



ความสัมพันธ์ระหว่างแรงจูงใจในการสอบกับคะแนนสอบและค่าการเดาของนักเรียน
ที่มีระดับความสามารถทางการเรียนแตกต่างกันในการทดสอบที่มีความสำคัญต่ำ
CORRELATIONS OF TEST-TAKING MOTIVATION, TEST SCORE AND GUESSING
PARAMETER WITH DIFFERENT STUDENT'S ACADEMIC ACHIEVEMENT
IN LOW-STAKES TESTING

นางสาวพรพิมล ค่อมสิงห์ *

Pornpimon Komsing

รศ.ดร.วรรณิ แกมเกตุ **

Assoc. Prof. Wannee Kaemkate, Ph.D.

บทคัดย่อ

การวิจัยมีวัตถุประสงค์ดังนี้ 1) เพื่อพัฒนาและตรวจสอบคุณภาพของแบบวัดแรงจูงใจในการสอบต่อข้อสอบในการทดสอบที่มีความสำคัญต่ำรายข้อ 2) เพื่อตรวจสอบคุณภาพของข้อสอบในการทดสอบที่มีความสำคัญต่ำ 3) เพื่อเปรียบเทียบระดับแรงจูงใจในการสอบของนักเรียนที่มีความสามารถทางการเรียนแตกต่างกัน 4) เพื่อศึกษาความสัมพันธ์ระหว่างแรงจูงใจในการสอบกับคะแนนสอบในการทดสอบที่มีความสำคัญต่ำและค่าการเดาของนักเรียนที่มีระดับความสามารถทางการเรียนแตกต่างกัน 5) เพื่อศึกษาความสัมพันธ์ระหว่างค่าความสามารถของผู้สอบ (θ) กับเกรดเฉลี่ยสะสมวิชาวิทยาศาสตร์ของนักเรียน กลุ่มตัวอย่างคือนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียนในสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 1 และ 2 จำนวน 662 คน เครื่องมือที่ใช้คือแบบวัดแรงจูงใจในการสอบและแบบสอบการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์

ผลวิจัยพบว่า 1) แบบวัดแรงจูงใจในการสอบมี 1 องค์ประกอบ มีค่าความเที่ยงแบบความสอดคล้องภายในเท่ากับ .734 2) ข้อสอบที่มีความสำคัญต่ำมีความเที่ยงแบบ KR20 เท่ากับ .785 มีข้อสอบผ่านเกณฑ์ตามทฤษฎีการทดสอบแบบดั้งเดิม 27 ข้อ และมีค่าความเที่ยงแบบ marginal เท่ากับ .838 มีข้อสอบผ่านเกณฑ์ตามทฤษฎีการตอบสนองข้อสอบ 22 ข้อ 3) ผู้สอบที่มีความสามารถทางการเรียนต่างกันมีค่าเฉลี่ยของแรงจูงใจในการสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 4) สัมประสิทธิ์สหสัมพันธ์ระหว่างแรงจูงใจในการสอบและคะแนนสอบมีค่าเท่ากับ .297 และไม่พบความสัมพันธ์ระหว่างแรงจูงใจในการสอบและค่าการเดาในกลุ่มผู้สอบที่มีความสามารถทางการเรียนต่างกัน 6) สัมประสิทธิ์สหสัมพันธ์ระหว่างค่าความสามารถของผู้สอบ (θ) และเกรดเฉลี่ยสะสมวิชาวิทยาศาสตร์เท่ากับ .632

* นิสิตมหาบัณฑิตสาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย

E-mail Address: Pornpimon@suriyothai.ac.th

**อาจารย์ประจำสาขาวิชาวิจัยการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

E-mail Address: Wannee.k@gmail.com

ISSN1905-4491

Abstract

The purposes of research were 1) to develop and investigate test-taking motivation questionnaires quality in low-stakes testing 2) to investigate test quality in low-stakes testing 3) to compare test-taking motivation of student with different student's academic achievement 4) to study the correlation among test-taking motivation, test score and guessing parameter with different student's academic achievement in low-stakes testing 5) to study the correlation between ability parameter (θ) and Science's grade point average. The participants were 662 eighth grade student in Bangkok. The research tools were test-taking motivation questionnaires and Low stakes Test (Trends in International Mathematics and Science Study: TIMSS). The research findings were as follows: 1) There was one factor of test-taking motivation questionnaires; called effort and Cronbach's Alpha reliability was .734 2) Low-stakes test had KR20 reliability was .785, Classical theory were accepted 27 items, had marginal reliability was .838, Item response theory were accepted 22 items. 3) There were significant consequences of test-taking motivation under different student's ability levels at .05 4) There was a correlation (.297) between test-taking motivation and test score 5) There was no correlation between test-taking motivation and guessing parameter under different student's ability 6) There was a correlation (.632) between ability parameter and Science's grade point average.

คำสำคัญ: แรงจูงใจในการสอบ, การทดสอบที่มีความสำคัญต่ำ, ค่าการเดา

KEYWORDS: TEST-TAKING MOTIVATION / LOW-STAKES TESTING / GUESSING PARAMETER

บทนำ

จากการศึกษาพบว่า การประเมินระดับชาติมี 2 รูปแบบ คือ การประเมินระดับชาติที่ไม่มีผลต่อการตัดสินผลการเรียนหรือการประเมินที่มีความสำคัญต่ำ (Low-Stakes Assessment) เป้าหมายของการประเมินแบบนี้คือการได้ข้อมูลเกี่ยวกับคุณภาพการศึกษาของชาติในลักษณะภาพรวม ผลการประเมินสามารถแสดงให้เห็นคุณภาพการศึกษา ณ เวลาที่ประเมินและแสดงแนวโน้มในระยะยาว การประเมินแบบนี้ได้แก่ การศึกษาแนวโน้มการศึกษาวិชาคณิตศาสตร์และวิทยาศาสตร์ระหว่างประเทศ (Trends in International Mathematics and Science Study: TIMSS) โครงการประเมินผลทางการศึกษานานาชาติ (The National Assessment of Educational Progress: NAEP) โครงการประเมินผลผู้เรียนระดับนานาชาติ (The Programme for International Student Assessment: PISA) ของสหรัฐอเมริกาและ Qualifications and Curriculum Authority's (QCE) National Test ของสหราชอาณาจักร และอีกประเภทหนึ่งคือการประเมินผลระดับชาติเพื่อตัดสินผลการเรียนรู้ของผู้เรียนระดับตัวบุคคลต่างๆ หรือการประเมินที่มีความสำคัญสูง (High-Stakes Assessment) เช่น การสอบ O-Level และ A-Level ของสหราชอาณาจักร การสอบ Higher school Certificate ในรัฐ New Wales ประเทศออสเตรเลีย การทดสอบดำเนินการโดยหน่วยงานที่รับผิดชอบ

เกี่ยวกับการประเมินผลสัมฤทธิ์การเรียนรู้ของประเทศต่างๆ เช่น สหรัฐอเมริกา (NAGB) สหราชอาณาจักร (QCA) และสิงคโปร์ (Singapore Ministry of Education) (นิพล พลกลาง, 2549)

การศึกษาแนวโน้มการศึกษาวិชาคณิตศาสตร์และวิทยาศาสตร์ระหว่างประเทศ (Trends in International Mathematics and Science Study: TIMSS) มีจุดมุ่งหมายเพื่อประเมินผลสัมฤทธิ์ในการเรียนรู้ของผู้เรียนในวิชาคณิตศาสตร์และวิทยาศาสตร์ตามหลักสูตรที่กำลังศึกษาอยู่ในสถานศึกษาของผู้เรียนชั้นประถมศึกษาปีที่ 4 และมัธยมศึกษาปีที่ 2 (สำนักงานเลขาธิการสภาการศึกษา, 2554) จากเอกสารสรุปผลการ วิจัยโครงการ TIMSS 2011 ชั้นมัธยมศึกษาปีที่ 2 พบว่าในวิชาวิทยาศาสตร์ คะแนนรวมของประเทศไทยอยู่ในลำดับที่ 25 จาก 45 ประเทศที่เข้าร่วม นักเรียนส่วนใหญ่มีความสามารถอยู่ในระดับ 1 (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2556)

