



การเปรียบเทียบผลการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนในวิชาชีววิทยา
ของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย ระหว่างแบบสอบวินิจฉัยแบบเลือกตอบสามระดับกับสี่ระดับ

A COMPARISON OF THE MISCONCEPTIONS DIAGNOSTIC RESULTS BETWEEN
THREE-TIER AND FOUR-TIER DIAGNOSTIC TESTS IN BIOLOGY
FOR SECONDARY SCHOOL STUDENTS

นายธนบดี อินหาดกรวด*

Dhanabodee Inhadkruad

ผู้ช่วยศาสตราจารย์ ดร.ณัฐภรณ์ หลาวทอง**

Asst. Prof. Nuttaporn Lawthong, Ph.D.

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์คือ 1) เพื่อเปรียบเทียบผลการวินิจฉัยระหว่างแบบสอบวินิจฉัยแบบเลือกตอบสามระดับกับแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับวิชาชีววิทยาสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย และ 2) เพื่อวิเคราะห์ความสอดคล้องของผลการวินิจฉัยระหว่างแบบสอบวินิจฉัยแบบเลือกตอบสามระดับและแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับวิชาชีววิทยาสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลายกับข้อมูลเชิงประจักษ์จากการสัมภาษณ์ตัวอย่างในการวิจัยคือนักเรียนระดับมัธยมศึกษาตอนปลายจำนวน 62 คน เครื่องมือวิจัยได้แก่ 1) แบบสอบวินิจฉัยแบบเลือกตอบสามระดับและสี่ระดับ 2) แบบสัมภาษณ์โดยใช้เทคนิคการคิดออกเสียง การวิเคราะห์ข้อมูลใช้การวิเคราะห์ค่าสถิติพื้นฐาน ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน ค่าสัมประสิทธิ์แอลฟาของครอนบาค และค่าสัมประสิทธิ์สหสัมพันธ์ของคราเมอร์ ผลการเปรียบเทียบการวินิจฉัยระหว่างแบบสอบวินิจฉัยแบบเลือกตอบสามระดับกับสี่ระดับด้วยการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ของคราเมอร์พบว่า ส่วนใหญ่มีความสอดคล้องกันในระดับปานกลาง (V มีค่าระหว่าง 0.322 ถึง 0.489) เมื่อเปรียบเทียบผลการวินิจฉัยด้วยแบบสอบวินิจฉัยแบบเลือกตอบสามระดับและสี่ระดับกับผลการสัมภาษณ์พบว่า แบบสอบวินิจฉัยแบบเลือกตอบสามระดับกับสี่ระดับให้ผลการวินิจฉัยสอดคล้องกับข้อมูลเชิงประจักษ์จากการสัมภาษณ์ไม่แตกต่างกัน

*สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

จุฬาลงกรณ์มหาวิทยาลัย ถนนพญาไท แขวงวังใหม่ เขตปทุมวัน กรุงเทพมหานคร 10330

E-mail Address: dhanabodee13178@gmail.com

**อาจารย์ประจำสาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

E-mail Address: Nuttaporn.L.@chula.ac.th

ISSN1905-4491

Abstract

The purposes of this study were 1) to compare diagnostic results between three-tier and four-tier diagnostic tests in biology for secondary school students, and 2) to analyze the congruence among diagnostic results from three-tier and four-tier diagnostic tests and the think aloud interviews. The participants consisted of 62 upper secondary students of the 2017 academic year. The instruments used in this research were 1) the three-tier and four-tier diagnostic tests, and 2) the think aloud interview. The data were analyzed by descriptive statistics, Pearson's correlation, Cronbach's alpha coefficient, and Cramer's V correlation coefficient. The results of this study showed the average correlation level of diagnostic results both of three-tier and four-tier diagnostic tests (Cramer' V range 0.322 to 0.489). The correlation coefficient between the diagnostic results of the four-tier diagnostic test and think aloud interviews was not different to those of the three-tier diagnostic test and think aloud interviews.

คำสำคัญ: แบบสอบวินิจฉัยแบบเลือกตอบสามระดับ / แบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับ / ผลการวินิจฉัย

KEYWORDS: THREE-TIER DIAGNOSTIC TESTS / FOUR-TIER DIAGNOSTIC TESTS / DIAGNOSTIC RESULT

บทนำ

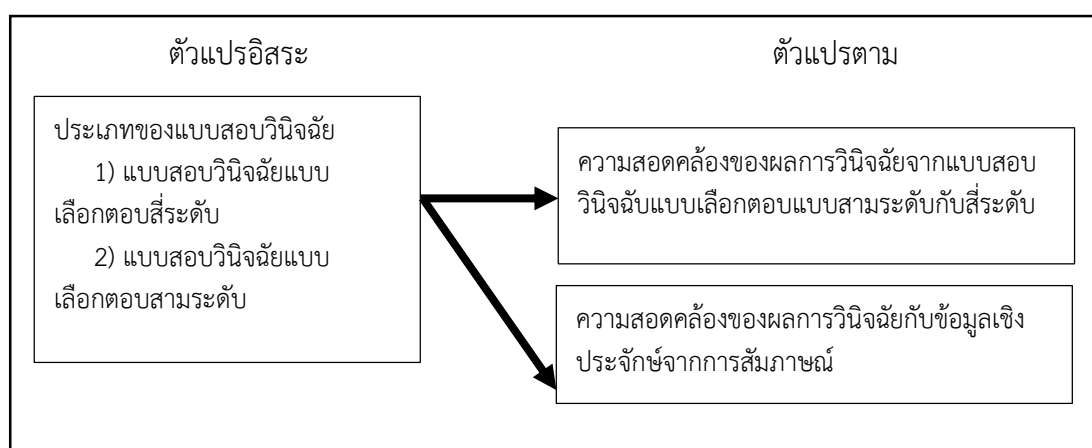
ปัจจุบันรัฐบาลมีวิสัยทัศน์ในการบริหารประเทศเพื่อให้ประเทศไทย “มั่นคง มั่งคั่ง ยั่งยืน” ทั้งนี้เพื่อให้สามารถรับมือกับโอกาสและการเปลี่ยนแปลงทางสังคมอย่างรวดเร็วในศตวรรษที่ 21 ซึ่งมุ่งเน้นพัฒนาประเทศให้เป็นประเทศเศรษฐกิจใหม่ ขับเคลื่อนเศรษฐกิจด้วยการสร้างนวัตกรรม การพัฒนาคนในชาติให้มีความสามารถในการคิดเพื่อสร้างนวัตกรรมนั้น สอดคล้องกับแนวทางของวิทยาศาสตร์ศึกษาที่เน้นพัฒนาวิธีคิด ทั้งความคิดเป็นเหตุเป็นผล คิดสร้างสรรค์ คิดวิเคราะห์ วิจัย มีทักษะสำคัญในการค้นคว้าหาความรู้ มีความสามารถในการแก้ปัญหาอย่างเป็นระบบ การจัดการเรียนรู้วิชาวิทยาศาสตร์ในปัจจุบัน เป็นการเน้นการจัดการเรียนการสอนที่มุ่งให้ผู้เรียนเป็นผู้สร้างองค์ความรู้ขึ้นมาด้วยตนเอง ตามแนวทางการศึกษากลุ่มคอนสตรัคติวิสต์ บทบาทของนักเรียนเป็นผู้สืบสอบ ค้นคว้าหาความรู้ สร้างองค์ความรู้ และเชื่อมโยงองค์ความรู้ที่มีอยู่เข้ากับความรู้เดิมของนักเรียน (พิมพันธ์ เดชะคุปต์, 2558) ครูจึงจำเป็นต้องคอยตรวจสอบความเข้าใจของนักเรียนอยู่เสมอเพื่อป้องกันไม่ให้เกิดความเข้าใจที่คลาดเคลื่อน หรือมีโนทัศน์ที่คลาดเคลื่อน (Misconceptions) ที่อาจเกิดขึ้นระหว่าง การเรียนรู้ของนักเรียน

จากผลการทดสอบทางการศึกษาระดับชาตินี้พื้นฐาน (Ordinary National Educational Test: O-NET) ซึ่งเป็นการวัดและประเมินผลความรู้ตามหลักสูตรแกนการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ในหลักสูตรมัธยมศึกษาตอนปลายของนักเรียนชั้นมัธยมศึกษาปีที่ 6 โดยสำนักทดสอบทางการศึกษา พบว่าตั้งแต่ปีการศึกษา 2551 ถึง 2558 โดยส่วนใหญ่แล้วมาตรฐาน ว1.2 มีผลคะแนนเฉลี่ยต่ำในหลายปีการศึกษา เนื้อหาตามมาตรฐาน ว1.2 แบ่งออกเป็น 5 มโนทัศน์หลักที่สัมพันธ์กัน ได้แก่ 1) กระบวนการถ่ายทอดลักษณะทางพันธุกรรม 2) การแปรผันทางพันธุกรรม 3) เทคโนโลยีชีวภาพทางพันธุกรรม 4) ความหลากหลายทางชีวภาพ และ 5) วิวัฒนาการ (กระทรวงศึกษาธิการ, 2551)

