assessing Bias in Test Items.* Washington, D.C. : Paper presented at the Meeting of the American Educational Research Association.

Scheuneman, J.D. (1979). A method of assessing bias in test items. *Journal of Educational Measurement*, 16(3), 143–152.

Scheuneman, J.D. and Bleistein, C.A. (1989) A Consumer's Guide to Statistics for Identifying Differential Item Functioning. *Applied Measurement in Education*, 2(3), 255-275.

Shealy, R. and Stout, W.F. (1993). *Differential item functioning : Theory and practice*. New Jersey : Lawrence Erlbaum.

Swaminathan, H. and Rogers, H.J. (1990). Detecting Differential Item Functioning Using Logistic Regression Procedures. *Journal of Educational Measurement*, 27(4), 361-370.

Taylor, C.S. and Lee, Y. (2012). Gender DIF in Reading and Mathematics Tests With Mixed Item Formats. *Applied Measurement in Education*, 25(3), 246-280.

Wolfe, E.W. and Smith, E.V. (2007). Instrument tools and activities for measure validation using rash models: Part II–Validation activities. *Journal of Applied Measurement*, 8(2), 204-234.

Woods, C.M. and Grimm, K.J. (2011). Testing for nonuniform differential item functioning with multiple indicator multiple cause models. *Applied Psychological Measurement*, 35, 339-361.

Zieky, M. (1993). Practical questions in the use of DIF statistics in item development. In P.W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 337–347). Hillsdale, NJ : Lawrence Erlbaum Associates.

ภาคผนวก

ภาคผนวก ก

รายชื่อผู้เชี่ยวชาญในการวิเคราะห์ความสอดคล้อง (IOC) ระหว่างนิยามและข้อคำถาม

แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

พหูน ปณฺฑิโต ชีเว

รายชื่อผู้เชี่ยวชาญในการวิเคราะห์ความสอดคล้อง (IOC) ระหว่างนิยามและข้อคำถาม แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

1. ผู้ช่วยศาสตราจารย์ ดร.สมทรง สิทธิ

    อาจารย์ประจำภาควิชาหลักสูตรและการสอน

    คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

2. นางสาวรัตนา เพ็งเพราะ

    อาจารย์ประจำภาควิชาชีววิทยา คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์

    กศ.ม.(ชีววิทยา) มหาวิทยาลัยมหาสารคาม

3. นายวีรพล เลพล

    อาจารย์ประจำภาควิศวกรรมเครื่องกล คณะเทคโนโลยีอุตสาหกรรม

    มหาวิทยาลัยราชภัฏพิบูลสงคราม

    กศ.ม. (วิศวกรรมศาสตร์) มหาวิทยาลัยมหาสารคาม

4. ดร.ตวัญจลักษณ์ พวงนิล

    ครูชำนาญการพิเศษ โรงเรียนสารคามพิทยาคม อำเภอเมืองมหาสารคาม

    จังหวัดมหาสารคาม

    ปร.ด.(บริหารการศึกษา) มหาวิทยาลัยขอนแก่น

5. นายตันติกร ขุนาพรม

    ศึกษานิเทศก์ สพป.มหาสารคาม เขต 1

    กศ.ม.(วิจัยและประเมินผลการศึกษา) มหาวิทยาลัยมหาสารคาม

พหูนํ ปณฺฑิโต ชีเว

ภาคผนวก ข

ตัวอย่างหนังสือขอความอนุเคราะห์เป็นผู้เชี่ยวชาญตรวจสอบเครื่องมือที่ใช้ในงานวิจัย

ที่ อว 0605.5(2)/ว1604

คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม
อำเภอเมือง จังหวัดมหาสารคาม 44000

14 พฤษภาคม 2563

เรื่อง ขอความอนุเคราะห์เป็นผู้เชี่ยวชาญตรวจสอบเครื่องมือที่ใช้ในการวิจัย

เรียน ผู้ช่วยศาสตราจารย์ ดร.สมทรง สิทธิ

ด้วย นายศราวุทธ โพธิ์สิงห์ นิสิตปริญญาโท สาขาวิชาวิจัยและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม กำลังศึกษาและทำวิทยานิพนธ์ เรื่อง : "การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3" ซึ่งเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร (กศ.ม.) วิจัยและประเมินผลการศึกษา โดยมี ผู้ช่วยศาสตราจารย์ ดร.วราพร เอราวรรณ์ เป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก ในครั้งนี้

เพื่อให้การจัดทำวิทยานิพนธ์เป็นไปด้วยความเรียบร้อย และบรรลุตามวัตถุประสงค์ คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม พิจารณาแล้วเห็นว่าท่านเป็นผู้มีความรอบรู้ ความสามารถ และมีประสบการณ์ ในเรื่องนี้เป็นอย่างดี จึงใคร่ขอความอนุเคราะห์จากท่านเป็นผู้เชี่ยวชาญตรวจสอบเครื่องมือที่ใช้ในการวิจัย เพื่อนิสิตจะนำข้อมูลที่ได้ไปดำเนินการในขั้นตอนต่อไป

จึงเรียนมาเพื่อโปรดพิจารณา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม หวังเป็นอย่างยิ่งว่าคงได้รับความกรุณาจากท่านด้วย และขอขอบคุณมา ณ โอกาสนี้

ขอแสดงความนับถือ

(รองศาสตราจารย์ ดร.รังสรรค์ โฉมยา)
รองคณบดี ปฏิบัติราชการแทน
คณบดีคณะศึกษาศาสตร์

งานวิชาการและบัณฑิตศึกษา คณะศึกษาศาสตร์
โทรศัพท์,โทรสาร 0-4371-3174
เบอร์โทรนิสิต 0956702692

ภาคผนวก ค

ตัวอย่างข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

แบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3

คำชี้แจง

1. แบบทดสอบนี้ เป็นแบบทดสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3
2. แบบทดสอบนี้ เป็นแบบทดสอบแบบเลือกตอบ จำนวน 54 ข้อ
3. ใช้เวลาในการสอบ 1 ชั่วโมง 30 นาที
4. ให้นักเรียนเลือกคำตอบที่ถูกที่สุดเพียงคำตอบเดียว จากตัวเลือก ก ข ค และ ง โดยทำเครื่องหมาย X ลงในกระดาษคำตอบให้ตรงกับข้อที่เลือก ดังตัวอย่าง

ข้อ	ก	ข	ค	ง
(0)	X			

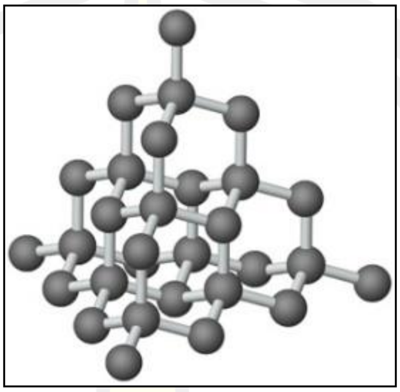
ถ้านักเรียนต้องการเปลี่ยนคำตอบใหม่ ให้ขีดทับเครื่องหมาย X เดิม แล้วจึงทำเครื่องหมาย X ลงในช่องตัวเลือกใหม่ ดังตัวอย่าง ถ้าต้องการเปลี่ยนคำตอบจาก ข เป็น ค ให้ทำดังนี้

ข้อ	ก	ข	ค	ง
(0)		~~X~~	X	

5. ก่อนลงมือทำข้อสอบให้เขียนชื่อ นามสกุล ลงในกระดาษคำตอบให้เรียบร้อย
6. ห้ามนักเรียนขีด เขียน หรือทำเครื่องหมายใดลงในแบบทดสอบนี้

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 1) จากภาพเป็นโครงสร้างของธาตุคาร์บอน (C) ชนิดหนึ่งที่มีอยู่ในธรรมชาติ จากการสันนิษฐานทางวิทยาศาสตร์ นักเรียนคิดว่าเป็นโครงสร้างของสารประกอบชนิดใด



ก. แกรไฟต์ (ไส้ดินสอ)

ข. ผงถ่าน

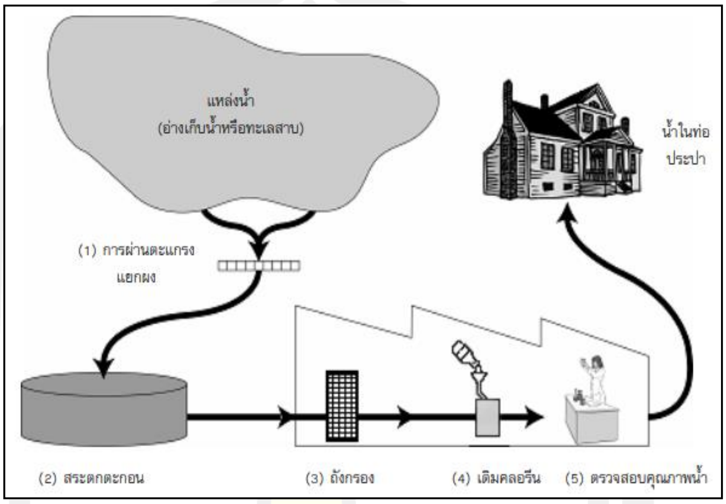
ค. เพชร

ง. พลอย

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	โครงสร้างของสสาร
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

รูปแสดงการผลิตน้ำใช้สำหรับบ้านที่อยู่ในเมืองให้สะอาดเพียงพอสำหรับบริโภค



ข้อ 2) การผลิตน้ำดื่มจำเป็นต้องมีแหล่งน้ำที่ดี น้ำที่พบอยู่ใต้ดินเรียกว่า น้ำใต้ดิน ทุกข้อเป็นเหตุผลที่ว่า แบคทีเรียและอนุภาคมลพิษในน้ำใต้ดินน้อยกว่าน้ำบนผิวดิน เช่น น้ำในทะเลสาบ และแม่น้ำ ยกเว้นข้อใด

ก. เมื่อน้ำซึมลงสู่พื้น น้ำจะถูกทำให้สะอาดโดยหินและทราย

ข. น้ำใต้ดินอยู่ในดิน มลพิษทางอากาศจึงไม่สามารถทำให้น้ำสกปรกได้

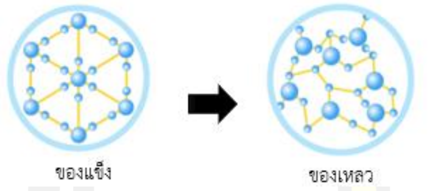
ค. น้ำใต้ดินผ่านที่กรองและเติมคลอรีน ผ่านที่กรองจนทำให้น้ำสะอาดมากที่สุด

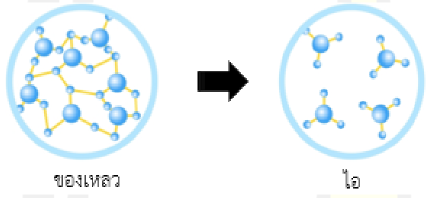
ง. น้ำในดินเป็นน้ำที่มีอาหารไม่มากพอสำหรับแบคทีเรีย แบคทีเรียจึงจึงมีชีวิตอยู่ในน้ำใต้ดินไม่ได้

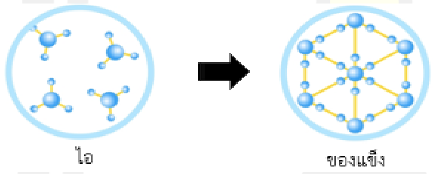
ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	โครงสร้างของสสาร
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	สุขภาพและโรคภัย
ระดับบริบท	ระดับท้องถิ่น/ชาติ

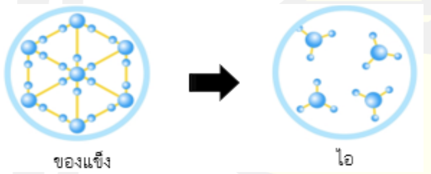
ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 3) ลูกเหม็นเป็นสสารชนิดหนึ่งอยู่ในรูปของสารประกอบ นักเรียนคิดว่ารูปแบบของการเปลี่ยนสถานะของลูกเหม็นเป็นไปตามข้อใด

ก. 

ข. 

ค. 

ง. 