นักวิจัยหลายคนให้ความสนใจศึกษาคะแนนที่ได้จากการสอบข้อสอบนานาชาติที่ผลการประเมินไม่ได้ถูกนำไปใช้ในการตัดสินผลการเรียนหรือนำไปใช้ในการคัดเลือกเข้ามหาวิทยาลัย ทำให้นักเรียนบางส่วนไม่ทุ่มเทความพยายามในการทำข้อสอบหรือไม่มีแรงจูงใจในการสอบ Baumert and Demmrich (2001) ได้ให้ความหมายของแรงจูงใจในการสอบว่าหมายถึง การเต็มใจที่จะเข้ามาเรียนรู้สิ่งที่เกี่ยวข้องกับการสอบ ทุ่มเทความสามารถ ความพยายามและมุ่งมั่นในการสอบ สอดคล้องกับ Wise and DeMars (2005) ได้ให้นิยามแรงจูงใจในการสอบว่าหมายถึง ความพยายามของบุคคลที่จะทำการสอบให้ดีที่สุด เป็นกิจกรรมที่มุ่งไปยังเป้าหมาย ทำให้การสอบเป็นตัวแทนที่แท้จริงของสิ่งที่บุคคลนั้นรู้โดยผ่านกระบวนการสอบและรักษาไว้ซึ่งระดับของความพยายามในการสอบ ทำให้ผลการสอบมีความตรงและแสดงถึงระดับความสามารถที่แท้จริงของผู้สอบ จากการสังเคราะห์งานวิจัยของ Eklöf (2008) ที่ศึกษาแรงจูงใจในการสอบของนักเรียนประเทศสวีเดน ในการทดสอบตามโครงการวิจัยนานาชาติ TIMSS 2003 พบว่านักเรียนมีแรงจูงใจในการสอบอยู่ในระดับดี และแรงจูงใจในการสอบมีความสัมพันธ์เชิงบวกในระดับต่ำ ($r = 0.25$) กับคะแนนสอบวิชาคณิตศาสตร์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 สอดคล้องกับ Eklöf and Nyroos's (2013) ได้วิเคราะห์ข้อมูลเกี่ยวกับคะแนนวิชาวิทยาศาสตร์ของการทดสอบระดับชาติของประเทศสวีเดนของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในปี ค.ศ. 2009 พบว่ามีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติระหว่างคะแนนสอบวิชาวิทยาศาสตร์และ (a) การรายงานแรงจูงใจ ($r = 0.25$), (b) การรับรู้ความสำคัญของการสอบ ($r = 0.20$) และ (c) ความกังวลในการสอบ ($r = -0.10$) สอดคล้องกับ Eklöf, Pavešić, and Grønmo (2014) ที่ได้ศึกษาเปรียบเทียบการรายงานแรงจูงใจในการสอบและคะแนนสอบวิชาคณิตศาสตร์ตามโครงการวิจัยนานาชาติ TIMSS ชั้นสูงระหว่างประเทศสวีเดน นอร์เวย์ และสโลวีเนีย พบว่าการรายงานแรงจูงใจในการสอบมีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติกับคะแนนสอบในกลุ่มตัวอย่างทั้งสาม พร้อมทั้งได้ให้ข้อเสนอแนะให้มีการควบคุมและชี้แนะให้นักเรียนพยายามทำข้อสอบในการทดสอบที่มีความสำคัญต่อผู้เรียนต่ำ โดยพฤติกรรมที่แสดงถึงการไม่มีแรงจูงใจในการสอบของนักเรียน ได้แก่ การเดาข้อสอบ การละเลยข้อสอบ และการตอบอย่างรวดเร็วเกินไป

การวัดแรงจูงใจในการสอบนั้นยังไม่มีแบบวัดที่เป็นมาตรฐาน นักวิจัยจึงต้องสร้างแบบวัดหรือปรับปรุงจากแบบวัดที่มีอยู่เพื่อนำมาใช้ได้อย่างเหมาะสมกับบริบทที่ต้องการศึกษา Eklöf (2010) ได้นำแบบวัดฉบับเดิมของ Wolf and Smith (1995) จำนวน 4 ข้อ มาปรับปรุงและได้เพิ่มข้อคำถามอีก 2 ข้อ ทำให้แบบวัดความพยายามประกอบด้วย 6 ข้อคำถาม แบบวัดฉบับนี้เป็นมาตรประมาณค่า 4 ระดับ การตอบมีตั้งแต่ 1 เห็นด้วยอย่างยิ่งถึง 4 ไม่เห็นด้วยอย่างยิ่ง ข้อคำถาม 3 ข้อในแบบวัดต้องปรับสเกลก่อนที่จะมีการวิเคราะห์ แล้วจึงนำ 6 ข้อคำถามเพิ่มในภูมิหลังของ TIMSS Student Questionnaire เมื่อนักเรียนทำข้อสอบเสร็จแล้ว จึงให้นักเรียนรายงานแรงจูงใจในการสอบ ผลการวิเคราะห์พบว่าค่าความเที่ยงแบบความสอดคล้องภายในของ 6 ข้อคำถามของแบบวัดความพยายามอยู่ในระดับที่ยอมรับได้ โดยกลุ่มตัวอย่างในประเทศสวีเดนมีค่าสัมประสิทธิ์แอลฟาเท่ากับ .79 ประเทศนอร์เวย์มีค่าสัมประสิทธิ์แอลฟาเท่ากับ .82 และในกลุ่มตัวอย่างประเทศสโลวีเนียมีค่าสัมประสิทธิ์แอลฟาเท่ากับ .82 ทั้ง 6 ข้อคำถามมีความสัมพันธ์ระหว่างข้ออย่างมีนัยสำคัญทางสถิติที่ระดับ .01 การวิเคราะห์องค์ประกอบเชิงสำรวจพบว่ามีเพียงองค์ประกอบเดียว หลังจากการสกัดองค์ประกอบและหมุนแกนพบว่าองค์ประกอบ 1 องค์ประกอบสามารถอธิบายความแปรปรวนของแรงจูงใจในการสอบของกลุ่มตัวอย่างในประเทศสวีเดนได้ 49% ประเทศนอร์เวย์และสโลวีเนีย 53% การวิเคราะห์องค์ประกอบเชิงยืนยันพบว่าโมเดลสอดคล้องกับข้อมูลเชิงประจักษ์ (model fit)

อีกสิ่งหนึ่งที่ส่งผลต่อความตรงของคะแนนคือ การเดาข้อสอบ (guessing) หากให้โอกาสผู้เข้าสอบเดาข้อสอบได้จะส่งผลต่อความตรงและความเที่ยงของแบบสอบ ทำให้เกิดแหล่งความคลาดเคลื่อนในการวัด Obinne (2012) ให้ความหมายของการเดา (guessing) ว่าหมายถึง การให้คำตอบหรือการทำบางสิ่งบางอย่างเกี่ยวกับการตัดสินใจที่ไม่มีการตรวจสอบข้อเท็จจริงทั้งหมด ข้อสอบปรนัยเลือกตอบหลายตัวเลือก (Multiple Choice: MC) เป็นรูปแบบของแบบสอบที่ผู้เข้าสอบสามารถเดาคำตอบได้ โดยโอกาสในการเดาคำตอบได้ถูกขึ้นกับจำนวนและประสิทธิภาพของตัวเลือก (ศิริชัย กาญจนวาสี, 2544)

การตรวจสอบคุณภาพของข้อสอบการศึกษาแนวโน้มการศึกษาวิชาคณิตศาสตร์และวิทยาศาสตร์ระหว่างประเทศ (TIMSS) และความสามารถของผู้สอบ โดยใช้ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ซึ่งอธิบายความสัมพันธ์ระหว่างความสามารถของผู้สอบและคำตอบของผู้สอบโดยใช้คุณลักษณะของข้อสอบ ได้แก่ ความยาก (difficulty parameter) อำนาจจำแนก (discrimination parameter) และค่าการเดา (guessing parameter) โดยแนวคิดของทฤษฎีนี้คือการนำข้อมูลจากการตอบมาประมาณค่าพารามิเตอร์ของผู้สอบและข้อสอบตามโมเดลที่ใช้ (ศิริชัย กาญจนวาสี, 2545)