แบบสอบวินิจฉัยเป็นแบบสอบที่ใช้ค้นหาความบกพร่องทางการเรียนของนักเรียนเป็นรายบุคคล ผลการตอบแบบสอบสามารถบอกได้ว่านักเรียนบกพร่องในทักษะใดและบกพร่องในจุดใด ผู้พัฒนาแบบสอบ

วินิจัยพยายามให้แบบสอบวินิจัยสามารถระบุสาเหตุของความบกพร่องนั้นได้อย่างถูกต้อง ข้อบกพร่องของนักเรียนอาจเป็นความบกพร่องที่เกิดจากนักเรียนเอง หรือเป็นข้อบกพร่องที่เกิดจากกระบวนการจัดการเรียนรู้ของครูผู้สอน (กระทรวงศึกษาธิการ, 2539) ผลการวินิจัยนำมาเพื่อการแก้ไขและส่งเสริมการเรียนรู้ของนักเรียนได้ถูกต้องและตรงจุด ตลอดจนปรับปรุงการสอนของครูให้มีประสิทธิภาพยิ่งขึ้น

จากการศึกษาที่ผ่านมาพบว่ามีแบบสอบวินิจัยแบบเลือกตอบสองระดับ สามระดับ และสี่ระดับ ซึ่งมีข้อดีสอดคล้องกันคือประหยัดเวลาในการวินิจัย การตรวจให้คะแนนเป็นปรนัย มีหลักฐานน่าเชื่อถือเกี่ยวกับความตรง และสามารถใช้สอบกับนักเรียนจำนวนมากได้พร้อม ๆ กัน ซึ่งสอดคล้องกับสภาพการทำงานของครูในปัจจุบัน แต่แบบสอบวินิจัยแบบเลือกตอบสองระดับมีข้อจำกัดคือไม่สามารถระบุกลุ่มนักเรียนที่ตอบผิดเพราะขาดความรู้ได้อย่างชัดเจน (Gurel, Eryilmaz, & McDermott, 2015) ซึ่งแบบสอบวินิจัยแบบเลือกตอบสามระดับสามารถจำแนกนักเรียนออกเป็นกลุ่มที่ตอบผิดออกเป็นตอบผิดเพราะมีมโนทัศน์ที่คลาดเคลื่อนและตอบผิดเพราะไม่มีความรู้ออกจากกันได้ แต่แบบสอบวินิจัยแบบเลือกตอบสี่ระดับสามารถจำแนกกลุ่มนักเรียนออกเป็นกลุ่มที่ตอบด้วยมโนทัศน์ที่ถูกต้อง ตอบผิดเพราะไม่มีความรู้ และตอบผิดด้วยมโนทัศน์ที่คลาดเคลื่อนออกจากกันได้ชัดเจนกว่าแบบสอบวินิจัยแบบเลือกตอบสามระดับ เนื่องจากการตอบความมั่นใจในการตอบของแบบสอบวินิจัยแบบเลือกตอบสามระดับ เป็นการตอบความมั่นใจในภาพรวมของการตอบในระดับคำตอบและระดับเหตุผล ส่วนการตอบระดับความมั่นใจในการตอบในแบบสอบวินิจัยแบบเลือกตอบสี่ระดับ เป็นการตอบแยกกันระหว่างความมั่นใจในการตอบระดับคำตอบ และความมั่นใจในการตอบระดับเหตุผล ทำให้ได้ข้อมูลการวินิจัยที่ละเอียดกว่าแบบสอบวินิจัยแบบเลือกตอบสามระดับ แต่แบบสอบวินิจัยแบบเลือกตอบสี่ระดับใช้เวลาสอบนานกว่าแบบสามระดับ ผู้วิจัยจึงสนใจที่จะเปรียบเทียบการวินิจัยมโนทัศน์ที่คลาดเคลื่อนด้วยแบบสอบวินิจัยระหว่างแบบสอบวินิจัยแบบเลือกตอบสี่ระดับกับสามระดับ เพื่อให้ได้ข้อสรุปเกี่ยวกับแนวทางการพัฒนาแบบสอบวินิจัยและผลการวินิจัยของแบบสอบวินิจัย และให้เกิดการนำไปใช้เพื่อวินิจัยมโนทัศน์ที่คลาดเคลื่อนได้อย่างมีประสิทธิภาพ และเหมาะสมกับสถานการณ์แนวคิดในการวิจัยครั้งนี้สามารถสรุปได้ดังแผนภาพ



แผนภาพกรอบแนวคิดการวิจัย

วัตถุประสงค์

1) เพื่อเปรียบเทียบผลการวินิจฉัยระหว่างแบบสอบวินิจฉัยแบบเลือกตอบสามระดับกับแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับวิชาชีววิทยาสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย

2) เพื่อวิเคราะห์ความสอดคล้องของผลการวินิจฉัยระหว่างแบบสอบวินิจฉัยแบบเลือกตอบสามระดับและแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับวิชาชีววิทยาสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลายกับข้อมูลเชิงประจักษ์จากการสัมภาษณ์

วิธีดำเนินการวิจัย

ประชากรและตัวอย่าง

ประชากรของการวิจัยครั้งนี้คือนักเรียนระดับมัธยมศึกษาตอนปลาย ปีการศึกษา 2560 จำนวน 12,875 คน จากโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 1 กรุงเทพมหานครที่มีผลการทดสอบระดับชาติขั้นพื้นฐานอยู่ในระดับปานกลาง จำนวน 12 โรงเรียน (ข้อมูลจากสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 1 เมื่อวันที่ 10 มิถุนายน 2560) สำหรับตัวอย่างในการวิจัยครั้งนี้คือนักเรียนระดับมัธยมศึกษาตอนปลายจากโรงเรียนราชวินิต มัธยม ได้มาด้วยวิธีการเลือกอย่างเจาะจง ที่มีระดับความสามารถแบ่งเป็นนักเรียนกลุ่มอ่อน กลุ่มปานกลาง และกลุ่มเก่ง รวมทั้งสิ้น จำนวน 62 คน

เครื่องมือที่ใช้เก็บรวบรวมข้อมูล

เครื่องมือที่ใช้ในการวิจัยครั้งนี้เป็นเครื่องมือที่มุ่งวัดความรู้ความเข้าใจของนักเรียนในโมทัศน์หลัก 5 โมทัศน์ ได้แก่ 1) การถ่ายทอดลักษณะทางพันธุกรรม 2) การแปรผันทางพันธุกรรม 3) เทคโนโลยีชีวภาพทางพันธุกรรม 4) ความหลากหลายทางชีวภาพ และ 5) วิวัฒนาการ โดยมีเครื่องมือวิจัยดังนี้

1) แบบสอบเพื่อสำรวจมโนทัศน์ที่คลาดเคลื่อนในวิชาชีววิทยา สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย เป็นข้อสอบแบบเขียนตอบจำนวน 36 ข้อ ข้อมูลที่ได้จากคำตอบของนักเรียนนำไปสู่การสร้างตัวเลือกและตัวลงในแบบสอบวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนแบบเลือกตอบสามระดับและสี่ระดับ

2) แบบสอบวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนแบบเลือกตอบสามระดับในวิชาชีววิทยา สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย ในแบบสอบประกอบด้วยข้อสอบจำนวน 33 ข้อ แต่ละข้อแบ่งการตอบเป็น 3 ระดับ ได้แก่ ระดับคำตอบ (Answer tier: A tier) เป็นข้อสอบแบบปรนัย 4 ตัวเลือก ระดับเหตุผล (Reason tier: R tier) เป็นข้อสอบแบบปรนัย 4 ตัวเลือก และ ระดับความมั่นใจ (Confidence tier: C tier) คือ การตอบระดับความมั่นใจในภาพรวมของการตอบข้อสอบข้อนั้น ๆ เป็นมาตราประมาณค่าแบบลิเคิร์ท 6 ระดับ

3) แบบสอบวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนแบบเลือกตอบสี่ระดับในวิชาชีววิทยาสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย ในแบบสอบประกอบด้วยข้อสอบ จำนวน 33 ข้อ แต่ละข้อแบ่งการตอบเป็น 4 ระดับ ได้แก่ ระดับคำตอบ (A tier) เป็นข้อสอบแบบปรนัย 4 ตัวเลือก ระดับความมั่นใจของการตอบระดับคำตอบ (Confidence of answer tier: CA tier) เป็นมาตราประมาณค่าแบบลิเคิร์ท 6 ระดับ ระดับเหตุผล (R tier) เป็นข้อสอบแบบปรนัย 4 ตัวเลือก และ ระดับความมั่นใจของการตอบระดับเหตุผล (Confidence of reason tier: CR tier) เป็นมาตราประมาณค่าแบบลิเคิร์ท 6 ระดับ

สำหรับคำถามและตัวเลือกในระดับคำตอบและระดับเหตุผลของแบบสอบวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนแบบเลือกตอบสามระดับและสี่ระดับเป็นคำถามและตัวเลือกชุดเดียวกัน แตกต่างกันที่ แบบสอบวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนแบบเลือกตอบสามระดับวัดความมั่นใจของการตอบในภาพรวม (C tier) ส่วนแบบสอบวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนแบบเลือกตอบสี่ระดับมีการวัดความมั่นใจของการตอบแยกเป็นความมั่นใจของการตอบระดับคำตอบ (CA tier) กับความมั่นใจของการตอบระดับเหตุผล (CR tier)