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	โครงสร้างของสสาร
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	สุขภาพและโรคภัย
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 4) สมชายกำลังทำงานซ่อมแซมบ้านเก่าหลังหนึ่ง เขานำขวดน้ำ ตะปูเหล็ก และไม้ชิ้นหนึ่งไว้ที่ท้ายรถยนต์หลังจากที่รถออกไปอยู่กลางแดดเป็นเวลา 4 ชั่วโมง อุณหภูมิภายในรถสูงถึงประมาณ 39 °C เกิดอะไรขึ้นกับวัตถุในท้ายรถยนต์

ก. วัตถุทุกชนิดมีอุณหภูมิเท่ากัน

ข. หลังจากระยะเวลาหนึ่งน้ำเริ่มเดือด

ค. หลังจากระยะเวลาหนึ่งตะปูเหล็กเริ่มร้อนแดง

ง. อุณหภูมิของน้ำสูงกว่าอุณหภูมิของตะปูเหล็กและไม้

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	สมบัติของสสาร
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

น้ำแข็งขั้วโลก

รายงานการวิเคราะห์ของนักวิทยาศาสตร์ ระบุว่าน้ำจะท่วมโลกเหตุจากน้ำแข็งขั้วโลกละลายเร็วขึ้น โดยในปัจจุบันน้ำแข็งกรีนแลนด์ละลายเร็วมากขึ้นกว่าปี 1990 ถึง 7 เท่า สูญเสียน้ำแข็งไปประมาณ 250 พันล้านตัน ซึ่งถ้าหากย้อนดูไปในช่วงปี 90 กรีนแลนด์สูญเสียน้ำแข็งเพียงแค่ 33 พันล้านตัน ส่วนในปีนี้นักวิทยาศาสตร์คาดว่าอาจสูงถึง 370 พันล้านตัน

นอกจากนี้ นักวิทยาศาสตร์ขั้วโลก ยังชี้ว่า จากการรวบรวมข้อมูลจากดาวเทียมในช่วง 26 ปี ที่ผ่านมาเหตุการณ์ดังกล่าวจะส่งผลกระทบต่อประชาชนในประเทศที่อยู่ในพื้นที่ชายฝั่งทะเล เนื่องจากระดับน้ำทะเลที่สูงขึ้นกว่าที่เป็นอยู่ โดยคาดว่า ระดับน้ำทะเลจะสูงขึ้นอีก 7 เซนติเมตร ในช่วงปี 2100 สาเหตุหลักมาจากมนุษย์ ที่เกิดจากมลพิษจากการใช้และการเผาเชื้อเพลิง มลพิษจากการผลิตพลังงานจากฟอสซิล ซึ่งบทบาทหลักของน้ำแข็งขั้วโลกคือ สะท้อนความร้อนมหาศาลจากดวงอาทิตย์ รักษาสมดุลทางภูมิอากาศสำหรับเกษตรกรรม รักษาระดับน้ำทะเล และเป็นถิ่นอาศัยของสัตว์ต่าง ๆ

อย่างไรก็ตามในปี 2018 ทีมนักวิทยาศาสตร์ คำนวณพบว่าขั้วโลกใต้ Antarctica ในปี 2100 จะต่ำกว่าที่เคยคาดการณ์ไว้ 10 เซนติเมตร และในส่วนของกรีนแลนด์ กล่าวว่าความหนาของน้ำแข็งจะต่ำกว่าที่คาดแล้ว 7 เซนติเมตร

ข้อ 5) จากการละลายของน้ำแข็งขั้วโลกข้อใดต่อไปนี้ไม่ได้รับผลกระทบจากการละลายของน้ำแข็งขั้วโลก

ก. สัตว์หลายชนิดไร้ที่อยู่อาศัย

ข. เกิดภัยพิบัติเพิ่มขึ้นเพราะโลกเสียสมดุล

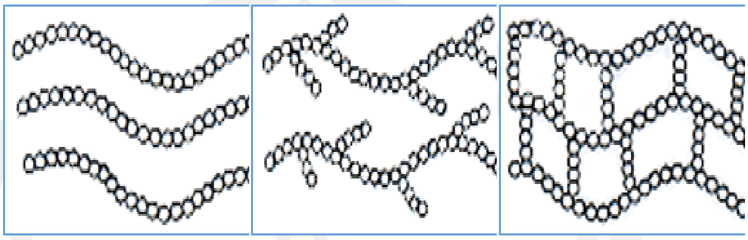
ค. มีแหล่งพลังงานทดแทนเพิ่มมากขึ้น

ง. แหล่งอาหารถดถอยเพราะอากาศแปรปรวน

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	สมบัติของสสาร
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	อันตราย
ระดับบริบท	ระดับโลก

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 6) พิจารณาโครงสร้างพอลิเมอร์ 4 ชนิด โครงสร้างกับสมบัติของพอลิเมอร์ต่อไปนี้ถูกต้อง



A B C

ก. ความยืดหยุ่นของพอลิเมอร์ชนิด B > A > C

ข. จุดเดือดของพอลิเมอร์ชนิด A > B > C

ค. พอลิเมอร์ที่นำมารีไซเคิลได้คือ ชนิด A และ B เท่านั้น

ง. พอลิเมอร์ที่มีความแข็งมากที่สุดคือ ชนิด A

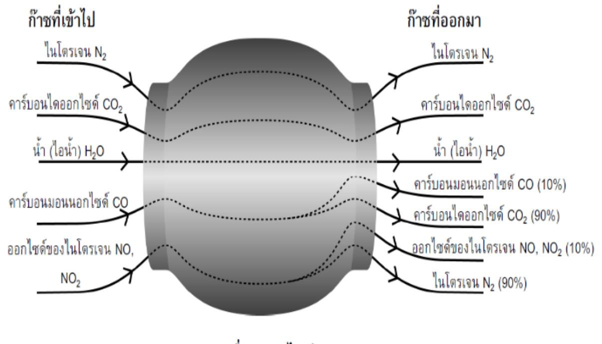
ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	สมบัติของสสาร
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับท้องถิ่น/ชาติ

พหุน ปณ ฑิโต ชีเว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

เครื่องแปลงไอเสีย

รถยนต์รุ่นใหม่ส่วนใหญ่ได้ติดตั้งเครื่องแปลงไอเสีย ซึ่งทำให้ไอเสียของรถยนต์เป็นอันตรายต่อคนและสิ่งแวดล้อมน้อยลงประมาณ 90% ของก๊าซอันตรายถูกแปลงเป็นก๊าซที่เป็นอันตรายน้อยลง ต่อไปนี้คือก๊าซบางชนิดที่เข้าไปในเครื่องแปลงและออกมาจากเครื่อง ดังรูป



เครื่องแปลงไอเสีย

ข้อ 7) จากแผนผังข้างบน การเปลี่ยนแปลงคาร์บอนมอนนอกไซด์หรือไนโตรเจนออกไซด์ไปเป็นสารประกอบอื่นๆ ข้อใดกล่าวไม่ถูกต้อง

ก. ไนโตรเจนออกไซด์ถูกเปลี่ยนเป็นไนโตรเจน

ข. คาร์บอนมอนนอกไซด์ถูกเปลี่ยนเป็นคาร์บอนไดออกไซด์

ค. คาร์บอนมอนนอกไซด์ถูกเปลี่ยนไปเป็นคาร์บอนไดออกไซด์และเขม่า

ง. คาร์บอนมอนนอกไซด์และไนโตรเจนออกไซด์ที่เป็นพิษถูกเปลี่ยนเป็นคาร์บอนไดออกไซด์และไนโตรเจนที่เป็นพิษน้อยลง

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	การเปลี่ยนแปลงทางเคมี
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	คุณภาพสิ่งแวดล้อม
ระดับบริบท	ระดับโลก

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ลิปมัน	ลิปสติก
ส่วนผสม : น้ำมันละหุ่ง 5.0 กรัม ไขผึ้ง 0.2 กรัม ไขมันปาล์ม 0.2 กรัม สีผสมอาหาร 1.0 ช้อนชา สารแต่งรสชาติ 1.0 หยด **วิธีทำ :** อุ่นน้ำมันและไขในอ่างน้ำจนผสมกันดี จากนั้นเติมสีผสมอาหารและสารแต่งรสชาติ แล้วผสมให้เข้ากัน	**ส่วนผสม :** น้ำมันละหุ่ง 5.0 กรัม ไขผึ้ง 1.0 กรัม ไขมันปาล์ม 1.0 กรัม สีผสมอาหาร 1.0 ช้อนชา สารแต่งรสชาติ 1.0 หยด **วิธีทำ :** อุ่นน้ำมันและไขในอ่างน้ำจนผสมกันดี จากนั้นเติมสีผสมอาหารและสารแต่งรสชาติ แล้วผสมให้เข้ากัน

ข้อ 8) ในการทำลิปมันและลิปสติก น้ำมันและไขถูกผสมเข้าด้วยกัน แล้วเติมสีผสมอาหารและสารแต่งรสชาติลิปสติกที่ทำจากส่วนผสมนี้จะแข็งและใช้ยาก นักเรียนจะเปลี่ยนสัดส่วนของส่วนผสมอย่างไร เพื่อทำให้ลิปสติกอ่อนลงกว่าเดิม

ก. เติมสีผสมอาหารมากขึ้น

ข. อุ่นของผสมให้นานขึ้นเพื่อให้อ่อนลง

ค. ควรใช้ไขผึ้งและไขมันปาล์มลดลงเล็กน้อย

ง. ควรใช้ไขผึ้งและไขมันปาล์มเพิ่มขึ้นเล็กน้อย

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	การเปลี่ยนแปลงทางเคมี
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	สุขภาพและโรคภัย
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

หลุมยุบ

หลุมยุบ เป็นปรากฏการณ์ธรรมชาติอย่างหนึ่งที่ดินยุบตัวลงเป็นหลุมลึกและมีเส้นผ่านศูนย์กลางตั้งแต่ 1-200 เมตร ลึกตั้งแต่ 1 ถึงมากกว่า 20 เมตร เมื่อแรกเกิดปากหลุมมีลักษณะเกือบกลมและมีน้ำขังอยู่ก้นหลุม ภายหลังน้ำจะกัดเซาะดินก้นหลุมกว้างขึ้น ลักษณะคล้ายลูกน้ำเต้า ทำให้ปากหลุมพังลงมาจนเหมือนกับว่าขนาดของหลุมยุบกว้างขึ้น โดยปกติหลุมยุบจะเกิดในบริเวณที่ราบใกล้กับภูเขาที่เป็นหินปูน เนื่องจากหินปูนมีคุณสมบัติละลายน้ำ ที่มีสภาพเป็นกรดอ่อนได้ประกอบกับภูเขาหินปูนมีรอยเลื่อนและรอยแตกมากมาย ดังจะสังเกตเห็นได้ว่าภูเขาหินปูนมีหน้าผาชัน หน้าผาเป็นรอยเลื่อนและรอยแตกในหินปูนนั่นเอง บริเวณใดที่รอยแตกของหินปูนตัดกันจะเป็นบริเวณที่ทำให้เกิดโพรงได้ง่าย โพรงหินปูนถ้าอยู่พ้นผิวดินก็คือถ้ำ ถ้าไม่โผล่เรียกว่าโพรงหินปูนใต้ดิน ซึ่งจำแนกเป็น 2 ระดับ คือ โพรงหินปูนใต้ดินระดับลึก (ลึกจากผิวดินมากกว่า 50 เมตร) และโพรงหินปูนใต้ดินระดับตื้น (ลึกจากผิวดินไม่เกิน 50 เมตร) ส่วนใหญ่หลุมยุบจะเกิดในบริเวณที่มีโพรงหินปูนใต้ดินระดับตื้น

ข้อ 9) หลุมยุบเกิดจากหินปูนถูกกัดกร่อนโดยน้ำที่มีสภาพเป็นกรดอ่อน หากน้ำฝนโดยทั่วไปมีค่า pH อยู่ระหว่าง 5-6 สถานที่ใดจะได้รับความเสียหายจากน้ำฝนมากที่สุด

ก.

ข.

ค.

ง.