จากที่กล่าวมาแสดงให้เห็นถึงความสำคัญของการศึกษาแรงจูงใจในการสอบและคะแนนสอบของการทดสอบที่มีความสำคัญต่ำ แม้ว่าจะมีข้อมูลเชิงประจักษ์จากการศึกษาความสัมพันธ์ระหว่างแรงจูงใจในการสอบและคะแนนสอบข้อสอบที่มีความสำคัญต่ำ แต่เป็นการศึกษาโดยวัดแรงจูงใจในการสอบที่มีต่อแบบสอบทั้งฉบับ ไม่มีรายงานการศึกษาแรงจูงใจในการสอบต่อข้อสอบรายข้อ และไม่มีรายงานการศึกษาความสัมพันธ์

ระหว่างแรงจูงใจในการสอบและค่าการเดาของนักเรียนที่มีระดับความสามารถทางการเรียนแตกต่างกัน ผู้วิจัยจึงสนใจศึกษาความสัมพันธ์ระหว่างแรงจูงใจในการสอบกับคะแนนสอบและค่าการเดาของนักเรียนที่มีระดับความสามารถทางการเรียนแตกต่างกันในการทดสอบที่มีความสำคัญต่ำ

วัตถุประสงค์

การวิจัยครั้งนี้มีจุดมุ่งหมายหลักเพื่อศึกษาความสัมพันธ์ระหว่างแรงจูงใจในการสอบกับคะแนนสอบและค่าการเดาของนักเรียนที่มีระดับความสามารถทางการเรียนแตกต่างกันในการทดสอบที่มีความสำคัญต่ำ โดยมีวัตถุประสงค์เฉพาะดังนี้

1. เพื่อพัฒนาและตรวจสอบคุณภาพของแบบวัดแรงจูงใจในการสอบต่อข้อสอบในการทดสอบที่มีความสำคัญต่ำรายข้อ
2. เพื่อตรวจสอบคุณภาพของข้อสอบในการทดสอบที่มีความสำคัญต่ำ
3. เพื่อเปรียบเทียบระดับแรงจูงใจในการสอบของนักเรียนที่มีระดับความสามารถทางการเรียนแตกต่างกัน
4. เพื่อศึกษาความสัมพันธ์ระหว่างแรงจูงใจในการสอบกับคะแนนสอบในการทดสอบที่มีความสำคัญต่ำและค่าการเดาของนักเรียนที่มีระดับความสามารถทางการเรียนแตกต่างกัน
5. เพื่อศึกษาความสัมพันธ์ระหว่างค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) กับเกรดเฉลี่ยสะสมวิชาวิทยาศาสตร์ของนักเรียน

วิธีดำเนินการวิจัย

การวิจัยนี้เป็นการวิจัยเชิงบรรยาย (Descriptive research) เพื่อศึกษาความสัมพันธ์ระหว่างแรงจูงใจในการสอบกับคะแนนสอบและค่าการเดาของนักเรียนที่มีระดับความสามารถทางการเรียนแตกต่างกันในการทดสอบที่มีความสำคัญต่ำ

กลุ่มตัวอย่างในการวิจัย

กลุ่มตัวอย่างในการวิจัยครั้งนี้เป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 2 ปีการศึกษา 2558 ของโรงเรียนในสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน สำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 1 และ 2 จำนวน 662 คน

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการศึกษาครั้งนี้ประกอบด้วย

1. แบบวัดแรงจูงใจในการสอบ เป็นมาตรวัดประมาณค่า (rating scale) 5 ระดับ จำนวน 3 ข้อ ซึ่งจะถูกนำไปใช้ในการวัดแรงจูงใจในการสอบต่อข้อสอบรายข้อ

การสร้างแบบวัดแรงจูงใจในการสอบ ผู้วิจัยได้ดำเนินการสร้างตามขั้นตอนดังนี้

- 1) ผู้วิจัยศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับแรงจูงใจในการสอบ เพื่อกำหนดนิยาม

ปฏิบัติการโครงสร้างของตัวแปรแรงจูงใจในการสอบ องค์ประกอบ และลักษณะการเขียนข้อคำถาม

2) ผู้วิจัยนำแบบวัด The Effort Scale ของ Eklöf (2010) ที่ประกอบด้วยข้อคำถาม 6 ข้อ คำถามและพัฒนาเพิ่ม 1 ข้อคำถาม รวม 7 ข้อคำถาม แล้วปรับเป็นมาตรประมาณค่า 5 ระดับ โดยมีตั้งแต่ 1 ไม่เห็นด้วยอย่างยิ่ง ถึง 5 เห็นด้วยอย่างยิ่ง ดังตาราง 1

3) การตรวจสอบความตรงเชิงเนื้อหา (content validity) โดยหาค่าดัชนีความสอดคล้องของ ข้อคำถาม (Item Objective Congruence: IOC) ผู้วิจัยนำเครื่องมือฉบับร่างพร้อมกับรายละเอียดหัวข้อการ วิจัย ให้ผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา ผู้เชี่ยวชาญด้านจิตวิทยา และผู้เชี่ยวชาญที่ปฏิบัติงาน ในโรงเรียน รวมจำนวน 5 ท่าน ตรวจสอบความตรงเชิงเนื้อหาที่เกี่ยวกับความสอดคล้องกับองค์ประกอบ ความถูกต้องเหมาะสมชัดเจนของการใช้ภาษา และข้อเสนอแนะอื่นๆ

ตาราง 1 ข้อคำถามในแบบวัดแรงจูงใจในการสอบ

ข้อที่	ข้อความ	ระดับความคิดเห็น				
		5	4	3	2	1
1	ฉันทำข้อสอบข้อนี้ด้วยความพยายามมากที่สุด					
2	ฉันไม่ให้ความสนใจขณะที่ทำข้อสอบข้อนี้					
3	ฉันใช้ความพยายามในการทำข้อสอบข้อนี้น้อยกว่าทำข้อสอบของโรงเรียน					
4	ฉันยังคงพยายามทำข้อสอบข้อนี้แม้ว่ามันจะยาก					
5	ฉันรู้สึกมีแรงจูงใจในการพยายามทำข้อสอบข้อนี้					
6	ขณะที่ทำข้อสอบข้อนี้ฉันน่าจะตั้งใจมากกว่านี้					
7	ฉันตั้งใจอ่านและทบทวนข้อสอบข้อนี้					

4) นำเครื่องมือไปทดลองใช้ (try out) กับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 2 โรงเรียนสตรีศรีสุริโยทัย จำนวน 234 คน

5) นำข้อมูลจากแบบวัดแรงจูงใจในการสอบมาวิเคราะห์รายข้อ ด้วยการทดสอบความสอดคล้องของข้อคำถาม (Homogeneity of the item) โดยการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างข้อคำถามแต่ละข้อกับคะแนนรวมทั้งหมด (corrected item-total correlation: CITC) โดยยอมรับข้อที่มีค่า CITC เท่ากับ 0.2 ขึ้นไป พิจารณาค่าน้ำหนักองค์ประกอบ (factor loading) โดยการวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) การสกัดปัจจัย (factor extraction) โดยใช้เทคนิค principal components analyses (PCA)

6) การแปลความหมายของคะแนน ผู้วิจัยกำหนดเกณฑ์เพื่อจัดระดับของค่าเฉลี่ยของแบบวัดแรงจูงใจในการสอบ เพื่อประโยชน์สำหรับการตีความของค่าเฉลี่ย ดังตาราง 2 และคำนวณคะแนนการรายงานแรงจูงใจในการสอบ ได้แก่ ค่าเฉลี่ยและร้อยละของนักเรียนที่เห็นด้วยหรือไม่เห็นด้วยในแต่ละข้อคำถาม หลังจากนั้นจึงตรวจสอบความสัมพันธ์ระหว่างการรายงานแรงจูงใจในการสอบและคะแนนสอบ โดยการวิเคราะห์สหสัมพันธ์ของเพียร์สัน (Pearson's r)

2. แบบสอบการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ วิชาวิทยาศาสตร์ จำนวน 30 ข้อ โดยเป็นข้อสอบเลือกตอบแบบหลายตัวเลือก

