

การเปรียบเทียบผลการประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบ สนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบทดสอบแบบทดสอบแบบทดสอบ เมื่อแบบสอบมีข้อสอบบางส่วนไม่เป็นอิสระต่อกัน

กมลทิพย์ ศรีหาเศษ¹

บทคัดย่อ

บทความวิจัยนี้มีวัตถุประสงค์หลักเพื่อเปรียบเทียบผลการประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบทดสอบแบบทดสอบเมื่อแบบสอบมีข้อสอบบางส่วนไม่เป็นอิสระต่อกัน ภายใต้ลักษณะของแบบสอบ ขนาดตัวอย่าง ความยาวแบบสอบ อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน และวิธีการประมาณค่าพารามิเตอร์ของข้อสอบที่ต่างกัน

ผลการวิจัย พบว่า (1) เมื่ออัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันเพิ่มขึ้น IRT มีแนวโน้มให้ค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าโอกาสในการเดาใกล้เคียงค่าพารามิเตอร์ข้อสอบที่เป็นอิสระจากกันมากกว่า TRT (2) ค่าพารามิเตอร์ของแบบสอบภายใต้การประมาณค่าด้วย IRT และ TRT ที่อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันต่างกันมีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ทุกคู่ (3) ค่าเฉลี่ยของค่าพารามิเตอร์ข้อสอบที่ประมาณค่าตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบทดสอบแบบทดสอบทุกคู่ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ และ (4) ค่า RMSE ของค่าฟังก์ชันสารสนเทศของแบบสอบภายใต้การประมาณค่าด้วย IRT มีแนวโน้มสูงกว่า TRT

คำสำคัญ: ทฤษฎีการตอบสนองข้อสอบ; ทฤษฎีการตอบสนองข้อสอบแบบทดสอบแบบทดสอบ; ทฤษฎีการตอบสนองข้อสอบแบบทดสอบ

ประเภทบทความ: บทความวิจัย

การอ้างอิง:

กมลทิพย์ ศรีหาเศษ. (2563). การเปรียบเทียบผลการประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบทดสอบแบบทดสอบ เมื่อแบบสอบมีข้อสอบบางส่วนไม่เป็นอิสระต่อกัน. *วารสารดุสิตบัณฑิตทางสังคมศาสตร์*, 10(3), 619-632.

¹ คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

2086 ถนนรามคำแหง แขวงหัวหมาก เขตบางกะปิ กรุงเทพมหานคร 10240, ประเทศไทย

อีเมล: ksrihaset@gmail.com

A Comparison of Parameter Estimation of Item Parameters under Item Response Theory and Testlet Response Theory Model with Some Items Lacking of Local Independence

Kamontip Srihasset¹

Abstract

This research article aims to compare item parameters estimates under item response theory and testlet response theory with some items lacking local independence. In this research, conditions included the difference of test types, sample group size, levels of test length, rate of increase of some items lacking local independence, and methods to estimate item parameter.

Findings are as follows: (1) IRT trend to estimate item parameters more nearly item parameters of local independence than TRT, (2) Estimated parameters of the IRT and TRT were significantly correlated in all aspects ($p < .05$), (3) Average of the item parameters estimated under 3PL IRT and the 3PL TRT were not significantly different in all aspects ($p > .05$), and (4) RMSE of test information function under IRT trend to over than TRT

Keywords: Item Response Theory; Testlet Response Theory; Testlet

Type of Article: Research Article

Cite this article as:

Srihasset, K. (2020). A Comparison of parameter estimation of item parameters under item response theory and Testlet Response Theory Model with some items lacking of local independence. *Ph.D. in Social Sciences Journal*, 10(3), 619-632.