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	การเปลี่ยนแปลงทางเคมี
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	อันตราย
ระดับบริบท	ระดับโลก

ข้อสอบการรู้เรื่องวิทยาศาสตร์

แรงเสียดทาน

แรงเสียดทานคือแรงที่ต้านการเคลื่อนที่ของวัตถุซึ่งเกิดขึ้นระหว่างผิวสัมผัสของวัตถุเกิดขึ้นทั้งวัตถุที่เคลื่อนที่และไม่เคลื่อนที่และจะมีทิศทางตรงกันข้ามกับการเคลื่อนที่ของวัตถุแรงเสียดทานจะมากหรือน้อยขึ้นอยู่กับแรงกระทำและพื้นผิวสัมผัสของวัตถุ

ข้อ 10) เด็กหญิงแสนดีเขียนหนังสือด้วยดินสอแท่งหนึ่งและรู้สึกเมื่อยมือจากการออกแรงเขียนหนังสือมาก ข้อใดไม่ใช่วิธีที่จะช่วยลดแรงเสียดทานจากการเขียนหนังสือของเด็กหญิงแสนดี

ก. ลดแรงกดกระดาษในการเขียน

ข. ใช้กระดาษที่มีผิวเรียบๆ

ค. เหลาดินสอให้แหลม

ง. ใช้ที่ช่วยจับดินสอ

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	การเคลื่อนที่และแรง
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	คุณภาพสิ่งแวดล้อม
ระดับบริบท	ระดับส่วนตัว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ภาพแสดงลักษณะของถนนที่ปูพื้นให้มีผิวหยาบ

ข้อ 11) ข้อใดคือตัวอย่างการใช้แรงที่เกี่ยวข้องกับการปูพื้นถนนไว้อย่างหยาบๆ มากที่สุด

ก. ใช้เข็มทิศเพื่อตรวจสอบทิศทาง

ข. ติดสปริงใต้อานจักรยาน

ค. ติดยางที่พื้นถุงเท้าเด็กเล็กเพื่อให้พื้นถุงเท้าขรุขระ

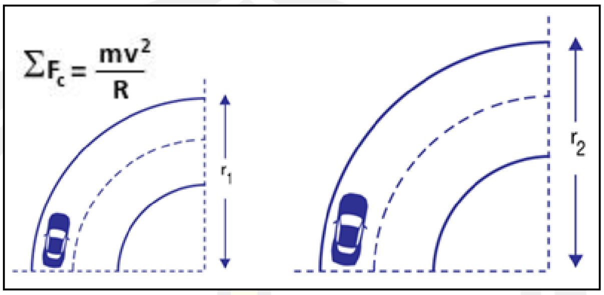
ง. บริเวณก่อสร้างมีการปล่อยลูกเหล็กลงมาแล้วตอกเสาเข็ม

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	การเคลื่อนที่และแรง
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	อันตราย
ระดับบริบท	ระดับท้องถิ่น/ชาติ

พหุนํ ปญฺ ฑิโต ชีเว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

เมื่อรถวิ่งเข้าโค้งจะมีรูปแบบการเคลื่อนที่แบบวงกลม ตามความสัมพันธ์



ข้อ 12) จากรูปพิจารณาข้อความต่อไปนี้ ข้อใดกล่าวถูกต้อง

ก. แรงเสียดทานระหว่างล้อกับถนน มีทิศทางตั้งฉากกับทิศทางการเคลื่อนที่

ข. หากเลี้ยวโค้งด้วยอัตราเร็วที่เท่ากัน โค้งรัศมี r_1 จะมีความปลอดภัยมากกว่าโค้งรัศมี r_2

ค. ความเร็วของรถ มีทิศขนานกันกับรัศมีของโค้ง

ง. แรงเสียดทานระหว่างล้อกับถนน มีทิศทางขนานกันกับทิศทางการเคลื่อนที่

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	การเคลื่อนที่และแรง
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับส่วนตัว

พหุนํ ปณฺฑิโต ชีเว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

การเผาไหม้

การเกิดปฏิกิริยาการเผาไหม้ของเชื้อเพลิงซึ่งมีธาตุคาร์บอนกับไฮโดรเจนเป็นองค์ประกอบส่วนใหญ่เมื่อเผาไหม้จะทำปฏิกิริยากับออกซิเจนในอากาศได้ 2 แบบดังนี้

1) เมื่อเกิดการเผาไหม้อย่างสมบูรณ์จะได้แก๊สคาร์บอนไดออกไซด์กับน้ำ

$$C_3H_8(g) + 5O_2(g) \longrightarrow 3CO_2(g) + 4H_2O\ (l)$$

(แก๊สโพรเพน) (แก๊สออกซิเจน) (แก๊สคาร์บอนไดออกไซด์) (น้ำ)

2) เมื่อเกิดการเผาไหม้ไม่สมบูรณ์จะได้เขม่า(ผงคาร์บอน) แก๊สคาร์บอนมอนอกไซด์และน้ำ

$$C_3H_8(g) + 3O_2(g) \longrightarrow 2CO_2(g) + C + 4H_2O\ (l)$$

(แก๊สโพรเพน) (แก๊สออกซิเจน) (แก๊สคาร์บอนไดออกไซด์) (คาร์บอน) (น้ำ)

การเผาไหม้ที่ไม่สมบูรณ์จะทำให้เกิดเขม่าสีดำสกปรกและแก๊สคาร์บอนไดออกไซด์ คาร์บอนมอนอกไซด์เป็นพิษต่อการหายใจถ้าสูดดมแก๊สชนิดนี้เพียงเล็กน้อยจะรู้สึกเวียนศีรษะถ้าสูดดมมากอาจทำให้ปวดศีรษะได้และถ้าสูดดมมากเกินไปจะทำให้สลบหรือตายได้

ข้อ 13) คำกล่าวอ้างข้อใดต่อไปนี้เป็นข้อดีของการใช้ภาชนะสแตนเลสและทรายก่อไฟเพื่อใช้ในการอยู่ไฟ

ก. สแตนเลสมีคุณสมบัติสามารถเก็บกักความร้อนได้ดี

ข. ทรายเป็นวัสดุที่ไม่ติดไฟ

ค. ทรายเป็นตัวพาความร้อนที่ดี

ง. สแตนเลสมีราคาถูก

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	พลังงานและการเปลี่ยนรูปพลังงาน
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	สุขภาพและโรคภัย
ระดับบริบท	ระดับส่วนตัว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

โยเกิร์ต

ในการทำโยเกิร์ตจะต้องนำนมสดไปอุ่นให้ร้อนแต่ไม่ถึงกับเดือดจากนั้นนำมาผสมกับหัวเชื้อ นั่นก็คือโยเกิร์ตที่เรารับประทานหลังจากผสมแล้วปิดฝาตั้งทิ้งไว้ประมาณ 4 -10 ชั่วโมง (ขึ้นกับอุณหภูมิห้อง) เพื่อให้เกิดการเปลี่ยนแปลงทางเคมี โดยการเจริญของจุลินทรีย์ร่วมกัน 2 ชนิดย่อยโปรตีนและน้ำตาลให้กลายเป็นกรด

ข้อ 14) การทำนมสดให้เป็นโยเกิร์ตข้อใดไม่ใช่เหตุที่ทำให้โยเกิร์ตมีรสเปรี้ยว

ก. โยเกิร์ตมีรสเปรี้ยว เพราะจุลินทรีย์เจริญเติบโต

ข. โยเกิร์ตมีรสเปรี้ยว เพราะโปรตีนถูกเปลี่ยนให้เป็นกรด

ค. โยเกิร์ตมีรสเปรี้ยว เพราะน้ำตาลถูกย่อยให้กลายเป็นกรด

ง. โยเกิร์ตมีรสเปรี้ยว เพราะไขมันเปลี่ยนสภาพกลายเป็นน้ำตาล

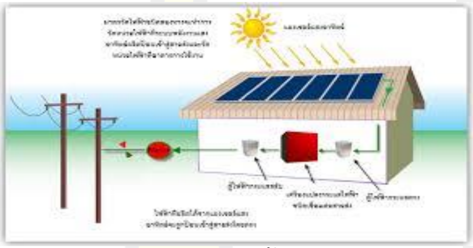
ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	พลังงานและการเปลี่ยนรูปพลังงาน
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	สุขภาพและโรคภัย
ระดับบริบท	ระดับส่วนตัว

พหูน ปณ ฑิโต ชีเว

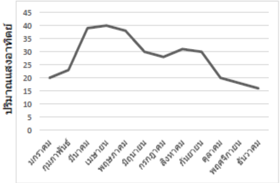
ข้อสอบการรู้เรื่องวิทยาศาสตร์

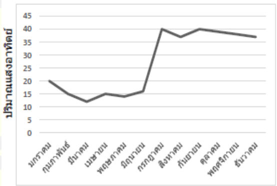
พลังงานแสงอาทิตย์

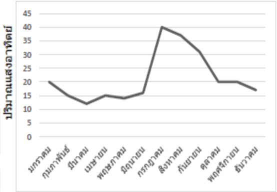
พลังงานแสงอาทิตย์ถูกใช้งานอย่างมากแล้วในหลายส่วนของโลก และมีศักยภาพในการผลิตพลังงานมากกว่าการบริโภคพลังงานของโลกในปัจจุบันหลายเท่าหากใช้ประโยชน์อย่างเหมาะสม พลังงานแสงอาทิตย์สามารถใช้โดยตรงเพื่อผลิตไฟฟ้าหรือสำหรับทำความร้อน หรือแม้แต่ทำความเย็น ศักยภาพในอนาคตของพลังงานแสงอาทิตย์นั้นถูกจำกัดโดยแค่เพียงความเต็มใจของเราที่จะคว้าโอกาสนั้นไว้มีวิธีการมากมายที่สามารถนำพลังงานจากแสงอาทิตย์มาใช้งานได้ พืชเปลี่ยนแสงอาทิตย์เป็นพลังงานทางเคมีโดยใช้การสังเคราะห์แสง เราใช้ประโยชน์จากพลังงานนี้โดยการกินพืชและเผาฟืน อย่างไรก็ตามคำว่า “พลังงานแสงอาทิตย์” หมายถึงการเปลี่ยนแสงอาทิตย์โดยตรงมากกว่าเปลี่ยนไปเป็นพลังงานความร้อนหรือพลังงานไฟฟ้าสำหรับใช้งาน ประเภทพื้นฐานของพลังงานแสงอาทิตย์ คือ “พลังความร้อนแสงอาทิตย์”และ“เซลล์แสงอาทิตย์”

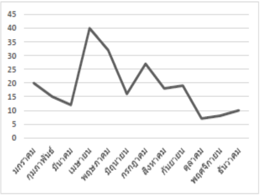


ข้อ 15) กราฟข้างล่างนี้ แสดงปริมาณแสงเฉลี่ยตลอดทั้งปีในสี่บริเวณที่แตกต่างกัน กราฟใดที่ชี้บอกบริเวณที่เหมาะสมในการตั้งเครื่องผลิตกระแสไฟฟ้าจากพลังแสงอาทิตย์

ก.


ข.


ค.


ง.


ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	พลังงานและการเปลี่ยนรูปพลังงาน
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

แผ่นดินไหว

ข้อมูลแสดงระดับและลักษณะความรุนแรงที่เกิดขึ้นจากแผ่นดินไหวตามมาตราเมอร์คัลลีที่ระดับความรุนแรง III ถึง VII เป็นดังนี้

ระดับความรุนแรง	ลักษณะความรุนแรง
III	คนที่อยู่บนอาคารรู้สึกได้ แต่คนส่วนใหญ่ยังไม่รู้สึกว่ามีแผ่นดินไหว
IV	ในเวลากลางวัน คนที่อยู่ในอาคารรู้สึกได้มาก แต่คนที่อยู่นอกอาคารรู้สึกบางคน จานหน้าต่าง ประตูสั่น ความรู้สึกเหมือนรถบรรทุกชนอาคาร
V	รู้สึกได้เกือบทุกคน หลายคนตกใจตื่น วัตถุที่ไม่มั่นคงล้มคว่ำ เสา ต้นไม้ แกว่งไกว
VI	ทุกคนรู้สึกได้ เครื่องเรือนเคลื่อน ปล่องไฟแตก เกิดความเสียหายเล็กน้อยกับอาคาร
VII	ทุกคนตกใจวิ่งออกนอกอาคาร อาคารที่ออกแบบไว้ดีไม่เสียหาย อาคารมาตรฐาน ปานกลาง เสียหายเล็กน้อยถึงปานกลาง อาคารที่ออกแบบไม่เสียหายมาก คนที่ขับรถรู้สึกมีแผ่นดินไหว

รายงานเหตุการณ์แผ่นดินไหวในเขตมินดาเนา ซึ่งอยู่ทางตอนใต้ของประเทศฟิลิปปินส์ “เมื่อวันที่ 29 ตุลาคม 2562 เกิดแผ่นดินไหวที่ความลึกประมาณ 7 กิโลเมตร จากระดับผิวโลก ส่งผลให้นักท่องเที่ยว ที่กำลังเล่นสกีหิมะทุกคนที่อยู่บริเวณยอดเขา รู้สึกได้ถึงการสั่นสะเทือน เมื่อมองเข้าไปในโรงอาหารของลานสกี พบว่าโคมไฟที่ห้อยอยู่ เกิดการแกว่งไปมา พบว่าตัวอาคารและสิ่งปลูกสร้างรอบ ๆ เกิดรอยร้าว แต่ไม่ถึงกับถล่มลงมา”