4) แบบสัมภาษณ์ แบบสัมภาษณ์นี้เป็นแบบสัมภาษณ์ที่ประยุกต์เทคนิคการคิดออกเสียงมาประกอบการสัมภาษณ์ ในแบบสัมภาษณ์มีคำถาม 15 ข้อ ที่ให้นักเรียนอธิบายความเข้าใจของนักเรียนเกี่ยวกับโมโนทัศน์หลักทั้ง 5 โมโนทัศน์ แล้ววินิจฉัยว่านักเรียนมีโมโนทัศน์ที่ถูกต้อง นักเรียนที่มีโมโนทัศน์ที่คลาดเคลื่อน และนักเรียนที่ไม่มีความรู้

ขั้นตอนการดำเนินการวิจัย

การดำเนินการวิจัยในครั้งนี้เริ่มจากการกำหนดขอบเขตของเนื้อหาและโมโนทัศน์ที่ต้องการวัด ด้วยการศึกษาเนื้อหาตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 จากนั้นกำหนดโมโนทัศน์ที่ต้องการวัดจากขอบข่ายเนื้อหา นำไปสู่การกำหนดวัตถุประสงค์เชิงพฤติกรรม

ขั้นตอนพัฒนาเครื่องมือวิจัยเริ่มจากการสร้างแบบสอบเพื่อสำรวจรูปแบบโมโนทัศน์ที่คลาดเคลื่อน ซึ่งเป็นคำถามแบบเขียนตอบ เพื่อนำคำตอบของนักเรียนมาใช้พัฒนาเป็นตัวเลือกในแบบสอบวินิจฉัยโมโนทัศน์ที่คลาดเคลื่อนแบบเลือกตอบสามระดับและสี่ระดับ

ลักษณะของแบบสอบวินิจฉัยแบบเลือกตอบสามระดับ ประกอบด้วยข้อสอบจำนวน 33 ข้อ แต่ละข้อประกอบด้วยสามส่วนได้แก่ ส่วนที่ 1 คำถามระดับความรู้หรือระดับคำตอบ (Answer tier: A tier) เป็นข้อสอบแบบปรนัย 4 ตัวเลือก ซึ่งนักเรียนจะต้องตอบความรู้ของนักเรียนที่เป็นคำตอบของคำถามข้อนั้น ส่วนที่ 2 การตอบระดับเหตุผล (Reason tier: R tier) เป็นข้อสอบแบบปรนัย 4 ตัวเลือก การตอบระดับเหตุผลนี้เป็นการระบุเหตุผลของนักเรียนที่อยู่เบื้องหลังการตอบในระดับคำตอบ (A tier) และส่วนที่ 3 การตอบระดับความมั่นใจ (Confidence tier: C tier) เป็นการระบุระดับความมั่นใจในภาพรวมของนักเรียนว่านักเรียนมั่นใจมากน้อยเพียงใดว่าการตอบระดับคำตอบและระดับเหตุผลของนักเรียนเป็นคำตอบที่ถูกต้อง ซึ่งเป็นมาตราประมาณค่า (Rating scale) แบบ 6 ระดับ

สำหรับแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับ ประกอบด้วยข้อสอบจำนวน 33 ข้อ แต่ละข้อประกอบด้วย 4 ส่วนได้แก่ ส่วนที่ 1 คำถามระดับความรู้หรือระดับคำตอบ (A tier) เป็นข้อสอบแบบปรนัย 4 ตัวเลือกเช่นเดียวกับแบบสอบวินิจฉัยแบบเลือกตอบสามระดับ ส่วนที่ 2 การตอบระดับความมั่นใจของระดับคำตอบ (Confidence of answer tier: CA tier) เป็นการระบุระดับความมั่นใจว่าการตอบระดับคำตอบของนักเรียนเป็นคำตอบที่ถูกต้อง ซึ่งเป็นมาตราประมาณค่า (Rating scale) แบบ 6 ระดับ ส่วนที่ 3 เป็นการตอบระดับเหตุผล (R tier) เช่นเดียวกับ ส่วนที่ 2 ของแบบสอบวินิจฉัยแบบเลือกตอบสามระดับ ส่วนที่ 4 การตอบระดับความมั่นใจของระดับเหตุผล (Confidence of reason tier: CR tier) เป็นการระบุระดับความมั่นใจว่าการตอบระดับเหตุผลของนักเรียนเป็นคำตอบที่ถูกต้อง ซึ่งเป็นมาตราประมาณค่า (Rating scale) แบบ 6 ระดับ

จากนั้นเป็นการตรวจสอบคุณภาพของแบบสอบวินิจฉัยแบบเลือกตอบด้วยการพิจารณาค่าดัชนีความสอดคล้องของวัตถุประสงค์เชิงพฤติกรรมกับเนื้อหาของข้อสอบ (Index of item-objective congruence: IOC) จากการพิจารณาของผู้ทรงคุณวุฒิ การทดลองเครื่องมือเพื่อตรวจสอบค่าความยากและอำนาจจำแนก จากนั้นปรับปรุงให้มีคุณภาพเป็นไปตามเกณฑ์ แล้วจึงนำแบบสอบวินิจฉัยไปใช้เก็บข้อมูลโดยแบ่งตัวอย่างออกเป็น 2 กลุ่มย่อย ในการวินิจฉัยรอบแรก กลุ่มย่อยที่ 1 วินิจฉัยด้วยแบบสอบวินิจฉัยแบบเลือกตอบสามระดับ กลุ่มย่อยที่ 2 วินิจฉัยด้วยแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับ แล้วเว้นระยะเวลา 2 สัปดาห์ แล้ววินิจฉัยรอบที่ 2 โดยกลุ่มย่อยที่ 1 วินิจฉัยด้วยแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับ และกลุ่มย่อยที่ 2 วินิจฉัยด้วยแบบสอบวินิจฉัยแบบเลือกตอบสามระดับ แล้วจึงสัมภาษณ์นักเรียนด้วยเทคนิคการคิดออกเสียง

คุณภาพของแบบสอบวินิจฉัยโมโนทัศน์ที่คลาดเคลื่อนแบบเลือกตอบ สำหรับความยากและอำนาจจำแนกของแบบสอบมีรายละเอียดดังตารางที่ 1

ตารางที่ 1 ค่าเฉลี่ยของค่าความยากและอำนาจจำแนกของแบบสอบวินิจฉัยแบบเลือกตอบสามระดับและสี่ระดับ

ชนิดของแบบสอบ	ค่าความยากเฉลี่ย		ค่าอำนาจจำแนกเฉลี่ย	
	A tier	R tier	A tier	R tier
แบบสอบวินิจฉัยแบบเลือกตอบสามระดับ	0.46	0.43	0.37	0.32
แบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับ	0.44	0.44	0.35	0.34

จากตารางที่ 1 พบว่า แบบสอบวินิจฉัยแบบเลือกตอบสามระดับ ระดับคำตอบ (A tier) มีค่าความยากเฉลี่ยเท่ากับ 0.46 แสดงว่า โดยเฉลี่ยของระดับคำตอบแบบสอบมีความยากระดับปานกลาง ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 0.43 แสดงว่า โดยเฉลี่ยระดับคำตอบของแบบสอบสามารถจำแนกนักเรียนได้ดี ระดับเหตุผล (R tier) มีค่าความยากเฉลี่ยเท่ากับ 0.43 แสดงว่าโดยเฉลี่ยระดับเหตุผลของแบบสอบมีความยากระดับปานกลาง ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 0.32 แสดงว่า โดยเฉลี่ยระดับเหตุผลของแบบสอบสามารถจำแนกนักเรียนได้ดี ส่วนแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับ ระดับคำตอบ (A tier) มีค่าความยากเฉลี่ยเท่ากับ 0.44 แสดงว่าโดยเฉลี่ยระดับคำตอบของแบบสอบมีความยากระดับปานกลาง ค่าอำนาจจำแนกเฉลี่ยมีค่าเท่ากับ 0.35 แสดงว่า โดยเฉลี่ยระดับคำตอบของแบบสอบสามารถจำแนกนักเรียนได้ดี ระดับเหตุผล (R tier) มีค่าความยากเฉลี่ยเท่ากับ 0.44 แสดงว่า โดยเฉลี่ยระดับเหตุผลของแบบสอบมีความยากระดับปานกลาง ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 0.34 แสดงว่า โดยเฉลี่ยระดับเหตุผลของแบบสอบสามารถจำแนกนักเรียนได้ดี

ความเที่ยงของแบบสอบซึ่งเป็นการพิจารณาความเที่ยงแบบความสอดคล้องภายใน ด้วยการพิจารณาค่าสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's alpha coefficient) แบบสอบวินิจฉัยแบบเลือกตอบสามระดับและสี่ระดับ มีค่าสัมประสิทธิ์แอลฟาของครอนบาคในภาพรวมทั้งฉบับเท่ากับ 0.83 และ 0.84 ตามลำดับ แสดงว่าในภาพรวมแบบสอบทั้ง 2 ฉบับมีความเที่ยงสูง