การคัดเลือกข้อสอบจากแบบสอบการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ วิชาวิทยาศาสตร์ ผู้วิจัยได้ดำเนินการคัดเลือกตามขั้นตอน ดังนี้

1) ศึกษากรอบการประเมินข้อสอบการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ วิชาวิทยาศาสตร์ และพิจารณาข้อสอบตามน้ำหนักคะแนนด้านเนื้อหาและพฤติกรรมการเรียนรู้ที่ประเมิน แล้วจึงคำนวณสัดส่วนของข้อสอบที่จะคัดเลือกเพื่อใช้ในการวิจัยตามเนื้อหาและพฤติกรรมการเรียนรู้ให้ได้จำนวน 30 ข้อ

2) สร้างผังข้อสอบ (Test Blueprint) และคัดเลือกข้อสอบวิชาวิทยาศาสตร์ จากแบบสอบการศึกษาแบบสอบการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ วิชาวิทยาศาสตร์ จำนวน 30 ข้อ ตามข้อ 2

การวิเคราะห์ข้อมูล

ก่อนการดำเนินการวิเคราะห์ข้อมูล ผู้วิจัยตรวจสอบความถูกต้องสมบูรณ์ของข้อมูลและตรวจสอบว่าข้อมูลครบถ้วนหรือไม่ หากพบว่ามีข้อมูลสูญหาย (missing data) ผู้วิจัยจะแทนที่ด้วยค่าเฉลี่ย

1. การวิเคราะห์ข้อมูลเบื้องต้นของกลุ่มตัวอย่างโดยการคำนวณค่าร้อยละ โดยใช้โปรแกรม SPSS for Windows

2. ตรวจสอบคุณสมบัติการวัดเพียงมิติเดียวของแบบวัดแรงจูงใจในการสอบโดยใช้วิธีวิเคราะห์องค์ประกอบ

วิเคราะห์ค่าสถิติ Bartlett's test of Sphericity เพื่อพิจารณาว่าตัวแปรมีความเหมาะสมในการวิเคราะห์องค์ประกอบหรือไม่ ค่าสถิติ Bartlett's test of Sphericity เป็นค่าสถิติทดสอบสมมติฐานว่าเมทริกซ์สหสัมพันธ์เป็นเมทริกซ์เอกลักษณะหรือไม่ โดยพิจารณาจากค่าระดับนัยสำคัญทางสถิติที่น้อยกว่าหรือเท่ากับ .05 แสดงว่าเมทริกซ์สหสัมพันธ์ไม่เป็นเมทริกซ์เอกลักษณะมีความเหมาะสมที่จะวิเคราะห์องค์ประกอบต่อไป

ค่าสถิติ Kaiser-Meyer-Olkin (KMO) เป็นดัชนีเปรียบเทียบขนาดสัมประสิทธิ์สหสัมพันธ์บางส่วน (partial correlation) ระหว่างตัวแปรแต่ละคู่ เมื่อจัดความแปรปรวนของตัวแปรอื่นๆ ออกไปแล้วว่ามีความสัมพันธ์ระหว่างตัวแปรมากพอที่จะนำไปวิเคราะห์หองค์ประกอบหรือไม่ หาก KMO มีค่าเข้าใกล้ 1 แสดงว่ามีความเหมาะสมมาก ส่วนค่าที่น้อยกว่า 0.50 เป็นค่าที่ไม่เหมาะสมที่จะนำไปวิเคราะห์หองค์ประกอบ

ทดสอบความสอดคล้องของข้อคำถาม (Homogeneity of the item) โดยการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างข้อคำถามแต่ละข้อกับคะแนนรวมทั้งหมด (corrected item-total correlation: CITC) โดยยอมรับข้อที่มีค่า CITC เท่ากับ 0.2 ขึ้นไป

3. ตรวจสอบให้คะแนน โดยข้อสอบแบบหลายตัวเลือกมีคะแนนเต็มข้อละ 1 คะแนน ตอบถูกได้ 1 คะแนน ตอบผิดได้ 0 คะแนน

4. วิเคราะห์ค่าความยาก (p) ค่าอำนาจจำแนก (r) โดยใช้ทฤษฎีทดสอบแบบดั้งเดิม โดยใช้โปรแกรม TAP และวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ เพื่อประมาณค่าพารามิเตอร์ความสามารถของผู้สอบและพารามิเตอร์ของข้อสอบการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ วิชาวิทยาศาสตร์ ได้แก่ ค่าความยาก (b) ค่าอำนาจจำแนก (a) ค่าการเดา (c) ตามทฤษฎีการสนองข้อสอบแบบเอกมิติ (Unidimensional Item Response Theory) และค่าความเที่ยงของแบบสอบโดยใช้การวิเคราะห์โมเดลการตอบสนองข้อสอบที่ใช้กับการตรวจให้คะแนนรายข้อแบบ 2 ค่า (Dichotomous IRT Models) หรือโมเดลที่มีการตรวจให้คะแนนแบบถูกให้ 1 ผิดให้ 0 (Dichotomous) โดยใช้โปรแกรม MULTILOG

5. จัดกลุ่มนักเรียนตามระดับความสามารถทางการเรียนต่ำ กลาง สูง ตามตำแหน่งเปอร์เซ็นต์ไทล์

6. ตรวจสอบให้คะแนนการรายงานแรงจูงใจในการสอบ และแปลความหมายให้คะแนนแรงจูงใจในการสอบของนักเรียนแต่ละคน คำนวณคะแนนการรายงานแรงจูงใจในการสอบ ได้แก่ ค่าเฉลี่ยและร้อยละของนักเรียนที่เห็นด้วยหรือไม่เห็นด้วยในแต่ละข้อคำถาม

7. นำคะแนนสอบสอบมาแปลงให้เป็นคะแนนมาตรฐานที่ (T-score) แล้วจึงวิเคราะห์ค่าสหสัมพันธ์ระหว่างแรงจูงใจในการสอบและคะแนนสอบของนักเรียนที่มีระดับความสามารถทางการเรียนสูง ปานกลาง และต่ำ โดยใช้วิธีวิเคราะห์สหสัมพันธ์ของเพียร์สัน (Pearson's r) โดยใช้โปรแกรม SPSS การแปลความหมายของค่าสัมประสิทธิ์สหสัมพันธ์ใช้เกณฑ์ของ Sakind และ Neil J. (อ้างถึงใน อวยพร เรื่องตระกูล, 2553) ดังนี้

ตาราง 2 การแปลความหมายของค่าสัมประสิทธิ์สหสัมพันธ์

ขนาดความสัมพันธ์	ความหมาย
0.0 – 0.2	มีความสัมพันธ์กันต่ำมาก
0.2 – 0.4	มีความสัมพันธ์กันต่ำ
0.4 – 0.6	มีความสัมพันธ์กันปานกลาง
0.6 – 0.8	มีความสัมพันธ์กันสูง
0.8 – 1.0	มีความสัมพันธ์กันสูงมาก

8. วิเคราะห์ค่าสหสัมพันธ์ระหว่างแรงจูงใจในการสอบและค่าการเดา (c) ของนักเรียนที่มีระดับความสามารถทางการเรียนสูง ปานกลาง และต่ำ โดยใช้วิธีวิเคราะห์สหสัมพันธ์ของเพียร์สัน (Pearson's r)
9. ตรวจสอบความสัมพันธ์ระหว่างค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) และเกรดเฉลี่ยสะสม วิชาวิทยาศาสตร์ โดยใช้วิธีวิเคราะห์ค่าสหสัมพันธ์ของเพียร์สัน (Pearson's r)
10. การทดสอบความแตกต่างของค่าเฉลี่ยระหว่างกลุ่ม (one-way ANOVAs) เพื่อทดสอบความแตกต่างระหว่างระหว่างการรายงานแรงจูงใจในการสอบของกลุ่มผู้สอบที่มีระดับความสามารถทางการเรียนสูง ปานกลาง และต่ำ
11. วิเคราะห์ความแปรปรวนโดยใช้การทดสอบภายหลัง (Games-Howell's post hoc) การหาขนาดอิทธิพลเพื่อประมาณค่าขนาดอิทธิพลของตัวแปรต้นที่มีผลต่อตัวแปรตาม โดยใช้ ω^2