ข้อ 16) จากข้อมูล ฟิลิปปินส์ได้รับผลกระทบจากแผ่นดินไหวรุนแรงในระดับใดตามมาตราเมอร์คัลลี

ก. ระดับความรุนแรง ระดับ III

ข. ระดับความรุนแรง ระดับ IV

ค. ระดับความรุนแรง ระดับ V

ง. ระดับความรุนแรง ระดับ VI

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	การปฏิสัมพันธ์ระหว่างพลังงานและสสาร
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

การเดินทางของเสียงผ่านตัวกลาง

เสียงเกิดจากการสั่นสะเทือนของวัตถุ การที่เราสามารถได้ยินเสียงแสดงว่าเสียงเดินทางจากแหล่งกำเนิดเสียงมาสู่หูของเราได้ นักเรียนสองคนชื่อ หวั่นไหว และ ใจอ่อน ทำการทดลองเกี่ยวกับการเดินทางของเสียง ด้วยอุปกรณ์ดังนี้ กระป๋องนม 2 ใบ เชือกว่าว เส้นเอ็นพลาสติก เส้นลวดเหล็กขนาดเล็กที่สุด ลวดเสียบกระดาษ ตะปูเข็ม ค้อน

ขั้นตอนการทดลอง ใช้ค้อนตอกตะปูเพื่อเจาะรูตรงกลางก้นกระป๋องนมทั้ง 2 ใบ ตัดเชือกว่าวออกมายาวประมาณ 6 เมตร สอดปลายเชือกด้านหนึ่งเข้าไปในกระป๋องใบแรก ผูกลวดเสียบกระดาษไว้ข้างในดึงเชือกออกมาเพื่อให้ลวดเสียบกระดาษติดที่ก้นกระป๋อง ทำเช่นเดียวกันกับกระป๋องใบที่สองที่ปลายอีกข้างของเชือก ดังภาพที่ 1 แยกกระป๋องทั้งสองใบออกจากกัน เพื่อให้เชือกระหว่างกระป๋องทั้งสองตึง หวั่นไหว และ ใจอ่อน ผลัดกันเป็นคนกระซิบผ่านกระป๋องใบหนึ่ง ขณะที่อีกคนคอยฟังผ่านกระป๋องอีกใบหนึ่ง ดังภาพที่ 2 ทำการทดลองซ้ำแบบเดิมอีกโดยเปลี่ยนสายที่ขึงระหว่างแก้วจากเชือกว่าว เป็นเส้นเอ็นพลาสติก และเส้นลวดเหล็กขนาดเล็ก ตามลำดับ ผลัดกันเป็นคนกระซิบผ่านแก้วใบหนึ่ง ขณะที่อีกคนคอยฟังผ่านแก้วอีกใบหนึ่ง

ภาพที่ 1

ภาพที่ 2

บันทึกเปรียบเทียบลักษณะความชัดเจนของเสียงที่ได้ยินทั้ง 3 ครั้ง ได้ดังนี้

วัสดุที่ใช้ร้อยระหว่างแก้วทั้งสอง	ระดับความชัดเจนของเสียงที่ได้ยิน
1. เชือกว่าว	ความชัดเจนของเสียงน้อยที่สุด
2. เส้นเอ็นพลาสติก	ความชัดเจนของเสียงปานกลาง
3. เส้นลวดเหล็กขนาดเล็ก	ความชัดเจนของเสียงมากที่สุด

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 17) จากผลการทดลองจากการใช้เชือกว่าว เส้นเอ็นพลาสติก และเส้นลวดเหล็กขนาดเล็ก ขึงตึงระหว่างกระป๋องนมที่ใช้แล้วทั้งสอง ข้อสรุปใดไม่ถูกต้อง

ก. เชือกว่าว เส้นเอ็นพลาสติก และเส้นลวดเหล็กขนาดเล็ก มีสถานะไม่เหมือนกัน

ข. เชือกว่าว เส้นเอ็นพลาสติก และเส้นลวดเหล็กขนาดเล็ก มีความหนาแน่นไม่เท่ากัน

ค. เส้นลวดเหล็กขนาดเล็กมีความหนาแน่นมากที่สุด จึงส่งผ่านพลังงานเสียงได้ดีที่สุด

ง. โมเลกุลของตัวกลางจะเคลื่อนที่ไปพร้อมกับพลังงานเสียงถึงหูผู้ฟัง

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	การปฏิสัมพันธ์ระหว่างพลังงานและสสาร
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับส่วนตัว

พหุนํ ปณฺ ฑิโต ชีเว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 18) หากนำไข่เป็ดไปทำให้สุกโดยใช้ไมโครเวฟ ไข่จะระเบิดดังภาพ นักเรียนคิดว่าไข่ที่ระเบิดน่าจะเกิดจากสาเหตุใดที่สมเหตุสมผลมากที่สุด

ก. อุณหภูมิสูงเกินไป

ข. แรงดันเพิ่มขึ้นไข่เป็ดจึงระเบิด

ค. เมื่ออุณหภูมิสูงขึ้นทำให้เกิดการขยายตัวของไอน้ำในฟองไข่จนเกิดระเบิดขึ้น

ง. ไข่มีน้ำเป็นส่วนประกอบ เมื่อได้รับคลื่นจากไมโครเวฟไข่จะมีอุณหภูมิสูงขึ้นทำให้ความดันภายในไข่สูงขึ้นจนในที่สุดไข่ระเบิด

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบทางกายภาพ
ความรู้เกี่ยวกับ	การปฏิสัมพันธ์ระหว่างพลังงานและสสาร
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	อันตราย
ระดับบริบท	ระดับส่วนตัว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

“สิ่งมีชีวิตทั้งหลายประกอบด้วยเซลล์ และเซลล์ คือ หน่วยพื้นฐานของสิ่งมีชีวิตทุกชนิด”

ข้อ 19) จากบทความข้างต้น ใครเป็นผู้คิดค้นทฤษฎีเซลล์

ก. ชาลล์ ดาร์วิน และ ชไลเดน

ข. เทโอดอร์ ชวันน์ และ ชไลเดน

ค. รอเบิร์ต ฮุก และ รอเบิร์ต บราวน์

ง. หลุยส์ ปาสเตอร์ และ ชาลล์ ดาร์วิน

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	เซลล์
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับส่วนตัว

พหุน ปณ ฑิโต ชีเว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

การทดลองเพื่อศึกษาการแพร่ของด่างทับทิม 3 ชุดทดลอง ได้ผลการสังเกตดังนี้

ชุดทดลอง	ตัวกลาง	ผลการสังเกต
1.	A	สีม่วงของด่างทับทิมกระจายไปทุกทิศทุกทางอย่างรวดเร็ว
2.	B	สีม่วงของด่างทับทิมค่อยๆ กระจายไปทุกทิศทุกทางอย่างช้าๆ
3.	ไม่มีตัวกลาง	อนุภาคของด่างทับทิมไม่มีการเคลื่อนที่

ข้อ 20) ตัวแปรต้นของการทดลองนี้คืออะไร

ก. ปริมาณของด่างทับทิม

ข. ขนาดอนุภาคของด่างทับทิม

ค. ชนิดของตัวกลาง

ง. อัตราเร็วในการแพร่ของด่างทับทิม

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	เซลล์
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับท้องถิ่น/ชาติ

พหุน ปณ ฑิโต ชีเว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

นำเซลล์ชนิดเดียวกัน ขนาดเท่ากัน ไปใส่ในสารละลาย A B และ C ที่มีความเข้มข้นแตกต่างกันเป็นเวลานานเท่าๆ กัน ได้ผลดังภาพ

ข้อ 21) จงเรียงลำดับความเข้มข้นของสารละลาย A B และ C จากความเข้มข้นน้อยที่สุดไปมากที่สุด

ก. A B C

ข. A C B

ค. B A C

ง. C A B

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	เซลล์
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 22) ข้อใดคือเกณฑ์การจำแนกสิ่งมีชีวิตเซลล์เดียวกับสิ่งมีชีวิตหลายเซลล์

ก. จำนวนของเซลล์

ข. รูปร่างของเซลล์

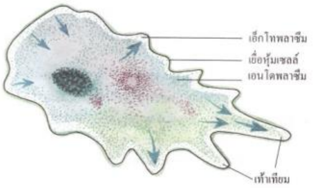
ค. ขนาดของเซลล์

ง. ส่วนประกอบของเซลล์

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	แนวความคิดเรื่องสิ่งมีชีวิต
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับส่วนตัว

พหุนํ ปณฺฑิโต ชีเว

ข้อสอบการรู้เรื่องวิทยาศาสตร์



ภาพการเคลื่อนที่อะมีบา

ข้อ 23) จงเรียงลำดับกลไกการเคลื่อนที่ของอะมีบา

1. การยื่นเยื่อหุ้มเซลล์ออกไป
2. การเปลี่ยนไซโทรพลาซึมจากโซลเป็นเจลและจากเจลเป็นโซล
3. การแยกตัวและการรวมตัวของโปรตีนแอกทิน
4. การไหลของเอนโดรพลาซึมออกไปด้านนอก

ก. 4 3 2 1

ข. 3 1 4 2

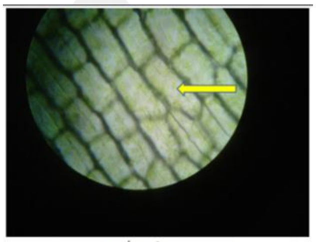
ค. 1 2 3 4

ง. 2 4 1 3

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	แนวความคิดเรื่องสิ่งมีชีวิต
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับส่วนตัว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 24) สมยศนำเซลล์ของใบสาหร่ายหางกระรอก และเซลล์เยื่อหอม มาส่องด้วยกล้องจุลทรรศน์ สังเกตเห็นเม็ดสีเขียวๆ ในเซลล์ของใบสาหร่ายหางกระรอก แต่ไม่พบในเซลล์เยื่อหอม หากสมยศต้องการเห็นภาพจำนวนเซลล์สาหร่ายหางกระรอกมากที่สุดควรเลือกใช้กล้องจุลทรรศน์ชนิดใด



เซลล์สาหร่ายหางกระรอก

	กล้องจุลทรรศน์	เลนส์ใกล้ตา	เลนส์ใกล้วัตถุ
ก.	A	4X	4X
ข.	B	4X	10X
ค.	C	10X	40X
ง.	D	20X	100X

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	แนวความคิดเรื่องสิ่งมีชีวิต
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับส่วนตัว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

เคมีบำบัด

เคมีบำบัด (Chemotherapy) หรือ คีโม เป็นวิธีการรักษาโรคชนิดหนึ่งที่สำคัญวิธีการหนึ่ง โดยการให้ยาที่ออกฤทธิ์ต้าน ชะลอ หรือทำลายเซลล์ร้าย ซึ่งเป้าหมายสำคัญคือเซลล์ที่มีการแบ่งตัวเร็ว กลไกการออกฤทธิ์ของเคมีบำบัด คือการยับยั้งขั้นตอนต่าง ๆ ของการแบ่งเซลล์ เมื่อแบ่งเซลล์ไม่ได้ เซลล์ร้ายก็ไม่สามารถโตต่อได้ เคมีบำบัด ออกฤทธิ์ได้ดีกับเซลล์ที่มีการแบ่งตัวเร็ว ซึ่งเป้าหมายที่แท้จริงคือเซลล์ร้าย แต่ในร่างกายยังมีเซลล์บางชนิดที่แบ่งตัวเร็วเช่นกัน การทำเคมีบำบัดจึงส่งผลกระทบต่อเซลล์ปกติทั่วไป และการทำงานของอวัยวะอื่น ๆ ซึ่งก่อให้เกิดผลข้างเคียงไม่พึงประสงค์ตามมา

ข้อ 25) เซลล์ร้ายที่กล่าวมาข้างต้นคือเซลล์ชนิดใด

ก. เซลล์เม็ดเลือดขาว

ข. เซลล์มะเร็ง

ค. เซลล์ผิวหนัง

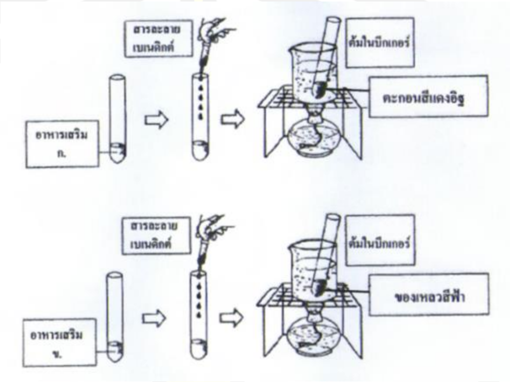
ง. เซลล์เม็ดเลือดแดง

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	มนุษย์
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับท้องถิ่น/ชาติ