การเปรียบเทียบผลการวินิจฉัยแบ่งออกเป็น 2 ส่วนได้แก่ ส่วนที่ 1 ผลการเปรียบเทียบการวินิจฉัยระหว่างแบบสอบวินิจฉัยแบบเลือกตอบสามระดับกับแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับในวิชาชีววิทยา สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย ด้วยการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันและค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างชั้นระหว่างคะแนนของนักเรียนที่ได้จากแบบสอบวินิจฉัยแบบเลือกตอบสามระดับและสี่ระดับ จากนั้นวิเคราะห์ความสอดคล้องของผลการวินิจฉัยด้วยแบบสอบวินิจฉัยแบบเลือกตอบสามระดับกับสี่ระดับ ด้วยการวิเคราะห์ค่าสหสัมพันธ์วีของคราเมอร์ และส่วนที่ 2 การวิเคราะห์ความสอดคล้องระหว่างผลการวินิจฉัยจากแบบสอบวินิจฉัยแบบเลือกตอบสามระดับและสี่ระดับกับข้อมูลเชิงประจักษ์จากการสัมภาษณ์ ด้วยการวิเคราะห์ค่าสหสัมพันธ์วีของคราเมอร์ และเปรียบเทียบค่าสัมประสิทธิ์สหสัมพันธ์ของตัวอย่างที่ทับซ้อนกัน (Comparing correlation coefficients of overlapping samples)

ผลการวิจัย

ผลการวินิจฉัยของแบบสอบวินิจฉัยแบบเลือกตอบสามระดับและสี่ระดับ สามารถจำแนกนักเรียนออกเป็น 6 กลุ่มได้แก่ 1) นักเรียนที่มีมโนทัศน์ที่ถูกต้อง (Correct concept: CC) 2) นักเรียนที่มีมโนทัศน์ที่

คลาดเคลื่อนแบบที่ 1 (False positive: FP) 3) นักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 2 (False negative: FN) 4) นักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 3 (Misconception: MC) 5) นักเรียนที่ตอบถูกเพราะการเดา (Lucky guess: LG) และ 6) ไม่มีความรู้ (Lack of knowledge: LK) ผลการวินิจฉัยเป็นดังตารางที่ 2

ตารางที่ 2 ร้อยละของนักเรียนจำแนกกลุ่มตามผลการวินิจฉัยด้วยแบบสอบวินิจฉัยแบบเลือกตอบสามระดับและสี่ระดับ

ผลการวินิจฉัย	แบบสอบวินิจฉัยแบบเลือกตอบสามระดับ		แบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับ	
	จำนวน (คน)	ร้อยละ	จำนวน (คน)	ร้อยละ
1) นักเรียนที่มีมโนทัศน์ที่ถูกต้อง (CC)	17	27.43	17	27.43
2) นักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 1 (FP)	1	1.61	0	0.00
3) นักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 2 (FN)	1	1.61	1	1.61
4) นักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 3 (MC)	16	25.81	11	17.73
5) นักเรียนที่ตอบถูกด้วยการเดา (LG)	1	1.61	0	0.00
6) นักเรียนที่ไม่มีความรู้ (LK)	26	40.94	33	53.23
รวม	62	100.00	62	100.00

จากตารางที่ 2 แบบสอบวินิจฉัยแบบเลือกตอบสามระดับพบว่า ร้อยละของจำนวนนักเรียนที่มีมโนทัศน์ที่ถูกต้อง (CC) เท่ากับ 27.43 ร้อยละของจำนวนนักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 1 (FP) เท่ากับ 1.61 ร้อยละของจำนวนนักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 2 (FN) เท่ากับ 1.61 ร้อยละของจำนวนนักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 3 (MC) เท่ากับ 25.81 ร้อยละของจำนวนนักเรียนที่ตอบถูกด้วยการเดา (LG) เท่ากับ 1.61 และร้อยละของจำนวนนักเรียนที่ไม่มีความรู้ (LK) เท่ากับ 40.94 ส่วนแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับพบว่า ร้อยละของจำนวนนักเรียนที่มีมโนทัศน์ที่ถูกต้อง (CC) เท่ากับ 27.43 ร้อยละของจำนวนนักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 1 (FP) เท่ากับ 0.00 ร้อยละของจำนวนนักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 2 (FN) เท่ากับ 1.61 ร้อยละของจำนวนนักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 3 (MC) เท่ากับ 17.73 ร้อยละของจำนวนนักเรียนที่ตอบถูกด้วยการเดา (LG) เท่ากับ 0.00 และร้อยละของจำนวนนักเรียนที่ไม่มีความรู้ (LK) เท่ากับ 53.23

ผลการเปรียบเทียบผลการวินิจฉัยจากแบบสอบวินิจฉัยแบบเลือกตอบสามระดับกับสี่ระดับด้วยการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ของคราเมอร์ ปรากฏผลดังตารางที่ 3

ตารางที่ 3 ค่าสัมประสิทธิ์สหสัมพันธ์วีของคราเมอร์ระหว่างผลทดสอบจากแบบสอบถามวิจัยแบบเลือกตอบสามระดับกับสี่ระดับ

หัวข้อที่พิจารณา	ค่าสัมประสิทธิ์สหสัมพันธ์ วีของคราเมอร์	การแปลผล
1) มโนทัศน์หลักเรื่องการถ่ายทอด ลักษณะทางพันธุกรรม	0.339	สัมพันธ์กันระดับต่ำ
2) มโนทัศน์หลักเรื่องการแปรผันทาง พันธุกรรม	0.322	สัมพันธ์กันระดับต่ำ
3) มโนทัศน์หลักเรื่อง เทคโนโลยีชีวภาพทางพันธุกรรม	0.336	สัมพันธ์กันระดับต่ำ
4) มโนทัศน์หลักเรื่องความหลากหลาย ทางชีวภาพ	0.489	สัมพันธ์กันระดับปานกลาง
5) มโนทัศน์หลักเรื่องวิวัฒนาการ	0.405	สัมพันธ์กันระดับปานกลาง
6) ในภาพรวมทั้งฉบับ	0.536	สัมพันธ์กันระดับปานกลาง
เฉลี่ย	0.405	สัมพันธ์กันระดับปานกลาง

จากตารางที่ 3 พบว่ามโนทัศน์หลักที่มีความสอดคล้องในระดับต่ำได้แก่ 1) มโนทัศน์หลักเรื่องการถ่ายทอดลักษณะทางพันธุกรรม 2) มโนทัศน์หลักเรื่องการแปรผันทางพันธุกรรม และ 3) มโนทัศน์หลักเรื่องเทคโนโลยีชีวภาพทางพันธุกรรม มโนทัศน์ที่มีความสอดคล้องในระดับปานกลางได้แก่ 1) มโนทัศน์หลักเรื่องความหลากหลายทางชีวภาพ และ 2) มโนทัศน์หลักเรื่องวิวัฒนาการ เมื่อพิจารณาผลการวิจัยในภาพรวมทั้งฉบับพบว่า ผลวิจัยด้วยแบบสอบถามวิจัยทั้งสองฉบับมีความสอดคล้องกันในระดับปานกลาง

เมื่อเปรียบเทียบผลความสอดคล้องของการวิจัยด้วยแบบสอบถามวิจัยแบบเลือกตอบสามระดับและสี่ระดับกับข้อมูลเชิงประจักษ์จากการสัมภาษณ์ ปรากฏผลดังตารางที่ 4

ตารางที่ 4 สัมประสิทธิ์สหสัมพันธ์วีของคราเมอร์จากการเปรียบเทียบผลการวิจัยด้วยแบบสอบถามวิจัยกับผลการสัมภาษณ์

การเปรียบเทียบ	Cramer's V				ความแตกต่าง ของค่า Cramer's V
	สามระดับ	ระดับ ความสัมพันธ์	สี่ระดับ	ระดับ ความสัมพันธ์	
1) การถ่ายทอดลักษณะ ทางพันธุกรรม	0.613	สูง	0.588	ปานกลาง	0.025 (p=.801)
2) การแปรผันทาง พันธุกรรม	0.456	ปานกลาง	0.493	ปานกลาง	0.037 (p=.772)
3) เทคโนโลยีชีวภาพทาง พันธุกรรม	0.424	ปานกลาง	0.609	สูง	0.185 (p=.121)

ตารางที่ 4 สัมประสิทธิ์สหสัมพันธ์วีของคราเมอร์จากการเปรียบเทียบผลการวินิจฉัยด้วยแบบสอบวินิจฉัยกับผลการสัมภาษณ์ (ต่อ)

การเปรียบเทียบ	Cramer's V				ความแตกต่าง
	สามระดับ	ระดับความสัมพันธ์	สี่ระดับ	ระดับความสัมพันธ์	ของค่า Cramer's V
4) ความหลากหลายทางชีวภาพ	0.515	ปานกลาง	0.572	ปานกลาง	0.057 (p=.585)
5) วิวัฒนาการ	0.320	ต่ำ	0.485	ปานกลาง	0.165 (p=.189)
ในภาพรวมทั้งฉบับ	0.569	ปานกลาง	0.558	ปานกลาง	0.011 (p=.905)