ผลการวิจัย

1. ผลการตรวจสอบคุณภาพของแบบวัดแรงจูงใจในการสอบต่อข้อสอบในการทดสอบที่มีความสำคัญต่ำรายข้อ พบว่า แบบวัดแรงจูงใจในการสอบประกอบด้วย 1 องค์ประกอบ ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างข้อคำถามแต่ละข้อกับคะแนนรวมทั้งหมด (Corrected Item-Total Correlation: CITC) มีค่า CITC ผ่านเกณฑ์ ทั้ง 3 ข้อคำถาม มีค่าความเที่ยงแบบความสอดคล้องภายใน (Cronbach's Alpha) เท่ากับ 0.734 การวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ของข้อคำถามในแบบวัดแรงจูงใจในการสอบ พบว่าข้อคำถามทั้ง 3 ข้อมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 โดยข้อคำถามที่ 1 และข้อคำถามที่ 3 มีความสัมพันธ์กันมากที่สุด โดยมีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.698

ตาราง 3 แสดงค่าสถิติพื้นฐานและความสัมพันธ์ระหว่างข้อกับคะแนนรวมเฉลี่ยของข้อคำถามในแบบวัด
แรงจูงใจในการสอบ

องค์ประกอบ	ข้อที่	ข้อคำถาม	คะแนนแรงจูงใจในการสอบเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐานเฉลี่ย	ความสัมพันธ์ระหว่างข้อกับคะแนนรวมเฉลี่ย
ความพยายาม	1	ฉันทำข้อสอบข้อนี้ด้วยความพยายามมากที่สุด	4.05	0.99	0.490
	2	ฉันไม่สนใจในการทำข้อสอบข้อนี้	3.82	1.17	0.281
	3	ฉันยังคงพยายามทำข้อสอบข้อนี้แม่ว่ามันจะยาก	3.73	1.11	0.457

2. ผลการตรวจสอบคุณภาพของข้อสอบในการทดสอบที่มีความสำคัญต่ำ พบว่าเมื่อวิเคราะห์คุณภาพของข้อสอบการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ โดยใช้ทฤษฎีทดสอบแบบดั้งเดิม โดยใช้โปรแกรม TAP มีค่าความเที่ยง KR20 (Alpha) เท่ากับ 0.785 ค่าความยาก (p) อยู่ระหว่าง 0.219 – 0.845 ค่าความยากเฉลี่ยเท่ากับ 0.487 และค่าอำนาจจำแนก (r) อยู่ระหว่าง 0.187 – 0.568 ค่าอำนาจจำแนกเฉลี่ยเท่ากับ 0.396 ข้อสอบที่ผ่านเกณฑ์การยอมรับคุณภาพ โดยใช้ทฤษฎีทดสอบแบบดั้งเดิมมี 27 ข้อ ได้แก่ ข้อ 1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30

เมื่อวิเคราะห์โดยใช้ทฤษฎีการตอบสนองข้อสอบ โดยใช้โปรแกรม MULTILOG มีค่าความเที่ยง (marginal reliability) เท่ากับ 0.8378 ค่าไลค์ลิฮูด (negative twice the loglikelihood) เท่ากับ 14750.7 ค่าอำนาจจำแนก (a) มีค่าระหว่าง 0.39 – 3.98 ค่าความยาก (b) มีค่าระหว่าง -1.75 – 2.38 ค่าการเดา (c) มีค่าระหว่าง 0.00 – 0.64 ผู้วิจัยใช้เกณฑ์ในการพิจารณาคุณภาพของข้อสอบรายข้อ โดยพิจารณาจากค่าอำนาจจำแนก (a) มีค่าอยู่ระหว่าง +0.50 ถึง +2.50 โดยถ้าอำนาจจำแนก (a) สูงแสดงว่าข้อสอบนั้นสามารถจำแนกผู้สอบที่มีความสามารถแตกต่างกันได้ดี ค่าความยาก (b) มีค่าอยู่ระหว่าง -2.50 ถึง +2.50 ถ้าค่าความยาก (b) ที่อยู่ใกล้ -2.50 แสดงว่าเป็นข้อสอบที่ง่าย ส่วนค่าความยากใกล้ +2.50 แสดงว่าเป็นข้อสอบที่ยาก ค่าการเดา (c) มีค่าไม่เกิน 0.30 (ศิริชัย กาญจนวาสี, 2535) จากเกณฑ์ข้างต้นสามารถสรุปได้ว่ามีข้อสอบที่ผ่านเกณฑ์การยอมรับคุณภาพจำนวน 22 ข้อ ได้แก่ ข้อ 2, 6, 7, 8, 9, 11, 12, 13, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 และ 30

จากการวิเคราะห์ค่าสารสนเทศของแบบสอบ พบว่า แบบสอบฉบับนี้เหมาะสมกับผู้สอบที่มีพารามิเตอร์ความสามารถเท่ากับ 1.2 เพราะจะให้ข้อมูลมากที่สุดและเกิดความคลาดเคลื่อนน้อยที่สุด

3. ผลการเปรียบเทียบระดับแรงจูงใจในการสอบของนักเรียน พบว่า คะแนนที่ได้จากการรายงาน

แรงจูงใจในการสอบในภาพรวม มีค่าเฉลี่ยในระดับมาก ค่าเฉลี่ยเท่ากับ 3.86 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยแรงจูงใจในการสอบระหว่างผู้สอบที่มีความสามารถทางการเรียนแตกต่างกัน พบว่า ผู้สอบที่มีความสามารถทางการเรียนต่างกันมีค่าเฉลี่ยของแรงจูงใจในการสอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($F=31.632$, $sig = .000$) ผลการเปรียบเทียบรายคู่ของระดับแรงจูงใจในการสอบในกลุ่มผู้สอบที่มีความสามารถทางการเรียนแตกต่างกัน พบว่า ระดับแรงจูงใจในการสอบเฉลี่ยของกลุ่มผู้สอบแตกต่างกันทุกคู่ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยมีรายละเอียดดังนี้ 1) กลุ่มผู้สอบที่มีความสามารถทางการเรียนสูง (H) จะมีแรงจูงใจในการสอบสูงกว่ากลุ่มผู้สอบที่มีความสามารถทางการเรียนปานกลาง (M) และต่ำ (L) 2) กลุ่มผู้สอบที่มีความสามารถทางการเรียนปานกลาง (M) จะมีแรงจูงใจในการสอบสูงกว่ากลุ่มผู้สอบที่มีความสามารถทางการเรียนต่ำ (L)

ตาราง 4 ผลการเปรียบเทียบความแตกต่างรายคู่ของระดับแรงจูงใจในการสอบ ของนักเรียนที่มีความสามารถต่างกัน ด้วยวิธี Games - Howell

ระดับความสามารถ	ค่าเฉลี่ยแรงจูงใจในการสอบ	MEAN DIFFERENCE		
		สูง (H)	ปานกลาง (M)	ต่ำ (L)
สูง (H)	4.05	-	0.322 [*]	0.695 [*]
ปานกลาง (M)	3.91	-	-	0.373 [*]
ต่ำ (L)	3.55	-	-	-

หมายเหตุ * $p < .05$

4. ผลการศึกษาความสัมพันธ์ระหว่างแรงจูงใจในการสอบกับคะแนนสอบในการทดสอบที่มีความสำคัญต่ำและค่าการเดาของนักเรียนที่มีระดับความสามารถทางการเรียนแตกต่างกัน พบว่า

4.1 ความสัมพันธ์ระหว่างแรงจูงใจในการสอบและคะแนนสอบในการทดสอบที่มีความสำคัญต่ำในภาพรวมพบความสัมพันธ์ระหว่างแรงจูงใจในการสอบและคะแนนสอบมีความสัมพันธ์กันเชิงบวกในระดับต่ำ ($r = .297$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อพิจารณาตามกลุ่มความสามารถทางการเรียนของผู้สอบ พบว่า ในกลุ่มผู้สอบที่มีความสามารถทางการเรียนต่ำ (L) แรงจูงใจในการสอบและคะแนนสอบมีความสัมพันธ์กันเชิงบวกในระดับต่ำมาก ($r = .167$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในกลุ่มผู้สอบที่มีความสามารถทางการเรียนสูง (H) แรงจูงใจในการสอบและคะแนนสอบมีความสัมพันธ์กันเชิงบวกในระดับต่ำ ($r = .271$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และไม่พบความสัมพันธ์ระหว่างแรงจูงใจในการสอบและคะแนนสอบในกลุ่มผู้สอบที่มีความสามารถทางการเรียนปานกลาง (M)