พหูน ปณ ฑิโต ชีเว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

โภชนากรทดสอบอาหารเสริมที่เป็นของเหลวใส 2 ชนิด ตามขั้นตอน ดังแสดงในภาพ



ข้อ 26) ถ้าโภชนากรต้องการจัดอาหารเสริมให้ผู้ป่วยเบาหวานควรเลือกอาหารเสริมชนิดใด เพราะเหตุใด

ก. อาหารเสริม ก. เพราะทำปฏิกิริยากับสารละลายเบเนดิกต์

ข. อาหารเสริม ข. เพราะทำปฏิกิริยากับสารละลายเบเนดิกต์

ค. อาหารเสริม ก. เพราะไม่ทำปฏิกิริยากับสารละลายเบเนดิกต์

ง. อาหารเสริม ข. เพราะไม่ทำปฏิกิริยากับสารละลายเบเนดิกต์

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	มนุษย์
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	สุขภาพและโรคภัย
ระดับบริบท	ระดับส่วนตัว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

พิจารณาปริมาณของสารอาหารที่เป็นองค์ประกอบของอาหารชนิดต่างๆ แล้วตอบคำถาม

ชนิดของอาหาร	ปริมาณสารอาหาร			
	โปรตีน (g)	คาร์โบไฮเดรต (g)	ไขมัน (g)	น้ำ (g)
ชนิดที่ 1	200	100	50	400
ชนิดที่ 2	150	120	100	300
ชนิดที่ 3	100	150	120	200
ชนิดที่ 4	80	200	110	180

กำหนดให้

โปรตีนให้พลังงาน 4 กิโลแคลอรีต่อกรัม

คาร์โบไฮเดรตให้พลังงาน 4 กิโลแคลอรีต่อกรัม

ไขมันให้พลังงาน 9 กิโลแคลอรีต่อกรัม

ข้อ 27) อาหารชนิดใดให้พลังงานสูงที่สุดและต่ำที่สุดตามลำดับ

ก. ชนิดที่ 1 ชนิดที่ 2

ข. ชนิดที่ 2 ชนิดที่ 3

ค. ชนิดที่ 3 ชนิดที่ 4

ง. ชนิดที่ 4 ชนิดที่ 1

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	มนุษย์
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	สุขภาพและโรคภัย
ระดับบริบท	ระดับส่วนตัว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 28) การที่ลิงแสมกับลิงกังมีลักษณะคล้ายกัน แต่ที่จริงแล้วเป็นลิงต่างชนิดกัน จัดเป็นความหลากหลายทางใด

ก. ความหลากหลายทางกายภาพ

ข. ความหลากหลายทางชนิดพันธุ์

ค. ความหลากหลายทางพันธุกรรม

ง. ความหลากหลายทางระบบนิเวศ

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	ประชากร
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	คุณภาพสิ่งแวดล้อม
ระดับบริบท	ระดับโลก

พหุ ปญ ฺ ฑิโต ชีเว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

พืชดัดแปลงพันธุกรรม

ข้าวโพดดัดแปลงพันธุกรรมควรถูกห้าม

กลุ่มอนุรักษ์พันธุ์พืชและสัตว์ป่ากำลังเรียกร้องให้ยกเลิกข้าวโพดดัดแปลงพันธุกรรม (GM)

ข้าวโพดดัดแปลงพันธุกรรม ถูกออกแบบมาไม่ให้เกิดผลกระทบจากการใช้สารฆ่าวัชพืชชนิดใหม่ซึ่งฆ่าข้าวโพดพันธุ์ดั้งเดิมได้ด้วย สารฆ่าวัชพืชชนิดใหม่นี้จะฆ่าวัชพืชเกือบทุกชนิดในไร่ข้าวโพด

นักอนุรักษ์บอกว่าเนื่องจากวัชพืชเป็นอาหารของสัตว์เล็กๆ โดยเฉพาะแมลง การใช้สารฆ่าวัชพืชชนิดใหม่กับข้าวโพดดัดแปลงพันธุกรรมจะเป็นผลเสียต่อสิ่งแวดล้อม แต่ผู้สนับสนุนการใช้ข้าวโพดดัดแปลงพันธุกรรมบอกว่า การศึกษาทางวิทยาศาสตร์ได้แสดงว่าสิ่งนี้จะไม่เกิดขึ้น

ต่อไปนี้คือ รายละเอียดของการศึกษาทางวิทยาศาสตร์ที่กล่าวถึงในบทความข้างบน :

มีการปลูกข้าวโพด 200 แปลงทั่วประเทศ แต่ละแปลงถูกแบ่งเป็นสองส่วน ครึ่งหนึ่งปลูกข้าวโพดดัดแปลงพันธุกรรม (GM) ที่ใช้ สารฆ่าวัชพืชชนิดใหม่ ส่วนข้าวโพดพันธุ์ดั้งเดิมที่ใช้สารฆ่าวัชพืชชนิดเดิมถูกปลูกลงแปลงอีกครึ่งหนึ่งที่เหลือ

จำนวนแมลงที่พบในแปลงข้าวโพดดัดแปลงพันธุกรรมที่ใช้สารฆ่าวัชพืชชนิดใหม่มีจำนวนพอๆ กับแมลงในแปลงที่ปลูกข้าวโพดพันธุ์ดั้งเดิมที่ใช้สารฆ่าวัชพืชชนิดเดิม

ข้อ 29) ข้าวโพดถูกปลูกในที่ต่างๆ 200 แปลงทั่วประเทศ เพราะเหตุใดนักวิทยาศาสตร์จึงใช้พื้นที่ปลูกมากกว่าหนึ่งแห่ง

ก. เพื่อเกษตรกรจำนวนมาก จะได้ลองปลูกข้าวโพดดัดแปลงพันธุกรรม

ข. เพื่อดูว่าข้าวโพดดัดแปลงพันธุกรรมจะเจริญเติบโตได้มากเพียงใด

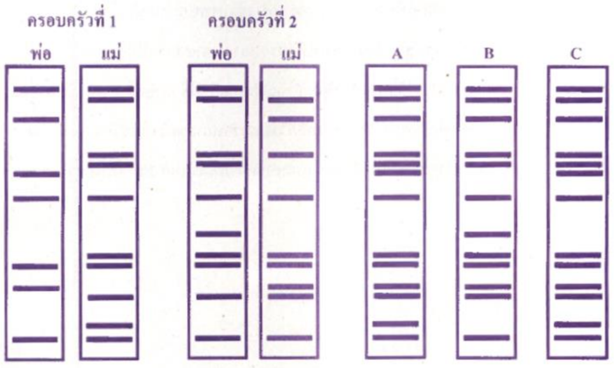
ค. เพื่อให้ข้าวโพดดัดแปลงพันธุกรรมครอบคลุมพื้นที่ให้มากที่สุดเท่าที่จะเป็นได้

ง. เพื่อให้มีสภาวะที่แตกต่างกันหลายๆ แบบในการเจริญเติบโตของข้าวโพด

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	ประชากร
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ลายพิมพ์ดีเอ็นเอของพ่อแม่สองครอบครัวและลูกสามคน ได้แก่ A B และ C เป็นดังภาพ



ข้อ 30) จากข้อมูล ข้อใดระบุความสัมพันธ์ของครอบครัวได้ถูกต้อง

ก. A กับ C เป็นลูกของครอบครัวที่ 1 และ B เป็นลูกของครอบครัวที่ 2

ข. A กับ C เป็นลูกของครอบครัวที่ 2 และ B เป็นลูกของครอบครัวที่ 1

ค. A เป็นลูกของครอบครัวที่ 1 และ B กับ C เป็นลูกของครอบครัวที่ 2

ง. A B และ C เป็นลูกของครอบครัวที่ 1

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	ประชากร
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์
ข้อ 31) ข้อใดจัดว่ามีรูปแบบของความสัมพันธ์ระหว่างสิ่งมีชีวิตแบบเดียวกัน ก. ดอกไม้กับแมลง, ไลเคน ข. นกเอี้ยงกับควาย, พลูด่างกับต้นไม้ ค. เสือกับกวาง, เสือกับสิงโตที่ล่าเหยื่อตัวเดียวกัน ง. กล้วยไม้กับต้นมะม่วง, ปลาฉลามกับเหาฉลาม

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	ระบบนิเวศ
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	คุณภาพสิ่งแวดล้อม
ระดับบริบท	ระดับโลก

ชาวบ้านในหมู่บ้านแห่งหนึ่งนิยมใช้ปุ๋ยเคมีและสารกำจัดแมลงปริมาณมากในการเพาะปลูกเป็นเวลานาน ส่งผลให้เกิดการสะสมของสารเคมีและทำให้ดินเสื่อมสภาพ แม้ปัจจุบันชาวบ้านจะเลิกใช้สารเคมีและปรับปรุงคุณภาพของดินที่ดีขึ้นแล้ว แต่ยังคงมีสารเคมีตกค้างในดินและในระบบนิเวศต่อไป

ข้อ 32) หากต้องการหลีกเลี่ยงการถ่ายทอดสารเคมีที่ตกค้างในดินสู่สิ่งมีชีวิตอื่น ชาวบ้านควรเลือกปลูกพืชเพื่อนำไปใช้ประโยชน์ในด้านใด

ก. การปลูกข้าวเพื่อส่งออก

ข. การปลูกมันเทศเพื่อแปรรูปเป็นอาหารสัตว์

ค. การปลูกมันสำปะหลังเพื่อผลิตน้ำมันเชื้อเพลิง

ง. การปลูกข้าวโพดเพื่อผลิตแป้งสำเร็จรูป

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	ระบบนิเวศ
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	อันตราย
ระดับบริบท	ระดับส่วนตัว

ข้อสอบการรู้เรื่องวิทยาศาสตร์
ข้อ 33) เมื่อพลังงานในสารอาหารถูกถ่ายทอดจากผู้ผลิตไปยังผู้บริโภคลำดับต่อไปได้เพียง 10% ถ้าหนอนตัวหนึ่งกินใบพืชจำนวน 10 กรัมต่อวัน ปริมาณพลังงานที่หนอนสามารถใช้สร้างเป็นเนื้อเยื่อเทียบได้กับใบพืชจำนวนเท่าใด ก. 0.1 กรัม ข. 1 กรัม ค. 10 กรัม ง. 100 กรัม

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	ระบบนิเวศ
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	คุณภาพสิ่งแวดล้อม
ระดับบริบท	ระดับโลก

ข้อ 34) การกระทำใดได้ชื่อว่าเพิ่มรายได้ให้ตนเอง โดยยึดหลักอนุรักษ์ทรัพยากรธรรมชาติ ก. เก็บกล้วยไม้และเฟินจากป่ามาขายให้คนในเมือง ข. จับม้าน้ำมาตากแห้งขายให้กับร้านขายยาโบราณ ค. เก็บเปลือกหอยและเศษปะการังตามชายหาดมาประดิษฐ์เป็นของที่ระลึกขายให้นักท่องเที่ยว ง. เก็บของพลาสติกที่มีคนทิ้งไว้ข้างทางมาสะสมไว้ขาย

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	ไบโอสเฟีย
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับโลก

ข้อสอบการรู้เรื่องวิทยาศาสตร์

วิธีการกำจัดขยะของครัวเรือนในชุมชนหนึ่งมีทั้งการทิ้งตามที่สาธารณะ การทิ้งลงแม่น้ำ และการฝังกลบ ซึ่งต่อมาพบว่า วิธีเหล่านี้ก่อให้เกิดปัญหามลพิษในชุมชน

ข้อ 35) ข้อใดไม่ใช่ปัญหามลพิษที่อาจเกิดขึ้นจากวิธีการกำจัดขยะของครัวเรือนในชุมชนนี้