จากตารางที่ 4 พบว่า เมื่อเปรียบเทียบผลการวินิจฉัยด้วยแบบสอบวินิจฉัยกับผลการสัมภาษณ์แล้ว ค่าสัมประสิทธิ์สหสัมพันธ์วีของคราเมอร์จากแบบสอบวินิจฉัยแบบเลือกตอบสามระดับมีค่าระหว่าง 0.320 ถึง 0.613 มีระดับความสัมพันธ์กับข้อมูลเชิงประจักษ์จากการสัมภาษณ์ในระดับต่ำถึงสูง ส่วนใหญ่สัมพันธ์ในระดับปานกลาง สำหรับแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับมีค่าระหว่าง 0.485 ถึง 0.609 มีระดับความสัมพันธ์กับข้อมูลเชิงประจักษ์จากการสัมภาษณ์ในระดับปานกลางถึงสูง ส่วนใหญ่อยู่ในระดับปานกลาง แต่เมื่อทดสอบความแตกต่างระหว่างค่าสัมประสิทธิ์สหสัมพันธ์วีของคราเมอร์ในรายมน็อตส์หลักและในภาพรวมทั้งฉบับพบว่า ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

กล่าวโดยสรุปแบบสอบวินิจฉัยแบบเลือกตอบสามระดับกับสี่ระดับ เป็นแบบสอบที่มุ่งวัดคุณลักษณะเดียวกัน แต่แบบสอบวินิจฉัยแบบเลือกตอบสามระดับกับสี่ระดับให้ผลการวินิจฉัยที่สอดคล้องกันในระดับต่ำถึงปานกลาง และแบบทั้งสองให้ผลการวินิจฉัยที่สอดคล้องกับข้อมูลเชิงประจักษ์จากการสัมภาษณ์ไม่แตกต่างกัน

อภิปรายผล

การวินิจฉัยมน็อตส์ที่คลาดเคลื่อนด้วยแบบสอบวินิจฉัยแบบเลือกตอบสามระดับและสี่ระดับในการวิจัยในครั้งนี้มุ่งวินิจฉัยในวิชาชีววิทยา แบ่งเนื้อหาออกเป็น 3 เรื่อง ได้แก่ เรื่องที่ 1) พันธุกรรม ประกอบด้วยมน็อตส์หลัก 3 มน็อตส์ได้แก่ 1.1) การถ่ายทอดลักษณะทางพันธุกรรม 1.2) การแปรผันทางพันธุกรรม 1.3) เทคโนโลยีชีวภาพทางพันธุกรรม เรื่องที่ 2) ความหลากหลายทางชีวภาพ ประกอบด้วยมน็อตส์หลักคือ 2.1) ความหลากหลายทางชีวภาพ และเรื่องที่ 3) วิวัฒนาการ ประกอบด้วยมน็อตส์หลักคือ 3.1) วิวัฒนาการรวมมน็อตส์หลักทั้งสิ้น 5 มน็อตส์สอดคล้องกับแนวคิดเรื่องลักษณะทั่วไปของแบบสอบวินิจฉัยของวนิดา ภู่อี่ยม (2550) ที่กล่าวว่ามีการวิเคราะห์เนื้อหาออกเป็นมน็อตส์ที่ต้องการวินิจฉัย การวินิจฉัยด้วยแบบสอบวินิจฉัยจัดเป็นการวินิจฉัยแบบเป็นทางการ (Formal diagnostic) (ศิริเดช สุชีวะ, 2538) และเป็นการวินิจฉัยระดับวิเคราะห์ (Analytical diagnostic) (สุปราณี บุระ, 2557)

เมื่อพิจารณาค่าความยาก (Difficulty) ของข้อสอบตามหลักการของการสร้างแบบสอบวินิจฉัยแบบเลือกตอบระบุว่า แบบสอบวินิจฉัยแบบเลือกตอบเป็นแบบสอบที่มีความยากในระดับปานกลางถึงค่อนข้างง่าย (วนิดา ภู่อี่ยม, 2550; วนิดา ช่อนคำ, 2551) สำหรับการวิจัยครั้งนี้พบว่า แบบสอบวินิจฉัยแบบเลือกตอบสามระดับ ในภาพรวมระดับคำตอบ (A tier) มีค่าความยากเฉลี่ย 0.46 ระดับเหตุผล (R tier) มีค่าความยากเฉลี่ย 0.43 ส่วนแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับ ในภาพรวมระดับคำตอบ (A tier) มีค่าความยากเฉลี่ย 0.44

ระดับเหตุผล (R tier) มีค่าความยากเฉลี่ย 0.44 ในภาพรวมแบบสอบวินิจัยแบบเลือกตอบสามระดับและสี่ระดับมีค่าความยากในระดับปานกลาง เมื่อพิจารณาค่าอำนาจจำแนก (Discriminant) ของข้อสอบ แบบสอบวินิจัยแบบเลือกตอบสามระดับ ในภาพรวมระดับคำตอบ (A tier) มีค่าอำนาจจำแนกเฉลี่ย 0.37 ระดับเหตุผล (R tier) มีค่าอำนาจจำแนกเฉลี่ย 0.32 ส่วนแบบสอบวินิจัยแบบเลือกตอบสี่ระดับในภาพรวมระดับคำตอบ (A tier) มีค่าอำนาจจำแนกเฉลี่ย 0.35 ระดับเหตุผล (R tier) มีค่าอำนาจจำแนกเฉลี่ย 0.34 แบบสอบวินิจัยแบบเลือกตอบสามระดับและสี่ระดับมีอำนาจจำแนกอยู่ในภาพรวมระดับพอใช้ ทั้งระดับคำตอบและระดับเหตุผล ซึ่งงานวิจัยนี้มีค่าความยากอยู่ในระดับปานกลาง อาจเกิดจากตัวอย่างในการวิจัยที่ส่วนใหญ่เป็นนักเรียนที่มีความรู้ในระดับปานกลางถึงอ่อน งานวิจัยอื่นที่มีค่าความยากสอดคล้องกับผลการวิจัยครั้งนี้ เช่น การวิจัยเรื่องการพัฒนาและการนำไปใช้ของแบบสอบวินิจัยแบบสองระดับสำหรับนักเรียนมัธยมศึกษาตอนปลายเพื่อวัดความเข้าใจเกี่ยวกับการเติบโตและการเจริญของพืชดอก ของ Lin Sheau-Wen (2004) ที่มีค่าความยากเฉลี่ย 0.53 เป็นค่าความยากในระดับปานกลาง (Lin Sheau-Wen, 2004) การพัฒนาและความตรงของแบบสอบวินิจัยสองระดับเพื่อวัดความเข้าใจเกี่ยวกับการลำเลียงในพืช และระบบหมุนเวียนเลือดในมนุษย์ ของ Wang Jing-Ru (2004) มีค่าความยากเฉลี่ย 0.40 เป็นความยากระดับปานกลาง (Wang Jing-Ru, 2004) สำหรับค่าอำนาจจำแนกของแบบสอบวินิจัยในการวิจัยครั้งนี้อยู่ในระดับดี สอดคล้องกับงานวิจัยเรื่องการพัฒนาและความตรงของแบบสอบวินิจัยสองระดับเพื่อวัดความเข้าใจเกี่ยวกับการลำเลียง ในพืช และระบบหมุนเวียนเลือดในมนุษย์ ของ Wang Jing-Ru (2004) มีค่าอำนาจจำแนกเฉลี่ย 0.37 อยู่ในระดับดี การใช้แบบสอบวินิจัยสี่ระดับเพื่อวัดธรรมชาติของมโนทัศน์ที่คลาดเคลื่อนของนักเรียนในเรื่องคลื่น ของ Caleon and Subramaniam (2010) มีค่าอำนาจจำแนกเฉลี่ย 0.38 อยู่ในระดับพอใช้ (Caleon & Subramaniam, 2010) จากข้อมูลดังกล่าวข้างต้นสอดคล้องกับข้อสรุปของ Gurel, Eryilmaz and McDermott (2015) ที่ระบุว่า การสร้างแบบสอบวินิจัยให้มีคุณภาพสูงเป็นวิธีการที่ทำได้ยาก (Gurel et al., 2015)