ตาราง 5 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างแรงจูงใจในการสอบและคะแนนสอบในการทดสอบที่มีความสำคัญต่ำ

ระดับความสามารถ	ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน	Sig (2-tailed)
ต่ำ (L)	.167*	0.40
ปานกลาง (M)	.096	.081
สูง (H)	.271**	.000
Total	.297**	.000

4.2 ไม่พบความสัมพันธ์ระหว่างแรงจูงใจในการสอบและค่าการเดาในทุกกลุ่มผู้สอบที่จำแนกตามระดับความสามารถต่างๆ

5. ผลการศึกษาความสัมพันธ์ระหว่างค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) กับเกรดเฉลี่ยสะสมวิชาวิทยาศาสตร์ของนักเรียน พบว่า ค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) และเกรดเฉลี่ยสะสมวิชาวิทยาศาสตร์มีความสัมพันธ์กันเชิงบวกในระดับสูง ($r = .632$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

อภิปรายผลและข้อเสนอแนะ

1. ระดับแรงจูงใจในการสอบของนักเรียนที่มีความสามารถทางการเรียนแตกต่างกัน

ผลการวิจัย พบว่า นักเรียนรายงานแรงจูงใจในการสอบในระดับมาก ค่าเฉลี่ยเท่ากับ 3.86 สอดคล้องกับ Eklöf (2008) ที่ศึกษาแรงจูงใจในการสอบในการสอบการศึกษาแนวโน้มนำการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ระหว่างประเทศ (TIMSS) ในประเทศสวีเดน ผลการศึกษาพบว่านักเรียนชั้นมัธยมศึกษาปีที่ 2 รายงานแรงจูงใจในการสอบในระดับสูง ซึ่งแตกต่างจากงานวิจัยของ Eklöf (2010) ที่ศึกษาการรายงานแรงจูงใจในการสอบการศึกษาแนวโน้มนำการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ระหว่างประเทศขั้นสูง (TIMSS Advance) ในประเทศสวีเดน ผลการศึกษา พบว่า นักเรียนชั้นมัธยมศึกษาปีที่ 6 รายงานแรงจูงใจในการสอบอยู่ในระดับต่ำ เนื่องจากนักเรียนรับรู้ว่าการทดสอบไม่มีความสำคัญต่อนักเรียนคะแนนที่ได้ไม่ถูกนำไปใช้ในการตัดสินผลการเรียน นักเรียนส่วนใหญ่จึงไม่ตั้งใจสอบ นอกจากนั้นยังพบว่าอายุของผู้สอบน่าจะมีส่วนเกี่ยวข้องกับการรายงานแรงจูงใจในการสอบ โดยนักเรียนระดับชั้นมัธยมศึกษาปีที่ 2 จะรายงานแรงจูงใจในการสอบในระดับสูง เนื่องจากนักเรียนไม่มีประสบการณ์สอบข้อสอบจากภายนอก จึงไม่รับรู้ว่าเป็นการทดสอบที่มีความสำคัญต่ำ ส่วนนักเรียนชั้นมัธยมศึกษาปีที่ 6 รายงานแรงจูงใจในการสอบในระดับต่ำ นอกจากนี้ Wolf, Smith, and Birnbaum (1995) พบว่าระดับแรงจูงใจในการสอบอยู่ในระดับต่ำหรือสูงจะแตกต่างกันไปตามความยากของข้อสอบ และหากข้อสอบมีความท้าทายความสามารถของนักเรียน นักเรียนจะมีแรงจูงใจในการสอบสูง ส่วนข้อสอบที่จะทำให้นักเรียนแสดงความสามารถออกมาได้มากที่สุดคือข้อสอบที่มีความยากอยู่ในระดับปานกลาง

เมื่อพิจารณาการรายงานแรงงูใจในการสอบของนักเรียนที่มีความสามารถทางการเรียนแตกต่างกัน ผลการวิจัยพบว่า ระดับแรงงูใจในการสอบเฉลี่ยของผู้สอบทุกกลุ่มแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 คือ กลุ่มผู้สอบที่มีความสามารถทางการเรียนสูงจะมีแรงงูใจในการสอบสูงกว่ากลุ่มผู้สอบที่มีความสามารถทางการเรียนปานกลางและต่ำ กลุ่มผู้สอบที่มีความสามารถทางการเรียนปานกลางจะมีแรงงูใจในการสอบสูงกว่ากลุ่มผู้สอบที่มีความสามารถทางการเรียนต่ำ ผลการวิจัยเป็นไปตามสมมติฐานที่ตั้งไว้

2. ความสัมพันธ์ระหว่างแรงงูใจในการสอบกับคะแนนสอบในการทดสอบที่มีความสำคัญต่ำ

ผลการวิจัย พบว่า แรงงูใจในการสอบและคะแนนสอบมีความสัมพันธ์กันเชิงบวกในระดับต่ำ ($r = .294$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ซึ่งข้อค้นพบนี้สอดคล้องกับผลการวิจัยเกี่ยวกับแรงงูใจในการสอบตามทฤษฎีความคาดหวัง-คุณค่า (expectancy-value) ได้แก่ งานวิจัยของ Wise and Demars (2005) ที่ศึกษาพบว่า นักเรียนกลุ่มที่มีแรงงูใจในการสอบสูงมีแนวโน้มที่จะได้รับคะแนนจากการสอบสูงกว่านักเรียนกลุ่มที่มีแรงงูใจในการสอบต่ำ และการศึกษาของ Eklöf (2008) ที่ศึกษาแรงงูใจในการสอบในการสอบการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ระหว่างประเทศ (TIMSS) ในประเทศสวีเดน ผลการศึกษาพบว่า แรงงูใจในการสอบและคะแนนสอบมีความสัมพันธ์เชิงบวกในระดับต่ำ ($r=0.25$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 สอดคล้องกับการศึกษาของ Eklöf and Nyroos (2013) ได้วิเคราะห์ข้อมูลเกี่ยวกับคะแนนวิชาวิทยาศาสตร์ของการทดสอบระดับชาติของประเทศสวีเดนของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในปี 2009 พบว่าการรายงานแรงงูใจในการสอบและคะแนนสอบวิชาวิทยาศาสตร์มีความสัมพันธ์เชิงบวกอย่างมีนัยสำคัญทางสถิติ ($r = 0.25$) การรายงานแรงงูใจในการสอบและการรับรู้ความสำคัญของการสอบมีความสัมพันธ์เชิงบวกอย่างมีนัยสำคัญทางสถิติ ($r = 0.20$) และ การรายงานแรงงูใจในการสอบและความกังวลในการสอบมีความสัมพันธ์เชิงลบอย่างมีนัยสำคัญทางสถิติ ($r=-0.10$) และสอดคล้องกับ Penk, Pöhlmann, and Roppelt (2014) ที่ศึกษาบทบาทของแรงงูใจในการสอบกับคะแนนสอบของนักเรียนในการสอบข้อสอบที่มีความสำคัญต่ำในโรงเรียนนอกระบบและในระบบ ผลการวิจัยพบว่าหลังจากที่มีอัตมโนทัศน์ (self-concept) ในวิชาคณิตศาสตร์ แรงงูใจในการสอบมีความสัมพันธ์กับคะแนนสอบอย่างมีนัยสำคัญแต่อยู่ในระดับค่อนข้างต่ำ นอกจากนั้น Steinmayr and Spinath (2009) ที่ศึกษาเกี่ยวกับระดับสติปัญญา ผลสัมฤทธิ์ทางการเรียนและแรงงูใจในการสอบ พบว่าผลสัมฤทธิ์ทางการเรียนมีความสัมพันธ์กับระดับสติปัญญาและแรงงูใจในการสอบ