ก. ปัญหาโรคทางเดินหายใจจากฝุ่นละอองขนาดเล็ก PM 2.5

ข. ปัญหาดินเสื่อมคุณภาพและเกิดการสะสมสารเคมีในโซ่อาหาร

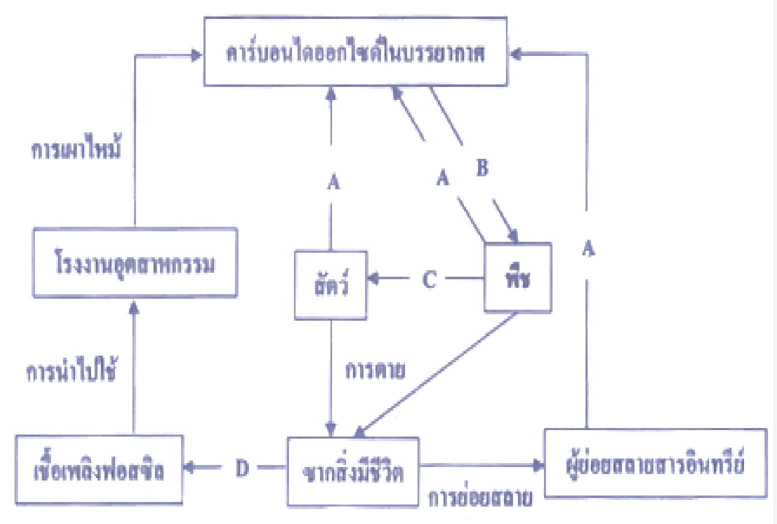
ค. ปัญหาการเน่าเสียของแหล่งน้ำส่งผลให้ปริมาณของสัตว์ลดลง

ง. ปัญหาด้านสุขภาพของประชาชนเนื่องจากน้ำไหลซึมผ่านกองขยะที่สาธารณะ

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	ไบโอสเฟีย
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	คุณภาพสิ่งแวดล้อม
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

A B C และ D คือ กระบวนการที่เกิดขึ้นในวัฏจักรคาร์บอนของบริเวณหนึ่ง ดังภาพ



ข้อ 36) จากแผนภาพ กระบวนการใดของวัฏจักรคาร์บอนนี้ที่ช่วยบรรเทาภาวะเรือนกระจกได้

ก. กระบวนการ A

ข. กระบวนการ B

ค. กระบวนการ C

ง. กระบวนการ D

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบสิ่งมีชีวิต
ความรู้เกี่ยวกับ	ไบโอสเฟีย
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	อันตราย
ระดับบริบท	ระดับโลก

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 37) โครงสร้างของโลก เมื่อแบ่งตามองค์ประกอบทางเคมีจะประกอบด้วยชั้นใดบ้าง

ก. ชั้นเปลือกโลก ชั้นผิวโลก และชั้นเนื้อโลก

ข. ชั้นเปลือกโลก ชั้นเนื้อโลก และชั้นแก่นโลก

ค. ชั้นเปลือกโลก แก่นโลกชั้นใน และชั้นแก่นโลก

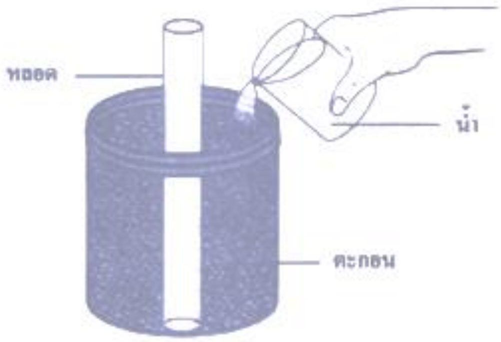
ง. ชั้นเปลือกโลก แก่นโลกชั้นนอก และแก่นโลกชั้นใน

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	โครงสร้างของโลกทั้งระบบ
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับส่วนตัว

พหูน ปณฺฑิโต ชีเว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ทดสอบความสามารถในการกักเก็บน้ำของตะกอนชนิดต่างๆ ที่มีสัดส่วนของอนุภาคกรวดและทราย เป็นส่วนผสมแตกต่างกัน โดยจัดชุดการทดลองดังภาพ จากนั้นเทน้ำปริมาตรเท่ากันลงในภาชนะ แล้ววัดความสูงของระดับน้ำในหลอด



ข้อ 38) จากการทดลอง การคาดคะเนผลการทดลองในข้อความใดไม่ถูกต้อง

ก. ตะกอนที่มีช่องว่างระหว่างตะกอนมาก ความสูงของระดับน้ำในหลอดจะต่ำกว่าตะกอนที่มีช่องว่างระหว่างตะกอนน้อย

ข. หากเทน้ำปริมาตรที่เท่ากันลงในตะกอนชนิดต่างๆ ตะกอนที่มีระดับน้ำในหลอดสูงกว่าจะมีสมบัติในการเป็นชั้นหินอุ้มน้ำดีกว่า

ค. หากเทน้ำลงในตะกอนชนิดต่างๆ ให้น้ำในหลอดมีระดับความสูงเท่ากัน ตะกอนที่เทน้ำลงไปมากกว่าจะมีสมบัติในการเป็นชั้นหินที่อุ้มน้ำดีกว่า

ง. ตะกอนที่มีลักษณะกลมมนและมีขนาดใหญ่ ความสูงของระดับน้ำในหลอดจะต่ำกว่าตะกอนที่มีลักษณะเป็นเหลี่ยมเป็นมุม และมีขนาดคละกันทั้งขนาดเล็กและขนาดใหญ่

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	โครงสร้างของโลกทั้งระบบ
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับส่วนตัว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ฝนกรด

รูปข้างล่างนี้ เป็นรูปแกะสลักที่เรียกว่า แคริยาทิด ซึ่งถูกสร้างไว้ที่มหาวิหารอโครโพลิสในกรุงเอเธนส์ เมื่อกว่า 2,500 ปีมาแล้ว รูปแกะสลักนี้ทำด้วยหินชนิดหนึ่งที่เรียกว่า หินอ่อน หินอ่อนประกอบด้วยแคลเซียมคาร์บอเนต

ในปี ค.ศ. 1980 รูปแกะสลักเดิมถูกย้ายมาอยู่พิพิธภัณฑ์ของอโครโพลิสและเอารูปแกะสลักจำลองไว้แทนที่ เนื่องจากรูปแกะสลักเดิมถูกกัดกร่อนจากฝนกรด

ผลของฝนกรดที่มีต่อหินอ่อน จำลองได้โดยใส่เศษหินอ่อนลงในน้ำส้มสายชู ตั้งทิ้งไว้ค้างคืน น้ำส้มสายชูและฝนกรดมีความเป็นกรดใกล้เคียงกัน เมื่อใส่เศษหินอ่อนลงในน้ำส้มสายชูจะมีฟองก๊าซเกิดขึ้น เราสามารถชั่งน้ำหนักของหินอ่อนแห้งก่อนและหลังการทดลองได้

ข้อ 39) หินอ่อนชิ้นเล็กๆ ก่อนใส่ลงในน้ำส้มสายชูมีมวล 2.0 กรัม เมื่อใส่ลงในน้ำส้มสายชู ตั้งทิ้งไว้ค้างคืน วันรุ่งขึ้นนำเศษหินขึ้นมาและทำให้แห้ง มวลของหินอ่อนที่แห้งแล้วควรเป็นเท่าใด

ก. น้อยกว่า 2.0 กรัม

ข. 2.0 กรัมเท่าเดิม

ค. ระหว่าง 2.0-2.4 กรัม

ง. มากกว่า 2.4 กรัม

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	โครงสร้างของโลกทั้งระบบ
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	อันตราย
ระดับบริบท	ระดับส่วนตัว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ฝุ่น

วิกฤตการณ์ฝุ่นที่เรากำลังเผชิญหน้ากันอยู่ในเวลานี้ไม่ใช่เรื่องธรรมดาอีกต่อไป แต่เริ่มส่งผลกระทบกับสุขภาพและการดำเนินชีวิตของคน ภาวะที่เกิดขึ้นขณะนี้ เรียกว่า “Smog Air Pollution” Smog เกิดจากหลายสาเหตุทั้งจากธรรมชาติและน้ำมือมนุษย์ สาเหตุจากธรรมชาติ ก็เช่น ไฟป่า เถ้าจากภูเขาไฟระเบิด เป็นต้น ส่วนที่เกิดจากน้ำมือมนุษย์ก็อย่างที่เราพอจะรู้กันว่าเป็นฝุ่นจากการก่อสร้าง การเผาไหม้เครื่องยนต์ ผ้าเบรก การเหยียบเบรกบ่อย ๆ รถติด เผาป่า โรงงานอุตสาหกรรมที่ปล่อยควันไม่ได้มาตรฐาน เป็นต้น Smog ที่เกิดไม่ได้มีแต่ฝุ่นดิน ฝุ่นปูนเท่านั้น แต่ยังมีละออง คาร์บอนไดออกไซด์ ไนตรัส คาร์บอนมอนอกไซด์ ปรอท แคดเมียม และสารอีกหลายอย่างปน ๆ กันไป ละอองบางอันมีขนาดใหญ่ ละอองบางอันมีขนาดเล็ก ฝุ่นยิ่งมีขนาดเล็กยิ่งอันตรายเพราะสามารถผ่านเข้าไปสู่ทางเดินหายใจและปอดได้โดยง่าย และยังส่งผลต่อร่างกายโดยรวมด้วย

ข้อ 40) จากข้อมูลข้างต้น ข้อความใดกล่าวได้ถูกต้อง

ก. วิกฤติฝุ่นควันสาเหตุทั้งหมดเกิดจากน้ำมือมนุษย์

ข. ฝุ่นอนุภาคใหญ่อันตรายกว่าฝุ่นอนุภาคขนาดเล็ก

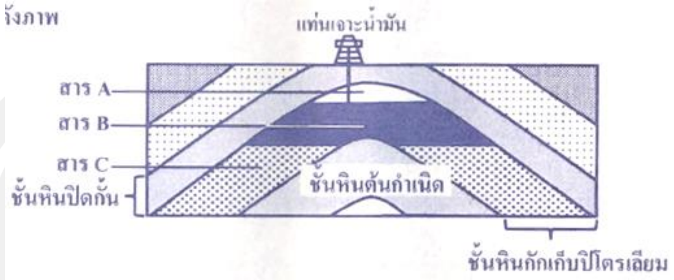
ค. คำว่า Smog มีทั้งสถานะของแข็งของเหลวและก๊าซ

ง. Smog มีสาเหตุมากจากธรรมชาติทั้งหมด

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	พลังงานในระบบโลก
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	คุณภาพสิ่งแวดล้อม
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

โครงสร้างชั้นหินของแหล่งกักเก็บปิโตรเลียม และสารในแหล่งกักเก็บปิโตรเลียม เรียงลำดับชั้นเป็นดังภาพ



สมบัติของโครงสร้างชั้นหินของแหล่งกักเก็บปิโตรเลียม เป็นดังนี้

ชั้นหิน	สมบัติ
หินต้นกำเนิด	มีสารอินทรีย์สะสมอยู่มาก ซึ่งสามารถเปลี่ยนสภาพไปเป็นปิโตรเลียมภายใต้สภาวะความร้อนและความดันที่เหมาะสม
หินกักเก็บปิโตรเลียม	มีความพรุน รอยแตกหรือโพรงให้ของเหลวและแก๊สไหลผ่าน และสามารถกักเก็บปิโตรเลียมได้
หินปิดกั้น	มีเนื้อละเอียด ของเหลวและแก๊สซึมได้น้อย จึงปิดกั้นไม่ให้ปิโตรเลียมรั่วไหลออกไปได้

ข้อ 41) จากข้อมูล ข้อสรุปใดต่อไปนี้ถูกต้อง

ก. หินที่มีสมบัติเป็นชั้นหินต้นกำเนิด คือ หินตะกอน

ข. หินที่มีสมบัติเป็นชั้นหินปิดกั้น คือ หินทราย

ค. หินที่มีสมบัติเป็นชั้นหินกักเก็บปิโตรเลียม คือ หินแปร

ง. เรียงลำดับความหนาแน่นของสารในแหล่งกักเก็บปิโตรเลียมจากมากไปน้อย คือ A B และ C ตามลำดับ

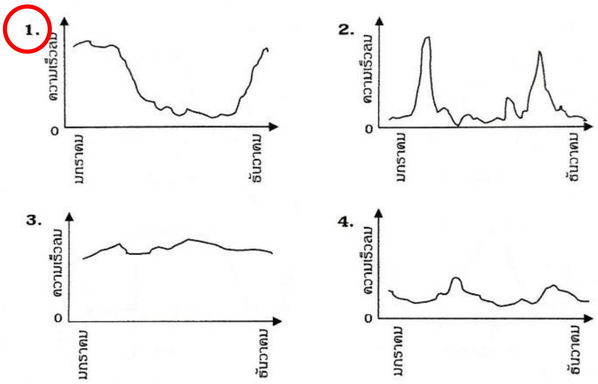
ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	พลังงานในระบบโลก
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