การตรวจสอบความเที่ยงของแบบสอบ พิจารณาความเที่ยง (Reliability) แบบความสอดคล้องภายใน ด้วยการคำนวณค่าสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's alpha coefficient) เมื่อพิจารณาค่าความเที่ยงในภาพรวมทั้งฉบับพบว่า แบบสอบวินิจัยแบบเลือกตอบสามระดับและสี่ระดับมีค่าสัมประสิทธิ์แอลฟาของครอนบาคเป็น 0.83 และ 0.84 ตามลำดับ หมายความว่าแบบสอบวินิจัยทั้งสองฉบับมีค่าความเที่ยงอยู่ในระดับสูง สอดคล้องกับการวิจัยเรื่องระดับความเชี่ยวชาญและระดับความมั่นใจของนักเรียนมัธยมศึกษาตอนปลายเกี่ยวกับมโนทัศน์เบื้องต้นของอิเล็กทรอนิกส์ ของ Sia Ding Teng and others (2012) ที่พิจารณาค่าความเที่ยงด้วยการคำนวณค่าสัมประสิทธิ์แอลฟาของครอนบาค มีค่า 0.85 มีความเที่ยงของแบบสอบอยู่ในระดับสูง และการประเมินภาคปฏิบัติในเด็กอายุ 10 – 11 ปี (Sia Ding Teng, Treagust, & Chandrasegaran, 2012) และมโนทัศน์ที่คลาดเคลื่อนเกี่ยวกับสำนักด้านตัวเลขโดยใช้แบบสอบวินิจัยสี่ระดับ ของ Yang Der-Ching and Lin Yang-Chi (2015) พิจารณาค่าความเที่ยงด้วยการคำนวณค่าสัมประสิทธิ์แอลฟาของครอนบาค มีค่า 0.98 มีความเที่ยงของแบบสอบอยู่ในระดับสูง (Yang Der-Ching & Lin Yung-Chi, 2015) และจากข้อมูลความเที่ยงของแบบสอบวินิจัยแบบเลือกตอบสามระดับและสี่ระดับมีค่าใกล้เคียงกัน อาจเกิดจากแบบสอบวินิจัยทั้งสองฉบับมีการวัดในเนื้อหาและมโนทัศน์เดียวกัน โดยเฉพาะอย่างยิ่งคำถามและตัวเลือกในระดับคำตอบและระดับเหตุผลของแบบสอบทั้งสองฉบับเป็นคำถามและตัวเลือกชุดเดียวกัน จึงทำให้ค่าสัมประสิทธิ์แอลฟาของครอนบาคของแบบสอบทั้งสองฉบับมีค่าใกล้เคียงกัน สอดคล้องกับงานวิจัยของมนัสสิริ อินทร์สวาท (2559) ที่พบว่าแบบสอบวินิจัยแบบเลือกตอบสามระดับจำนวน 2 ฉบับที่มีคำถามและตัวเลือกระดับคำตอบและระดับเหตุผลเหมือนกัน แต่ใช้ระดับการวัดความมั่นใจแตกต่างกัน มีค่าสัมประสิทธิ์แอลฟาของครอนบาคของแบบสอบสองฉบับไม่แตกต่างกัน

เมื่อพิจารณาความสอดคล้องของแบบสอบทั้งสองฉบับด้วยการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intraclass correlation) พบว่ามีค่าเท่ากับ 0.841 หมายความว่าแบบสอบวินิจฉัยแบบเลือกตอบสามารถบ่งชี้กับแบบสอบวินิจฉัยแบบเลือกตอบที่ระดับมีความสัมพันธ์ในทิศทางเดียวกันในระดับสูง เป็นอีกหลักฐานหนึ่งที่สนับสนุนว่า แบบสอบวินิจฉัยทั้งสองฉบับมีการวัดในเนื้อหาและมิติที่เหมือนกัน โดยเฉพาะอย่างยิ่งคำถามและตัวเลือกในระดับคำตอบและระดับเหตุผลของแบบสอบทั้งสองฉบับเป็นคำถามและตัวเลือกชุดเดียวกัน จึงทำให้ค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้นมีความสัมพันธ์กันสูง

จากรูปแบบการตอบของนักเรียนในแบบสอบวินิจฉัยแบบเลือกตอบสามระดับและสี่ระดับ สามารถจำแนกนักเรียนออกเป็น 6 กลุ่ม สอดคล้องกับการศึกษาของ Gurel, Eryilmaz, and McDermott (2015) ได้แก่ 1) นักเรียนที่มีมโนทัศน์ที่ถูกต้อง (Correct concept: CC) 2) นักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 1 (False positive: FP) 3) นักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 2 (False negative: FN) 4) นักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 3 (Misconception: MC) 5) นักเรียนที่ตอบถูกด้วยการเดา (Lucky guess: LG) และ 6) นักเรียนที่ไม่มีความรู้ (Lack of knowledge: LK) (Gurel et al., 2015) สำหรับผลจากการวินิจฉัยด้วยแบบสอบสามระดับพบว่า มีผลวินิจฉัยใน 4 กลุ่ม เรียงลำดับจากมากไปน้อย ได้แก่ 1) นักเรียนที่ไม่มีความรู้ (LK) ร้อยละ 40.94 2) นักเรียนที่มีมโนทัศน์ที่ถูกต้อง (CC) ร้อยละ 27.43 3) นักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 3 (MC) ร้อยละ 25.81 4) นักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 1 (FP) แบบที่ 2 (FN) และ นักเรียนที่ตอบถูกด้วยการเดา (LG) แต่ละกลุ่มมีค่าร้อยละเท่ากับ 1.61 ส่วนแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับมีผลวินิจฉัยใน 6 กลุ่ม เรียงลำดับจากมากไปน้อย ได้แก่ 1) นักเรียนที่ไม่มีความรู้ (LK) ร้อยละ 53.23 2) นักเรียนที่มีมโนทัศน์ที่ถูกต้อง (CC) ร้อยละ 27.43 3) นักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 3 (MC) ร้อยละ 17.73 และ 4) นักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 2 (FN) ร้อยละ 1.61 ส่วนนักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อนแบบที่ 1 (FP) และนักเรียนที่ตอบถูกด้วยการเดา (LG) ไม่พบในผลการทดสอบด้วยแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับ สอดคล้องกับผลการศึกษาของ Arshlan (2012) ที่ระบุว่า ความน่าจะเป็นของการตอบแบบมโนทัศน์ที่คลาดเคลื่อนแบบที่ 1 (FP) และแบบมโนทัศน์ที่คลาดเคลื่อนแบบที่ 2 (FN) ควรจะน้อยกว่าร้อยละ 10 แบบสอบวินิจฉัยที่มีค่าร้อยละของการตอบแบบมโนทัศน์ที่คลาดเคลื่อนแบบที่ 1 (FP) และการตอบแบบมโนทัศน์ที่คลาดเคลื่อนแบบที่ 2 (FN) ต่ำ แสดงว่ามีค่าความตรงเชิงเนื้อหาในระดับสูง (Arslan, Cigdemoglu, & Moseley, 2012) จากข้อมูลดังกล่าวเป็นหลักฐานสนับสนุนว่า แบบสอบวินิจฉัยแบบเลือกตอบสามระดับและสี่ระดับในการวิจัยนี้เป็นแบบสอบที่มีความตรงเชิงเนื้อหา

ผลการเปรียบเทียบการวินิจฉัยของแบบสอบวินิจฉัยแบบเลือกตอบสามระดับกับสี่ระดับในประเด็นความสัมพันธ์ของการตอบระดับคำตอบ (Answer tier) ระดับเหตุผล (Reason tier) และระดับความมั่นใจ (Confidence tier) โดยการพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson's product moment correlation coefficient) พบว่า ในแบบสอบวินิจฉัยแบบเลือกตอบสามระดับ มีความสัมพันธ์ระหว่างการตอบระดับคำตอบ (A tier) กับการตอบระดับเหตุผล (R tier) มีค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันเท่ากับ 0.782 หมายความว่า การตอบระดับคำตอบกับการตอบระดับเหตุผลของนักเรียนมีความสัมพันธ์สอดคล้องไปในทิศทางเดียวกันในระดับสูงอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เมื่อพิจารณาความสัมพันธ์ระหว่างการตอบระดับคำตอบ (A tier) กับการตอบระดับความมั่นใจ (C tier) พบว่ามีค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันเท่ากับ 0.316 หมายความว่า การตอบระดับคำตอบกับการตอบระดับเหตุผลของนักเรียนมีความสัมพันธ์ไปในทิศทางเดียวกันในระดับต่ำอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และเมื่อพิจารณาความสัมพันธ์ระหว่างการตอบระดับเหตุผล (R tier) กับการตอบระดับความมั่นใจ (C tier) พบว่ามีค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันเท่ากับ 0.471 หมายความว่า การตอบระดับคำตอบกับการตอบระดับเหตุผลของนักเรียนมี