เมื่อพิจารณาตามกลุ่มความสามารถทางการเรียนของนักเรียน พบความสัมพันธ์ระหว่างแรงงูใจในการสอบและคะแนนสอบมีความสัมพันธ์กันทางบวกในระดับต่ำมาก ($r = .167$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในกลุ่มผู้สอบที่มีความสามารถทางการเรียนต่ำ แต่ผลการวิจัยไม่พบความสัมพันธ์ระหว่างแรงงูใจในการสอบและคะแนนสอบในกลุ่มผู้สอบที่มีความสามารถทางการเรียนสูงและปานกลางเป็นไปตามสมมติฐานที่ตั้งไว้ ซึ่งสาเหตุที่กลุ่มผู้สอบที่มีความสามารถทางการเรียนปานกลางและสูงไม่พบความสัมพันธ์ระหว่างแรงงูใจในการสอบและคะแนนสอบอาจจะเป็นเพราะนักเรียนต้องรายงานแรงงูใจ

ในการสอบ 3 ข้อคำถามต่อข้อสอบ 1 ข้อ ข้อสอบมีทั้งหมด 30 ข้อ ทำให้เกิดความเบื่อหน่ายในการรายงาน
แรงจูงใจในการสอบ รายงานแรงจูงใจโดยที่ไม่ได้ตรงกับความคิดเห็น

3. ความสัมพันธ์ระหว่างแรงจูงใจในการสอบและค่าการเดา

ผลการวิจัย พบว่า แรงจูงใจในการสอบและค่าการเดาไม่มีความสัมพันธ์กันในภาพรวมและใน
การแบ่งกลุ่มผู้สอบตามระดับความสามารถทางการเรียน งานวิจัยก่อนหน้านี้ได้ศึกษาพบว่าพฤติกรรมเดา
อย่างรวดเร็ว (rapid guessing behavior) เป็นพฤติกรรมที่บ่งบอกถึงแรงจูงใจในการสอบที่ต่ำที่ความสัมพันธ์
กับความสำเร็จของการสอบ (test stakes) พบมากในการทดสอบที่มีความสำคัญต่ำ สาเหตุที่ไม่พบ
ความสัมพันธ์ของแรงจูงใจในการสอบและค่าการเดามาจาก 1) จำนวนตัวอย่างในแต่ละกลุ่มความสามารถมี
น้อยเกินไปที่ใช้วิเคราะห์หาค่าการเดา (guessing parameter) โดยจำนวนกลุ่มตัวอย่างอย่างน้อยในการ
วิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบโมเดลโลจิสติกส์ 1 พารามิเตอร์ (1PL) ต้องใช้ตัวอย่างอย่าง
น้อย 200 คน (Downing, 2003) ในการวิจัยนี้ใช้การวิเคราะห์โมเดลโลจิสติกส์ 3 พารามิเตอร์ (3PL) จึงต้องมี
ตัวอย่างในแต่ละกลุ่มความสามารถอย่างน้อยกลุ่มละ 600 คน จึงจะให้ค่าการเดาที่ถูกต้อง 2) ความเบื่อหน่าย
จากการรายงานแรงจูงใจในการสอบ เนื่องจากนักเรียนต้องรายงานแรงจูงใจในการสอบ 3 ข้อคำถามต่อ
ข้อสอบ 1 ข้อ ข้อสอบมีทั้งหมด 30 ข้อ ทำให้เกิดความล่าช้าและเบื่อหน่าย นักเรียนจึงไม่ตั้งใจอ่านข้อคำถามและ
รายงานไม่ตรงกับความเป็นจริง โดยที่ผู้วิจัยได้นำแบบวัดแรงจูงใจในการสอบของ Eklöf (2010) ที่มี 1
องค์ประกอบ ประกอบด้วย 6 ข้อคำถาม มีค่าความเที่ยงเท่ากับ .82 และโมเดลสอดคล้องกับข้อมูลใน
ประเทศสวีเดน สโลวีเนียและนอร์เวย์ สาเหตุเนื่องมาจากการปรับแบบวัดแรงจูงใจในการสอบให้เหลือเพียง 3
ข้อคำถามซึ่งน้อยเกินไปและนำมาใช้วัดกับข้อสอบรายข้อ แม้จะมีการตรวจสอบการแปลภาษาของข้อคำถาม
และตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญในกระบวนการพัฒนาเครื่องมืออย่างดีแล้ว แต่ซึ่งอาจไม่ได้
อธิบายแรงจูงใจในการสอบได้

4. การวิเคราะห์ความสัมพันธ์ระหว่างค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) กับเกรดเฉลี่ย
สะสมวิชาวิทยาศาสตร์ ผลการวิจัย พบว่า ค่าพารามิเตอร์ความสามารถของผู้สอบ (θ) และเกรดเฉลี่ย
สะสมวิชาวิทยาศาสตร์มีความสัมพันธ์กันเชิงบวกในระดับสูง ($r = .632$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01
เป็นไปตามสมมติฐานที่ตั้งไว้ แสดงให้เห็นว่าการพิจารณาผลการเรียนของโรงเรียนมีความน่าเชื่อถือ เพราะ
นักเรียนที่มีเกรดเฉลี่ยสะสมวิชาวิทยาศาสตร์สูงจะได้คะแนนสอบสูง มีระดับความสามารถจากการวิเคราะห์
ตามทฤษฎีการตอบสนองในระดับสูง

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. หน่วยงานที่รับผิดชอบในการประเมินการทดสอบที่มีความสำคัญต่ำ จะต้องกระตุ้นให้นักเรียนเกิด
แรงจูงใจในการสอบ โดยการกระตุ้นให้นักเรียนใช้ความพยายามอย่างเต็มที่ในการทำข้อสอบ ซึ่งจะเกิดขึ้นได้
เมื่อนักเรียนเห็นคุณค่าของการสอบ กล่าวคือจะต้องมีการชี้แจงให้นักเรียนเห็นความสำคัญของการสอบ การ
สอบส่งผลต่อการเรียนและการทำงานในอนาคตของนักเรียนอย่างไร การเพิ่มผลของการสอบต่อตัวผู้สอบ

เช่น การใช้คะแนนเป็นส่วนหนึ่งในการพิจารณาการจบการศึกษา จะทำให้ผู้สอบเกิดความวิตกกังวล สามารถเป็นสาเหตุที่ทำให้ได้คะแนนที่ไม่ตรงกับความเป็นจริง กลุ่มผู้สอบที่ควรได้รับการกระตุ้นให้นักเรียนเกิดแรงจูงใจในการสอบมากที่สุดคือกลุ่มผู้สอบที่มีระดับความสามารถทางการเรียนต่ำ เพราะเป็นกลุ่มที่มีความสัมพันธ์ระหว่างแรงจูงใจในการสอบและคะแนนสอบอย่างมีนัยสำคัญทางสถิติ นั้นแสดงว่าถ้าหากไม่มีการกระตุ้นแรงจูงใจนักเรียนกลุ่มนี้จะได้รับคะแนนต่ำกว่าความจริง

นอกจากนั้นยังมีวิธีที่สามารถลดพฤติกรรมแรงจูงใจในการสอบต่ำ เช่น การเดาหรือการเฉลยข้อสอบที่มักจะเกิดขึ้นในการทดสอบที่มีความสำคัญต่ำ คือ multiple-choice computer-based tests (CBTs) แนวคิดหลักของวิธีนี้คือ เมื่อผู้เข้าสอบทำข้อสอบจะมีพฤติกรรมการแก้ปัญหา (solution behavior) ผู้เข้าสอบพยายามตัดสินใจเลือกข้อถูก หรือมีพฤติกรรมการเดาอย่างรวดเร็ว (rapid guessing behavior) ผู้เข้าสอบเลือกคำตอบอย่างรวดเร็วโดยไม่พยายามหาคำตอบที่ถูกต้อง การใช้ Response Time Effort (RTE) คือ คะแนนที่เป็นตัวแทนของความเร็วในการตอบสนองข้อสอบ (มีค่า 0 – 1, 0 หมายถึง มีความพยายามต่ำ, 1 หมายถึง มีความพยายามสูง) ในการแยกความแตกต่างระหว่างการตอบสนองทั้ง 2 แบบ โดยการกำหนดเกณฑ์ของเวลา (time threshold) ของข้อสอบแต่ละข้อ การตอบที่เกิดขึ้นอย่างรวดเร็วกว่าเกณฑ์เวลาของข้อสอบที่ตั้งไว้จะจัดว่าการเป็นเดาคำตอบ ส่วนการตอบที่ช้ากว่าเกณฑ์เวลาที่ตั้งไว้จัดเวลาเป็นพฤติกรรมการแก้ปัญหา เกณฑ์ในการกำหนดเวลามีวิธีการแตกต่างกัน รวมถึงการพิจารณาจำนวนค่าของข้อสอบแต่ละข้อ หรือวิธีการส่งข้อความเตือนจากคอมพิวเตอร์เมื่อสามารถรับรู้ว่าการเดา ซึ่งจะทำให้คะแนนมาจากความสามารถที่แท้จริงและมีความตรงมากกว่า computer-based test (Kong, Wise, Harnes, & Yang, 2006; Wise, Bhola, et. Al. 2006)