การผลิตพลังงานลม

คนจำนวนมากเชื่อว่าลมสามารถเป็นแหล่งพลังงานทดแทนน้ำมันและถ่านหินซึ่งเป็นแหล่งผลิตกระแสไฟฟ้าในรูปกังหันลมที่ใช้ลมหมุนใบพัด การหมุนนี้ทำให้พลังงานไฟฟ้าที่เกิดขึ้นโดยเครื่องกำเนิดไฟฟ้าที่ถูกหมุนด้วยกังหัน

ข้อ 42) กราฟข้างล่างนี้ แสดงความเร็วลมเฉลี่ยตลอดปีในสี่บริเวณต่างกัน กราฟในข้อใดชี้บอกบริเวณที่เหมาะสมในการติดตั้งเครื่องผลิตกระแสไฟฟ้าจากพลังลม



ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	พลังงานในระบบโลก
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 43) เทือกเขาหิมาลัย เกิดจากปรากฏการณ์ทางธรณีภาคแบบใด

ก. การเกิดแผ่นดินไหว

ข. การแยกตัวของแผ่นเปลือกโลก

ค. การชนกันของแผ่นเปลือกโลก

ง. การระเบิดของภูเขาไฟ

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	การเปลี่ยนแปลงในระบบโลก
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับโลก

พหูน ปณ ฑิโต ชีเว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

คลื่นสึนามิ

คลื่นสึนามิในประเทศไทยเมื่อวันที่ 26 ธันวาคม พ.ศ. 2547 เกิดขึ้นเนื่องจากแผ่นดินไหวที่บริเวณร่องลึกซุนดรา (Sundra trench) เกิดการยุบตัวของเปลือกโลกบริเวณรอยต่อของแผ่นธรณีอินเดีย (India plate) กับแผ่นธรณีพม่า (Burma microplate) ทำให้เกิดแรงสั่นสะเทือน 9.1 ริกเตอร์ โดยมีจุดเหนือศูนย์เกิดแผ่นดินไหวอยู่ทางทิศตะวันตกเฉียงเหนือของเกาะสุมาตรา เหตุการณ์นี้ทำให้คนตายมากกว่า 226,000 คน ตามชายฝั่งของมหาสมุทรอินเดีย ในจำนวนนี้เป็นคนไทยไม่น้อยกว่า 5,300 คน

ข้อ 44) ข้อใดเรียงลำดับจากเหตุไปหาผลได้ถูกต้อง

1. ใต้แผ่นเปลือกโลก มีหินหนืด ความร้อนสูง
2. แผ่นเปลือกโลกชนกัน
3. เกิดสึนามิ
4. เกิดแผ่นดินไหว
5. แผ่นเปลือกโลกมีการเคลื่อนที่

ก. 5 → 3 → 4 → 2 → 1

ข. 1 → 5 → 2 → 4 → 3

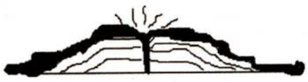
ค. 5 → 4 → 3 → 1 → 2

ง. 3 → 2 → 1 → 5 → 4

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	การเปลี่ยนแปลงในระบบโลก
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	อันตราย
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

จากข้อมูลต่อไปนี้ใช้ตอบคำถาม

ประเภท	ลักษณะ
แบบที่ 1	รูปร่างเหมือนกรวยคว่ำ มีความสูงชันมาก มีขนาดเล็ก หินหนืดปะทุระเบิดแล้วเย็นตัวอย่างรวดเร็ว การไหลไม่ต่อเนื่อง
แบบที่ 2	รูปร่างเป็นกรวยสูง มีทางไหลของลาวาหลายทาง มีฐานแผ่ขยายใหญ่ หินหนืดมีความหนืดสูง ลาวาอาจมีกาเปลี่ยนแปลงการปะทุอย่างกะทันหัน
แบบที่ 3 	ลักษณะเตี้ยแต่มีขนาดใหญ่คล้ายรูปโล่คว่ำ หินหนืดมีอุณหภูมิสูง ลาวาไหลออกจากปากปล่องและไม่กองสูงชัน

ข้อ 45) จากประเภทของภูเขาไฟและลักษณะดังกล่าว ภูเขาไฟแบบใดที่ลาวาไหลเร็วและไปได้ไกลที่สุด

ก. แบบที่ 1

ข. แบบที่ 2

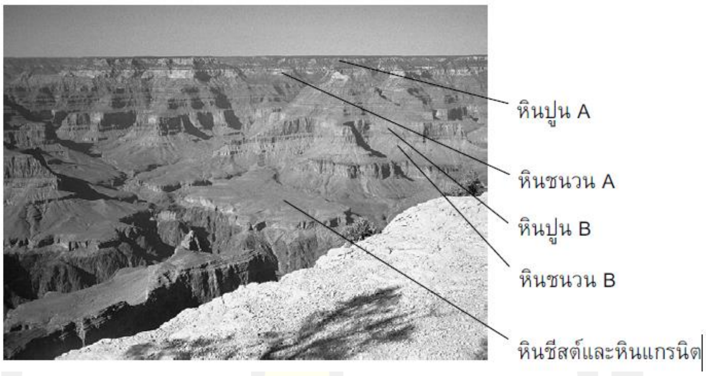
ค. แบบที่ 3

ง. แบบที่ 2 และ 3

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	การเปลี่ยนแปลงในระบบโลก
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

จงดูรูปข้างล่างซึ่งเป็นรูปของแกรนด์แคนยอนที่ถ่ายจากขอบด้านทิศใต้ สามารถเห็นชั้นหินที่แตกต่างกันหลายชั้นได้ชัดตามแนวผนังของหุบเขา



ข้อ 46) มีฟอสซิลของซากสัตว์ทะเลหลายชนิด เช่น หอยกาบ ปลา และปะการัง อยู่ในชั้นหินปูน A ของแกรนด์แคนยอน มีอะไรเกิดขึ้น เมื่อหลายล้านปีก่อนที่อธิบายว่าทำไมฟอสซิลเหล่านี้จึงถูกพบที่นั่น

ก. ในสมัยโบราณผู้คนนำอาหารทะเลจากมหาสมุทรเข้ามาในบริเวณนี้
ข. ครั้งหนึ่งมหาสมุทรมีคลื่นรุนแรงมากและคลื่นยักษ์พัดพาสิ่งมีชีวิตในทะเลขึ้นมาบนบก
ค. ในสมัยก่อนบริเวณนั้นเป็นบริเวณที่มีมหาสมุทรปกคลุม และได้เหือดแห้งไปในตอนหลัง
ง. สัตว์ทะเลบางชนิดครั้งหนึ่งมีชีวิตอยู่บนบกก่อนที่จะอพยพลงสู่ทะเล

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	ประวัติศาสตร์ของโลก
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับโลก

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 47) ซากดึกดำบรรพ์ที่สมบูรณ์มักเป็นสัตว์ทะเลมากกว่าสัตว์จากแหล่งอื่นเพราะเหตุใด

1. ซากสัตว์ทะเลจะจมสู่ท้องทะเล มีโคลนและตะกอนละเอียด
2. แร่ธาตุหลายชนิดที่อยู่ในน้ำทะเลจะซึมเข้าในช่องตัวสัตว์ ทำให้ซากทนต่อการผุพัง
3. น้ำทะเลมีความเค็มย่อมรักษาสภาพสัตว์ให้คงสภาพเดิมไว้ได้

ก. 1 2

ข. 2 3

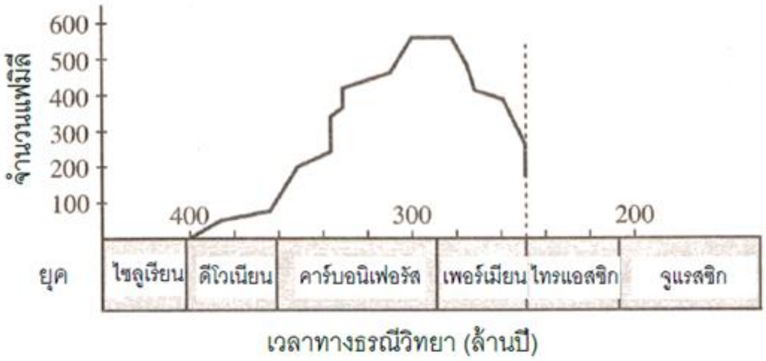
ค. 3 1

ง. 1 2 3

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	ประวัติศาสตร์ของโลก
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับโลก

ข้อสอบการรู้เรื่องวิทยาศาสตร์

จากการศึกษาซากสัตว์ประเภทหอยตระกูลหนึ่ง แล้วเขียนกราฟแสดงความสัมพันธ์ระหว่างจำนวน แฟมิลีกับเวลาทางธรณีวิทยาได้ดังรูป



ข้อ 48) จากข้อมูลข้างต้น ข้อใดต่อไปนี้ถูกต้อง

1. สูญพันธุ์เมื่อปลายยุคเพอร์เมียน
2. กำเนิดขึ้นในโลกประมาณ 400 ล้านปีมาแล้ว
3. มีจำนวนแฟมิลีมากที่สุดเมื่อ 500 ล้านปีมาแล้ว
4. มีชีวิตอยู่ระหว่างต้นยุคดีโวเนียนถึงปลายยุคเพอร์เมียน

ก. 1 2 3

ข. 1 2 4

ค. 2 3 4

ง. 1 2 3 4

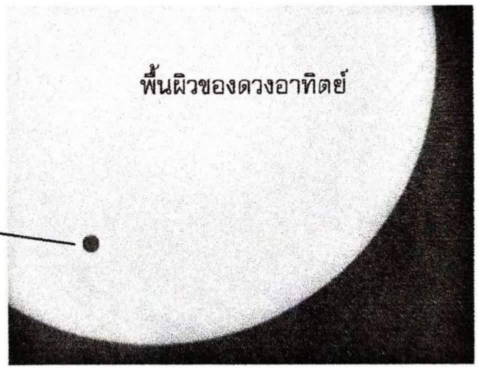
ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	ประวัติศาสตร์ของโลก
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	ทรัพยากรธรรมชาติ
ระดับบริบท	ระดับโลก

ข้อสอบการรู้เรื่องวิทยาศาสตร์

การเคลื่อนผ่านของดาวศุกร์

วันที่ 8 มิถุนายน ค.ศ. 2004 สามารถมองเห็นดาวศุกร์เคลื่อนที่ผ่านดวงอาทิตย์ได้ในหลายบริเวณของโลก เรียกว่าปรากฏการณ์นี้ว่า “การเคลื่อนผ่าน” ของดาวศุกร์ และจะเกิดขึ้นเมื่อวงโคจรของดาวศุกร์มาอยู่ระหว่างดวงอาทิตย์และโลก การเคลื่อนผ่านของดาวศุกร์ครั้งที่แล้วเกิดขึ้นในปี ค.ศ. 1882 และมีการทำนายว่าครั้งต่อไปจะเกิดขึ้นในปี ค.ศ. 2012

รูปข้างล่าง แสดงถึงการเคลื่อนผ่านของดาวศุกร์ในปี ค.ศ. 2004 โดยส่องกล้องโทรทรรศน์ไปที่ดวงอาทิตย์และฉายภาพลงบนกระดาษขาว



ข้อ 49) ทำไมการสังเกตการเคลื่อนผ่านของดาวศุกร์จึงต้องฉายภาพลงบนกระดาษขาวแทนที่จะมองผ่านกล้องโทรทรรศน์ด้วยตาเปล่าโดยตรง

ก. แสงอาทิตย์สว่างมากเกินไปที่จะมองเห็นดาวศุกร์ได้

ข. ดวงอาทิตย์มีขนาดใหญ่มากจนมองเห็นได้โดยไม่ต้องขยาย

ค. การมองดวงอาทิตย์ผ่านกล้องโทรทรรศน์อาจเป็นอันตรายต่อดวงตา

ง. ต้องทำภาพให้เล็กลงด้วยกันฉายลงบนกระดาษ

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	โลกในอวกาศ
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	สุขภาพและโรคภัย
ระดับบริบท	ระดับท้องถิ่น/ชาติ

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ยิงจรวด 2 ลูก ขนาดเดียวกันจากฐานยิง A และ B บนพื้นผิวโลกที่ความสูงต่างระดับกัน โดย B อยู่สูงกว่า A ปรากฏว่าจรวดทั้งคู่สามารถนำยานอวกาศไปโคจรรอบดวงอาทิตย์ได้