ความสัมพันธ์ไปในทิศทางเดียวกันในระดับปานกลางอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนในแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับ มีความสัมพันธ์ของการตอบระดับคำตอบ (A tier) กับการตอบระดับเหตุผล (R tier) มีค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันเท่ากับ 0.816 หมายความว่า การตอบระดับคำตอบกับการตอบระดับเหตุผลของนักเรียนมีความสัมพันธ์ไปในทิศทางเดียวกันในระดับสูงมากอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เมื่อพิจารณาความสัมพันธ์ของการตอบระดับคำตอบ (A tier) กับการตอบระดับความมั่นใจของระดับคำตอบ (CA tier) มีค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันเท่ากับ 0.509 หมายความว่า การตอบระดับคำตอบกับการตอบระดับความมั่นใจของระดับคำตอบของนักเรียนมีความสัมพันธ์ไปในทิศทางเดียวกันในระดับปานกลาง อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เมื่อพิจารณาความสัมพันธ์ของการตอบระดับเหตุผล (R tier) กับการตอบระดับความมั่นใจของระดับเหตุผล (CR tier) มีค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันเท่ากับ 0.417 หมายความว่า การตอบระดับเหตุผลกับการตอบระดับความมั่นใจของระดับเหตุผลของนักเรียนมีความสัมพันธ์ในทิศทางเดียวกันในระดับปานกลางอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และเมื่อพิจารณาความสัมพันธ์ของการตอบระดับความมั่นใจของระดับคำตอบ (CA tier) กับการตอบระดับความมั่นใจของระดับเหตุผล (CR tier) พบว่า มีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.947 หมายความว่า การตอบระดับความมั่นใจของระดับคำตอบกับการตอบระดับความมั่นใจของระดับเหตุผลมีความสัมพันธ์ในทิศทางเดียวกันในระดับสูงมากอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จากข้อมูลดังกล่าวสรุปได้ว่า แบบสอบทั้งสองฉบับมีความสัมพันธ์ระหว่างการตอบระดับคำตอบ (Answer tier) ระดับเหตุผล (Reason tier) และระดับความมั่นใจ (Confidence tier) อยู่จริง และเป็นความสัมพันธ์ทางบวก สอดคล้องกับการศึกษาของ Caleon and Subramaniam (2010) ที่พบว่าผลการวัดระดับความมั่นใจจะสอดคล้องกับคำตอบ โดยที่ระดับความมั่นใจจะสูง เมื่อคำตอบนั้นเป็นคำตอบที่ถูกต้อง และระดับความมั่นใจจะต่ำลงเมื่อคำตอบข้อนั้นเป็นคำตอบที่ผิด (Caleon & Subramaniam, 2010)

เมื่อพิจารณาความสอดคล้องของการวินิจฉัยด้วยแบบสอบวินิจฉัยแบบเลือกตอบสามระดับกับสี่ระดับ ด้วยการวิเคราะห์ความสัมพันธ์สหสัมพันธ์วีของคราเมอร์ (Cramer's V correlation coefficient) ในราย McNamara's test ทั้ง 5 McNamara's test และในภาพรวม พบว่า ส่วนใหญ่มีความสอดคล้องกันในระดับต่ำถึงปานกลาง แสดงว่าโอกาสที่นักเรียนคนหนึ่งจะถูกจัดกลุ่มอยู่ในกลุ่มเดียวกันทั้งจากการวินิจฉัยด้วยแบบสอบวินิจฉัยแบบเลือกตอบสามระดับและสี่ระดับมีความเป็นไปได้ในระดับต่ำถึงปานกลาง และเมื่อพิจารณาร้อยละของนักเรียนที่ไม่มีความรู้ (LK) จากผลการวินิจฉัยด้วยแบบสอบแบบเลือกตอบสามระดับ มีค่าเท่ากับ 50.00 ส่วนการวินิจฉัยด้วยแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับ มีค่าเท่ากับ 56.45 สอดคล้องกับข้อเสนอแนะของ Gurel, Eryilmaz and Mcdermott (2015) ที่ระบุว่าแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับควรให้การแปลผลที่ชัดเจนกว่าแบบสอบวินิจฉัยแบบเลือกตอบสามระดับ เนื่องจากการวัดระดับความมั่นใจที่แตกต่างกัน โดยที่แบบสอบวินิจฉัยแบบเลือกตอบสามระดับเป็นการวัดความมั่นใจในภาพรวม ส่วนแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับเป็นการตอบระดับความมั่นใจ ที่แยกกันชัดเจนระหว่างการตอบระดับคำตอบและการตอบระดับเหตุผล ทำให้สามารถแยกกลุ่มนักเรียนที่ตอบโดยไม่รู้คำตอบออกจากกลุ่มนักเรียนที่มี McNamara's test ที่คลาดเคลื่อนทั้ง 3 ได้ดีกว่าแบบสอบวินิจฉัยแบบเลือกตอบสามระดับ (Gurel et al., 2015)

ผลการวินิจฉัยด้วยแบบสอบวินิจฉัยแบบเลือกตอบสามระดับกับผลการสัมภาษณ์ มีค่าสัมประสิทธิ์สหสัมพันธ์วีของคราเมอร์ อยู่ระหว่าง 0.320 ถึง 0. ส่วนผลการวินิจฉัยด้วยแบบสอบวินิจฉัยแบบเลือกตอบสี่ระดับกับผลการสัมภาษณ์ มีค่าสัมประสิทธิ์สหสัมพันธ์วีของคราเมอร์ อยู่ระหว่าง 0.485 ถึง 0.609 เมื่อเปรียบเทียบค่าสัมประสิทธิ์สหสัมพันธ์ของตัวอย่างที่ทับซ้อนกัน (Comparing correlation coefficients of

overlapping samples) พบว่าผลวินิจฉัยจากแบบสอบถามวินิจฉัยทั้งสองฉบับมีความสอดคล้องกับข้อมูลเชิงประจักษ์จากการสัมภาษณ์ไม่แตกต่างกัน

ข้อเสนอแนะ

ผู้วิจัยมีข้อเสนอแนะในการนำผลการวิจัยในครั้งนี้ไปใช้ มีดังนี้

1) แบบสอบถามวินิจฉัยแบบเลือกตอบสามารถระดับและสักระดับให้ผลการวินิจฉัยที่สอดคล้องกับข้อมูลเชิงประจักษ์ไม่แตกต่างกัน ดังนั้นจึงสามารถนำวิธีการสร้างแบบสอบถามวินิจฉัยแบบเลือกตอบสามารถระดับและสักระดับในการวิจัยนี้ไปประยุกต์ใช้สร้างแบบสอบถามวินิจฉัยแบบเลือกตอบในมโนทัศน์เรื่องอื่น ๆ เพื่อวินิจฉัยว่านักเรียนถูกจำแนกอยู่ในกลุ่มใด ได้แก่ นักเรียนที่มีมโนทัศน์ที่ถูกต้อง นักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อน นักเรียนที่ตอบถูกด้วยการเดา หรือนักเรียนที่ไม่มีความรู้ ทั้งนี้เพื่อที่จะเป็นข้อมูลที่ครูผู้สอนสามารถนำไปวางแผนเพื่อส่งเสริมศักยภาพนักเรียน หรือหาทางแก้ไขนักเรียนที่มีมโนทัศน์ที่คลาดเคลื่อน รวมทั้งพัฒนานักเรียนที่ไม่มีความรู้ให้มีความรู้และมีมโนทัศน์ที่ถูกต้องในเรื่องนั้น ๆ ซึ่งควรเลือกใช้แบบสอบถามวินิจฉัยแบบเลือกตอบสามารถระดับมากกว่าแบบสักระดับ เนื่องจากแบบสอบถามวินิจฉัยแบบเลือกตอบสามารถระดับเป็นแบบสอบที่ประหยัดเวลาในการสร้างและประหยัดเวลาในการสอบมากกว่าแบบสอบถามวินิจฉัยแบบเลือกตอบสักระดับ

2) การนำแบบสอบถามวินิจฉัยแบบเลือกตอบไปใช้ทดสอบเพื่อวินิจฉัยการเรียนรู้ของนักเรียน อาจแบ่งทดสอบเป็นรายมโนทัศน์หลัก จะทำให้สามารถประหยัดเวลาการทำแบบสอบและลดความเหนื่อยล้าของนักเรียน ทำให้ได้ผลทดสอบที่มีความคลาดเคลื่อนน้อยลงด้วย อนึ่งการทดสอบด้วยแบบสอบถามวินิจฉัยเป็นการทดสอบเพื่อมุ่งหวังการวินิจฉัยการเรียนรู้ของนักเรียน ไม่ควรใช้เพื่อการตัดสินผลการเรียน แต่ควรใช้การแปลผลการวินิจฉัยเพื่อซ่อมเสริมและส่งเสริมการเรียนรู้ของนักเรียน

3) การทดสอบด้วยแบบสอบถามวินิจฉัยแบบเลือกตอบสักระดับเป็นสิ่งที่นักเรียนไม่คุ้นเคย อีกทั้งในแบบสอบแต่ละข้อมีการวัดระดับความมั่นใจของการตอบ ซึ่งเป็นการประเมินค่าความมั่นใจภายในของนักเรียนแต่ละคน การตอบระดับความมั่นใจมีผลต่อการวินิจฉัยนักเรียนออกเป็นกลุ่มต่าง ๆ ดังนั้นการนำแบบสอบถามวินิจฉัยแบบเลือกตอบสักระดับไปใช้ควรอธิบายวิธีการทำแบบสอบอย่างละเอียดให้นักเรียนเข้าใจอย่างชัดเจนเพื่อป้องกันความสับสน ทำให้นักเรียนตอบข้อสอบในแต่ละระดับได้ครบถ้วน และเน้นย้ำให้นักเรียนทำแบบสอบด้วยความตั้งใจ หรือสร้างแรงจูงใจให้นักเรียนตั้งใจทำแบบสอบ จะทำให้ผลวินิจฉัยแม่นยำยิ่งขึ้น