2. จากการเปรียบเทียบกรอบการประเมินของข้อสอบการศึกษาแนวโน้มนโยบายการจัดการศึกษาวิชาวิทยาศาสตร์และคณิตศาสตร์กับมาตรฐานตัวชี้วัดของหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ. 2551 พบว่า เนื้อหาที่ชี้วัดส่วนใหญ่มีความสอดคล้องกัน แต่เนื่องจากการประเมินการศึกษาแนวโน้มนโยบายการจัดการศึกษาคณิตศาสตร์ถูกนำไปใช้ในการประเมินผลกับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ทำให้มีเนื้อหาบางส่วนที่นักเรียนยังไม่ได้เรียน ซึ่งอาจจะเป็นสาเหตุที่ทำให้คะแนนที่ได้ไม่ตรงกับระดับความสามารถที่แท้จริงของนักเรียน ผลการประเมินไม่สามารถสะท้อนคุณภาพการศึกษาที่แท้จริงได้ หน่วยงานที่เกี่ยวข้องอันได้แก่ กระทรวงศึกษาธิการ สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี ควรพิจารณาว่าประเทศไทยควรเข้าร่วมโครงการประเมินนี้หรือไม่ หรือจะมีแนวทางในการแก้ปัญหาโดยการนำกรอบประเมินวิชาวิทยาศาสตร์ของข้อสอบการศึกษาแนวโน้มนโยบายการศึกษาคณิตศาสตร์และวิทยาศาสตร์มาวิเคราะห์และนำไปปรับเข้ากับหลักสูตรแกนกลางการศึกษาของประเทศไทย เพื่อให้นักเรียนศึกษาเนื้อหาให้ครบก่อนที่จะเข้ารับการประเมิน

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

1. ทฤษฎีการตอบสนองข้อสอบมาตรฐาน (standard IRT) อาจจะไม่เหมาะสำหรับการวิเคราะห์

ข้อสอบที่มีความสำคัญต่ำ Wise and Demars (2006) กล่าวว่า IRT modification เหมาะกับการวิเคราะห์ข้อสอบที่มีความสำคัญต่ำมากกว่า เพราะสามารถกำหนดโดเมนลักษณะข้อสอบสำหรับพฤติกรรม การแก้ปัญหา (solution behavior) หรือพฤติกรรมเดาอย่างรวดเร็ว (rapid guessing behavior) ได้โดยโมเดลจะสอดคล้องกับข้อมูลเชิงประจักษ์มากกว่า ในการวิจัยครั้งต่อไปควรใช้ IRT modification ในการวิเคราะห์ข้อมูลเพื่อให้ได้ผลการวิเคราะห์ที่ถูกต้อง

2. ในการวิจัยครั้งต่อไปควรศึกษาเปรียบเทียบแรงจูงใจในการสอบระหว่างเพศ และขนาดของโรงเรียน เพื่อทดสอบว่าแรงจูงใจในการสอบระหว่างเพศหญิงและเพศชายแตกต่างกันหรือไม่ และทดสอบว่าแรงจูงใจในการสอบของนักเรียนที่อยู่ในโรงเรียนขนาดแตกต่างกันจะมีแรงจูงใจในการสอบแตกต่างกันหรือไม่

3. แบบวัดแรงจูงใจในการสอบรายข้อยังไม่สอดคล้องกับข้อมูล ดังนั้นควรมีการเพิ่มจำนวนข้อคำถาม, พัฒนาแบบวัดขึ้นมาใหม่ หรืออาจพิจารณาลดจำนวนข้อสอบลงเพื่อไม่ให้นักเรียนรายงานเกิดความเบื่อหน่ายและเหนื่อยล้าในการรายงานแรงจูงใจในการสอบ

รายการอ้างอิง

ภาษาไทย

- นิพล พลกลาง (2549). *ปัจจัยที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้นมัธยมศึกษาปีที่ 2 โรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษาเพชรบูรณ์ เขต 1 ปีการศึกษา 2549*. รายงานการวิจัย. กลุ่มงานวัดและประเมินผลการศึกษา กลุ่มนิเทศ ติดตามและประเมินผลการจัดการศึกษาสำนักงานเขตพื้นที่การศึกษาเพชรบูรณ์ เขต 1.
- ศิริชัย กาญจนวาสี. (2544). *ทฤษฎีการทดสอบแนวดั้งเดิม (Classical test theory)*. (พิมพ์ครั้งที่ 4). กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2545). *ทฤษฎีการทดสอบแนวใหม่ (Modern test theory)*. (พิมพ์ครั้งที่ 2). กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2556). *สรุปผลการวิจัยโครงการ TIMSS 2011 ชั้นมัธยมศึกษาปีที่ 2*. สมุทรปราการ: บริษัท แอดวานซ์ พรินติ้ง เซอร์วิส จำกัด.
- สำนักงานเลขาธิการสภาการศึกษา (2554). *คู่มือสร้างความรู้ความเข้าใจแก่ผู้บริหาร ครู และนักเรียนเพื่อเตรียมความพร้อมรองรับการประเมินตามโครงการวิจัยนานาชาติ (PISA และ TIMSS)*. กรุงเทพมหานคร. 45/2554.
- อวยพร เรืองตระกูล. (2553). *สถิติประยุกต์ทางพฤติกรรมศาสตร์*. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ภาษาอังกฤษ

- Baumert, J. & Demmrich, A. (2001). Test motivation in the assessment of student skills: The effects of incentives on motivation and performance. *European Journal of Psychology of Education*, 16, 441–462.
- Downing S.M. (2003) Item response theory: applications of modern test theory for assessments in medical education. *Medical Education*, 37:739-745.
- Eklöf, H. (2008). Test-taking motivation on low-stakes tests: A Swedish TIMSS 2003 example. *Issues and methodologies in large-scale assessment*, IERI monograph series, Vol. 1, pp. 9-21.
- Eklöf, H. (2010). Skill and will: Test-taking motivation and assessment quality. *Assessment in Education: Principles, Policy & Practice*, 17, 345–356.
- Eklöf, H. & Nyroos, M. (2013). Pupil perceptions of national tests in terms of perceived importance, invested effort and test anxiety. *European Journal of Psychology of Education*, 28, 497-510.
- Eklöf, H., Pavešič, B. J., & Grønmo, L. S. (2014). A cross-national comparison of reported effort and mathematics performance in TIMSS advanced. *Applied Measurement in Education*, 27(1), 31-45.
- Kong,X.J., Wise,S.L., Harmes,J.C., & Yang,S. (2006). Motivational effects of praise inresponse-time based feedback: A follow-up study of the effort-monitoring CBT. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.
- Obinne, A. D. E. (2012). Using IRT in determining test item prone to guessing. *World Journal of Education*, 2(1), 91-95.
- Steinmayr, R. & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*,19, 80-90.
- Wise, S. L. & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment*, 10, 1-17.
- Wise, S. L., Bhola, D., & Yang, S. (2006). Taking the time to improve the validity of low-stakes tests: The effort-monitoring CBT. *Educational Measurement: Issues and Practice*, 25(2), 21–30.
- Wolf, L. F., & Smith, J. K. (1995). The consequence of consequence: Motivation, anxiety, and test performance. *Applied Measurement in Education*, 8, 227–242.

Wolf, L. F., Smith, J. K., & Birnbaum, M. E. (1995). Consequence of performance, test motivation, and mentally taxing items. *Applied Measurement in Education*, 8, 341–351.