ข้อ 50) ข้อใดถูก

1. จรวดที่ A ถูกยิงด้วยความเร็วของการผละหนีมากกว่าที่ B
2. เมื่อเริ่มเคลื่อนที่ จรวดที่ยิงจากฐาน A มีแรงขับดันเท่ากับจรวดที่ยิงจากฐาน B
3. จรวดที่ A ถูกยิงด้วยความเร็วของการผละหนีน้อยกว่าที่ B
4. จรวดทั้งสองหลุดพ้นแรงโน้มถ่วงของโลกได้

ก. 1 2

ข. 2 3

ค. 3 4

ง. 4 1

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	โลกในอวกาศ
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับโลก

พหูน ปณ ฑิโต ชีเว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 51) ถ้าโลกมีแรงโน้มถ่วงมากกว่าดวงจันทร์ 6 เท่า ดังนั้นคนที่มีน้ำหนักตัวบนพื้นผิวโลก 60 กิโลกรัม เมื่อไปชั่งน้ำหนักบนดวงจันทร์จะหนักเท่าไหร่

ก. 10 กิโลกรัม

ข. 12 กิโลกรัม

ค. 60 กิโลกรัม

ง. 180 กิโลกรัม

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	โลกในอวกาศ
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับโลก

พหูน ปณ ฑิโต ชีเว

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 52) ข้อใดอธิบายเกี่ยวกับทฤษฎีบิกแบงได้ถูกต้องที่สุด

ก. เป็นการระเบิดอย่างรุนแรงของสารที่ร้อนจัด ซึ่งมีความเร็วสูงมาก และแผ่กระจายออกไปอย่างรวดเร็ว ซึ่งทำให้เกิดเอกภพขึ้น

ข. เป็นการระเบิดอย่างรุนแรงของสารที่ร้อนจัด ซึ่งมีความเร็วสูงมาก จากนั้นจึงรวมตัวกันอย่างรวดเร็ว ซึ่งทำให้เกิดเอกภพขึ้น

ค. เป็นการระเบิดอย่างรุนแรงของสารที่ร้อนจัด ซึ่งมีความหนาแน่นสูงมาก และแผ่กระจายออกไปอย่างรวดเร็ว ซึ่งทำให้เกิดเอกภพขึ้น

ง. เป็นการระเบิดอย่างรุนแรงของสารที่ร้อนจัด ซึ่งมีความหนาแน่นสูงมาก จากนั้นจึงรวมตัวกันอย่างรวดเร็ว ซึ่งทำให้เกิดเอกภพขึ้น

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	ประวัติศาสตร์และขนาดของจักรวาล
สมรรถนะทางวิทยาศาสตร์	การอธิบายปรากฏการณ์ในเชิงวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับโลก

ข้อสอบการรู้เรื่องวิทยาศาสตร์

ข้อ 53) ทฤษฎีบิกแบงสามารถใช้อธิบายอุณหภูมิของเอกภพและปริมาณของอนุภาคในเอกภพได้ตามข้อใด

1. อุณหภูมิของเอกภพลดลง
2. อุณหภูมิของเอกภพเพิ่มขึ้น
3. ปริมาณของอนุภาคน้อยกว่าปฏิอนุภาค
4. ปริมาณของอนุภาคมากกว่าปฏิอนุภาค

ก. 1 และ 3

ข. 2 และ 3

ค. 1 และ 4

ง. 2 และ 4

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	ประวัติศาสตร์และขนาดของจักรวาล
สมรรถนะทางวิทยาศาสตร์	การประเมินและออกแบบกระบวนการสืบเสาะหาความรู้ทางวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับโลก

พหูน ปณฺฑิโต ชีเว

ข้อสอบการรู้เรื่องวิทยาศาสตร์
ข้อ 54) ระยะทางจากโลกถึงดวงอาทิตย์ มีค่าประมาณ 150 ล้านกิโลเมตร เราทราบว่าแสงเดินทางด้วยอัตราเร็ว 300,000 กิโลเมตรต่อวินาที แสดงว่าแสงเดินทางจากดวงอาทิตย์มายังโลกใช้เวลาเท่าไร ก. ประมาณ 10 ปี ข. ประมาณ 14 ปี ค. ประมาณ 8 นาที ง. ประมาณ 8 เดือน

ลักษณะเฉพาะของข้อสอบ	
ความรู้ด้านเนื้อหาวิทยาศาสตร์	ระบบของโลกและอวกาศ
ความรู้เกี่ยวกับ	ประวัติศาสตร์และขนาดของจักรวาล
สมรรถนะทางวิทยาศาสตร์	การแปลความหมายข้อมูลและการใช้ประจักษ์พยานในเชิงวิทยาศาสตร์
บริบท	ความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยี
ระดับบริบท	ระดับโลก

ภาคผนวก ง

ตัวอย่าง Print Out ผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบการรู้เรื่องวิทยาศาสตร์ ระดับชั้นมัธยมศึกษาปีที่ 3 ด้วยวิธี MIMIC โดยโปรแกรม Mplus

พหูน ปณ ฑิโต ชีเว

```
Mplus VERSION 8.3
MUTHEN & MUTHEN
08/07/2021  11:29 PM
INPUT INSTRUCTIONS

TITLE:  BIO DIF
 DATA:
    FILE IS "C:\Users\ASUS\Desktop\2.ÃÐººÊÔè§ÁÕªÕÇÔµ\BIO DIF.dat";

  VARIABLE:
    NAMES ARE Item16-Item30 Gender;
    USEVARIABLES ARE Item16-Item30 Gender;
    CATEGORICAL ARE Item16-Item30;

  ANALYSIS:
    TYPE IS GENERAL;
    ESTIMATOR IS WLSMV;
    ITERATIONS = 1000;
    CONVERGENCE = 0.00005;

  MODEL:
    BIO BY Item16-Item30;
    Item16-Item30 ON Gender;
    ITEM27   WITH ITEM26;
    ITEM27   WITH ITEM23;
    ITEM28   WITH ITEM27;
    ITEM30   WITH ITEM29;
    ITEM30   WITH ITEM27;
    ITEM29   WITH ITEM28;
    ITEM29   WITH ITEM27;
    ITEM30   WITH ITEM23;
```

```
ITEM29   WITH ITEM23;
ITEM28   WITH ITEM23;
ITEM26   WITH ITEM23;
ITEM30   WITH ITEM26;
ITEM30   WITH ITEM28;
ITEM27   WITH ITEM25;
ITEM24   WITH ITEM23;
ITEM28   WITH ITEM26;
```

MODEL FIT INFORMATION

Number of Free Parameters   61

Chi-Square Test of Model Fit

Value	127.301*
Degrees of Freedom	74
P-Value	0.0001

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.032
90 Percent C.I.	0.023 0.042
Probability RMSEA <= .05	0.999

CFI/TLI

CFI	0.999
TLI	0.999

Chi-Square Test of Model Fit for the Baseline Model

Value	70552.926
Degrees of Freedom	120
P-Value	0.0000

SRMR (Standardized Root Mean Square Residual)

Value 0.024

Optimum Function Value for Weighted Least-Squares Estimator

Value 0.38579951D-01

STANDARDIZED MODEL RESULTS

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
BIO BY				
ITEM16	0.971	0.013	74.501	0.000
ITEM17	0.954	0.016	60.889	0.000
ITEM18	0.966	0.013	73.198	0.000
ITEM19	0.948	0.016	57.620	0.000
ITEM20	0.965	0.013	73.257	0.000
ITEM21	0.937	0.017	55.069	0.000
ITEM22	0.947	0.016	59.751	0.000
ITEM23	0.824	0.032	25.415	0.000
ITEM24	0.707	0.038	18.827	0.000
ITEM25	0.427	0.053	8.085	0.000
ITEM26	0.803	0.030	26.419	0.000
ITEM27	0.715	0.040	18.069	0.000
ITEM28	0.894	0.022	40.373	0.000
ITEM29	0.840	0.025	34.274	0.000
ITEM30	0.907	0.020	44.462	0.000

ITEM16 ON				
GENDER	0.232	0.051	4.584	0.000
ITEM17 ON				
GENDER	0.256	0.054	4.721	0.000
ITEM18 ON				
GENDER	0.196	0.057	3.445	0.001
ITEM19 ON				
GENDER	0.199	0.056	3.544	0.000
ITEM20 ON				
GENDER	0.167	0.057	2.930	0.003
ITEM21 ON				
GENDER	0.287	0.052	5.469	0.000
ITEM22 ON				
GENDER	0.229	0.055	4.136	0.000
ITEM23 ON				
GENDER	0.122	0.048	2.575	0.010
ITEM24 ON				
GENDER	0.190	0.046	4.134	0.000
ITEM25 ON				
GENDER	0.041	0.048	0.858	0.391
ITEM26 ON				
GENDER	0.149	0.047	3.150	0.002
ITEM27 ON				
GENDER	0.241	0.045	5.380	0.000
ITEM28 ON				
GENDER	0.228	0.048	4.706	0.000
ITEM29 ON				
GENDER	0.413	0.040	10.212	0.000
ITEM30 ON				
GENDER	0.248	0.048	5.176	0.000

ITEM27 WITH				
ITEM26	-0.923	0.167	-5.519	0.000
ITEM23	-0.825	0.175	-4.723	0.000
ITEM28	-0.918	0.175	-5.229	0.000
ITEM30	-0.928	0.193	-4.815	0.000
ITEM29	-0.704	0.206	-3.428	0.001
ITEM25	-0.262	0.098	-2.669	0.008
ITEM28 WITH				
ITEM29	0.425	0.102	4.145	0.000
ITEM23	-0.793	0.236	-3.361	0.001
ITEM30	0.911	0.027	33.918	0.000
ITEM26	-0.505	0.154	-3.267	0.001
ITEM30 WITH				
ITEM29	0.492	0.101	4.888	0.000
ITEM23	-0.975	0.289	-3.380	0.001
ITEM26	-0.756	0.168	-4.514	0.000
ITEM29 WITH				
ITEM23	-0.888	0.249	-3.561	0.000
ITEM26 WITH				
ITEM23	-0.563	0.167	-3.369	0.001
ITEM24 WITH				
ITEM23	-0.361	0.134	-2.689	0.007

Thresholds

ITEM16$1	1.102	0.065	17.086	0.000
ITEM17$1	1.341	0.070	19.172	0.000

ITEM18$1	1.319	0.072	18.332	0.000
ITEM19$1	1.292	0.071	18.209	0.000
ITEM20$1	1.262	0.071	17.665	0.000
ITEM21$1	1.342	0.069	19.570	0.000
ITEM22$1	1.336	0.071	18.825	0.000
ITEM23$1	0.433	0.059	7.390	0.000
ITEM24$1	0.446	0.057	7.786	0.000
ITEM25$1	0.285	0.059	4.827	0.000
ITEM26$1	0.523	0.059	8.904	0.000
ITEM27$1	0.436	0.056	7.758	0.000
ITEM28$1	0.921	0.061	15.052	0.000
ITEM29$1	0.995	0.056	17.822	0.000
ITEM30$1	0.950	0.061	15.579	0.000

Variances

BIO	1.000	0.000	999.000	999.000

พหุน ปณ ฑิโต ชีเว

ประวัติผู้เขียน

ชื่อ	นายศราวุทธ โพธิ์สิงห์
วันเกิด	วันที่ 19 กันยายน พ.ศ. 2530
สถานที่เกิด	อำเภอกันทรวิชัย จังหวัดมหาสารคาม
สถานที่อยู่ปัจจุบัน	บ้านเลขที่ 111 หมู่ที่ 9 ตำบลนาสีนวน อำเภอกันทรวิชัย จังหวัดมหาสารคาม รหัสไปรษณีย์ 44150
ตำแหน่งหน้าที่การงาน	ครู
สถานที่ทำงานปัจจุบัน	โรงเรียนเพชรวิทยาคาร อำเภอภูเขียว จังหวัดชัยภูมิ รหัสไปรษณีย์ 36110

ประวัติการศึกษา	พ.ศ. 2548	มัธยมศึกษาปีที่ 6 โรงเรียนสารคามพิทยาคม จังหวัดมหาสารคาม
	พ.ศ. 2552	ปริญญาวิทยาศาสตรบัณฑิต (วท.บ.) สาขาวิชาชีววิทยา มหาวิทยาลัยมหาสารคาม
	พ.ศ. 2554	ประกาศนียบัตรบัณฑิต (ป.บัณฑิต) สาขาวิชาชีพครู มหาวิทยาลัยราชภัฏมหาสารคาม
	พ.ศ. 2565	ปริญญาการศึกษามหาบัณฑิต (กศ.ม.) สาขาวิชาวิจัยและประเมินผลการศึกษา มหาวิทยาลัยมหาสารคาม