4) จากการวินิจฉัยในครั้งนี้พบว่า มีนักเรียนส่วนใหญ่ไม่มีความรู้ในมโนทัศน์หลักเรื่องการถ่ายทอดลักษณะทางพันธุกรรม และการแปรผันทางพันธุกรรม เป็นสิ่งที่ครูผู้สอนควรเร่งพัฒนานักเรียนให้มีความรู้ส่วนมโนทัศน์หลักที่นักเรียนมีมโนทัศน์ที่คลาดเคลื่อนได้แก่ เทคโนโลยีชีวภาพทางพันธุกรรม ความหลากหลายทางชีวภาพ และวิวัฒนาการ ซึ่งครูควรวางแผนเพื่อให้นักเรียนมีมโนทัศน์ที่ถูกต้อง รวมทั้งหาสาเหตุของการมีมโนทัศน์ที่คลาดเคลื่อน มโนทัศน์หลักทั้ง 5 ในการวิจัยนี้ จัดอยู่ในมาตรฐาน ว1.2 ซึ่งมีผลการทดสอบระดับชาติขั้นพื้นฐาน (O-NET) ค่อนข้างต่ำในทุก ๆ ปีการศึกษา การที่นักเรียนมีผลการทดสอบระดับต่ำเกิดจากการไม่มีความรู้ ซึ่งอาจเกิดจากการที่นักเรียนไม่เข้าใจเนื้อหาที่ครูสอนในห้องเรียน หรือนักเรียนได้เรียนรู้แล้วแต่เมื่อเวลาผ่านไปนักเรียนลืมเนื้อหาดังกล่าว ส่วนการที่นักเรียนมีมโนทัศน์ที่คลาดเคลื่อนอาจเกิดจากการสร้างมโนทัศน์ที่คลาดเคลื่อนด้วยตัวนักเรียนเอง หรือเกิดจากการจัดการเรียนรู้ของครู ปัญหาที่พบนี้เป็นสิ่งที่ครูจะได้ใช้เป็นข้อมูลในการป้องกันและแก้ไขปัญหานี้ในอนาคตต่อไป

สำหรับข้อเสนอแนะในการทำวิจัยครั้งต่อไปมีดังนี้

1) การวัดความรู้ระดับเหตุผลในครั้งนี้เป็นการวัดด้วยตัวเลือกที่ผู้วิจัยกำหนดให้ ในการวิจัยครั้งต่อไป อาจเพิ่มตัวเลือกที่ให้นักเรียนสามารถอธิบายเหตุผลของนักเรียนในกรณีที่เหตุผลของนักเรียนไม่มีในตัวเลือกที่ผู้วิจัยเสนอให้ จะทำให้ได้ผลวิจัยที่แตกต่างจากครั้งนี้หรือไม่

2) การวินิจฉัยด้วยแบบสอบครั้งนี้เป็นการทดสอบด้วยการตอบลงในแบบสอบ (paper-pencil test) ในครั้งต่อไปถ้าออกแบบให้เป็นการทดสอบออนไลน์ผ่านระบบคอมพิวเตอร์ โดยกำหนดรูปแบบการทดสอบให้มีความตื่นเต้นเร้าใจ ด้วยภาพและเสียง หรือสื่อวีดิทัศน์ อาจจะทำให้ผู้เข้าสอบรู้สึกสนใจในการทำแบบสอบมากขึ้น หรือการเพิ่มการให้ข้อมูลป้อนกลับทันทีแก่นักเรียนหลังจากการทดสอบ ก็จะทำให้ให้นักเรียนทราบว่านักเรียนถูกจัดกลุ่มอยู่ในกลุ่มใด เพื่อการพัฒนานักเรียนด้วยตัวของนักเรียนเอง

3) การวิจัยในครั้งต่อไปควรศึกษาการวินิจฉัยด้วยทฤษฎีการตอบสนองข้อสอบ (Item response theory: IRT) เพื่อลดปัญหาค่าพารามิเตอร์ความยากและอำนาจจำแนก และการเดาของข้อสอบแต่ละข้อที่แปรผันตามกลุ่มของผู้เข้าสอบ แต่ในการวิจัยครั้งนี้มีจำนวนตัวอย่างในการวินิจฉัยน้อยจึงได้ใช้การทดสอบตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical test theory)

4) การวิจัยในครั้งนี้พบว่านักเรียนกลุ่มที่ไม่มีความรู้ส่วนใหญ่ใหญ่เป็นนักเรียนที่ไม่ใช่แผนการเรียนวิทยาศาสตร์ ในการวิจัยครั้งต่อไปอาจเพิ่มแผนการเรียนของนักเรียนเป็นตัวแปรอิสระ แล้วศึกษาว่าในแต่ละแผนการเรียนแบบสอบวินิจฉัยแบบเลือกตอบสามารถบ่งชี้ระดับให้ผลการวินิจฉัยที่แตกต่างหรือไม่ อย่างไร ซึ่งในการวิจัยในครั้งนี้เนื่องจากข้อจำกัดด้านจำนวนของตัวอย่างในแต่ละแผนการเรียน จึงได้ศึกษาผลการวินิจฉัยในภาพรวม

รายการอ้างอิง

ภาษาไทย

กระทรวงศึกษาธิการ. (2539). *แนวทางการสร้างแบบสอบวินิจฉัยเพื่อพัฒนาการเรียนการสอน*.

กรุงเทพมหานคร: ศุภสภา.

กระทรวงศึกษาธิการ. (2551). *ตัวชี้วัดและสาระการเรียนรู้แกนกลางกลุ่มสาระการเรียนรู้วิทยาศาสตร์ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551*. กรุงเทพมหานคร: โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย.

พิมพ์พันธ์ เดชะคุปต์. (2558). *การจัดการเรียนรู้ในศตวรรษที่ 21*. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

มนัสสิริ อินทร์สวาท. (2559). การเปรียบเทียบคุณภาพของแบบสอบวินิจฉัยแบบสามระดับสำหรับมโนทัศน์ที่คลาดเคลื่อนทางคณิตศาสตร์ที่ใช้ระดับความมั่นใจที่ต่างกัน (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต ไม่ได้ตีพิมพ์). จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพมหานคร.

วนิดา ภู่อี่ยม. (2550). *การเปรียบเทียบสัมประสิทธิ์การสรุปอ้างอิงและความแปรปรวนของความคลาดเคลื่อนระหว่างแบบสอบวินิจฉัยวิชาคณิตศาสตร์แบบประเพณีนิยมกับแบบสองระดับ* (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต ไม่ได้ตีพิมพ์). จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพมหานคร.

วิดา ซ่อนจำ. (2551). *การสร้างแบบทดสอบวินิจฉัยจุดบกพร่องในการเรียนวิชาคณิตศาสตร์ เรื่องจำนวนและการดำเนินการ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3* (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต ไม่ได้ตีพิมพ์). มหาวิทยาลัยศรีนครินทรวิโรฒ, กรุงเทพมหานคร.

- ศิริเดช สุชีวะ. (2538). *การพัฒนาวิธีการวินิจฉัยสำหรับตรวจสอบโนทัศน์ที่คลาดเคลื่อนทางคณิตศาสตร์* (วิทยานิพนธ์ปริญญาคุชฎบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพมหานคร.
- สุปราณี บุระ. (2557). *การพัฒนาแบบสอบเชิงวินิจฉัยทางพุทธิปัญญาเรื่องการดำเนินการเลขคณิตพื้นฐานโดยใช้คอมพิวเตอร์: การประยุกต์โมเดลดีไอเอ็นเอ* (วิทยานิพนธ์ปริญญาคุชฎบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพมหานคร.

ภาษาอังกฤษ

- Arslan, H. O., Cigdemoglu, C., & Moseley, C. (2012). A three-tier diagnostic test to assess pre-service teachers' misconceptions about global warming, greenhouse effect, ozone depletion, and acid rain. *International Journal of Science and mathematical Education*, 34(11), 1667-1686.
- Caleon, I. S., & Subramaniam, R. (2010). Do students know what they know and what they don't know? Using a four-tier diagnostic test to assess the nature of students' alternative conceptions. *Research in Science Education*, 40(3), 313-317.
- Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science. *Eurasia Journal of Mathematics, Science & Technology Education*, 11(5), 989-1008.
- Lin Sheau-Wen. (2004). Development and application of two-tier diagnostic test for high school students' understanding of flowering plant growth and development. *International Journal of Science and mathematical Education*, 2(2), 175-199.
- Sia Ding Teng, Treagust, D. F., & Chandrasegaran, A. L. (2012). High school students' proficiency and confidence levels in displaying of basic electrolysis concepts. *International Journal of Science and mathematical Education*, 10(6), 1325-1345.
- Wang Jing-Ru. (2004). Development and validation of a two-tier instrument to examine understanding of internal transport in plants and the human circulatory system. *International Journal of Science and mathematics Education*, 2(2), 131-154.
- Yang Der-Ching, & Lin Yung-Chi. (2015). Assessing 10- to 11-year-old children's performance and misconceptions in number sense using a four-tier diagnostic test. *Educational research*, 57(4), 368-388.