¹ Faculty of Education, Ramkhamhaeng University
2086 Ramkhamhaeng Road, Haumark, Bangkok, Bangkok 10240, Thailand
E-mail: ksrihasset@gmail.com

ความเป็นมาและความสำคัญของปัญหา

ทฤษฎีการตอบข้อสอบ (Item Response Theory--IRT) เป็นทฤษฎีที่นิยมใช้ในการวิเคราะห์ข้อสอบในปัจจุบันแทนที่ทฤษฎีการทดสอบดั้งเดิม (Classical Test Theory--CTT) ที่ใช้กันมานาน เนื่องจาก CTT มีข้อบกพร่องในเรื่องของค่าพารามิเตอร์ (parameter) ของข้อสอบ ได้แก่ ค่าความยาก และค่าอำนาจจำแนกของข้อสอบไม่คงที่ แต่จะแปรเปลี่ยนไปตามตัวอย่างที่ทำข้อสอบนั้น (Hambleton & Swaminathan, 1985, pp. 1-2) และคะแนนของผู้สอบก็แปรเปลี่ยนไปตามค่าความยากของข้อสอบ โดยทฤษฎีการตอบข้อสอบ (IRT) มีข้อตกลงเบื้องต้นที่เข้มงวด (strong assumption) ในการใช้ทฤษฎีที่สำคัญ 4 ข้อ คือ (1) ความเป็นเอกมิติของแบบสอบ (unidimensionality) (2) ความเป็นอิสระต่อกัน (local independence) (3) โค้งแสดงความสัมพันธ์ระหว่างความสามารถของผู้สอบและโอกาสที่จะตอบข้อสอบได้ถูกต้อง (item characteristic curve) และ (4) ต้องไม่ใช่แบบสอบที่เป็น Speed Test

สิ่งที่น่าสนใจสำหรับการใช้ทฤษฎี IRT วิเคราะห์ข้อสอบปรนัยแบบเลือกตอบเมื่อมีคำถามต่อเนื่องจากสถานการณ์เดียวกัน คือ ลักษณะของข้อสอบไม่สอดคล้องกับข้อตกลงเบื้องต้นเรื่องความเป็นอิสระต่อกัน (local independence) จากความเชื่อพื้นฐานของทฤษฎีการตอบข้อสอบ (IRT) ที่กล่าวถึงพฤติกรรมที่บุคคลตอบสนองต่อข้อสอบว่าถูกกำหนดขึ้นมาด้วยลักษณะบางอย่างภายในตัวของบุคคล (latent trait) ซึ่งไม่สามารถสังเกตได้โดยตรง โดยพยายามที่จะอธิบายความสัมพันธ์ระหว่าง ลักษณะภายในตัวบุคคล (trait) กับพฤติกรรมการตอบข้อสอบ ด้วยฟังก์ชันทางคณิตศาสตร์ ภายใต้เงื่อนไขสำคัญประการหนึ่งคือ ความเป็นอิสระต่อกันในการตอบข้อสอบ (local independence) ทั้งนี้ในบริบทของสถานการณ์

การทดสอบ ในกรณีที่แบบสอบที่ใช้มีชุดข้อสอบที่มีเนื้อหาสัมพันธ์กันภายใต้เนื้อเรื่องหรือสถานการณ์ย่อยใดๆ หรือที่รู้จักกันทั่วไปว่า "เทสต์เลท (testlets)" ซึ่งถือว่าเป็นการฝ่าฝืนข้อตกลงเบื้องต้นของทฤษฎีการตอบสองข้อสอบ ซึ่งการประมาณค่าพารามิเตอร์ตามแนวคิดทฤษฎีการตอบสองข้อสอบ การประมาณค่าจะมีความถูกต้องแม่นยำมากที่สุดเมื่อข้อสอบแต่ละข้อมีความเป็นอิสระต่อกัน แต่สำหรับข้อสอบที่สัมพันธ์กันนั้น มีแนวโน้มทำให้การประมาณค่าพารามิเตอร์มีความสูงขึ้น ซึ่งการฝ่าฝืนข้อตกลงของทฤษฎีการตอบสองข้อสอบกรณีที่มีข้อมูลแบบเทสต์เลทจะนำไปสู่ความลำเอียงในการประมาณค่าพารามิเตอร์ (Sireci, Thissen, & Wainer, 1991)

จากเหตุผลที่กล่าวมาข้างต้น ผลของการฝ่าฝืนข้อตกลงเรื่องความเป็นอิสระต่อกันของข้อสอบตามแนวคิดของทฤษฎีการตอบสองข้อสอบเมื่อแบบสอบที่ใช้มีข้อสอบบางส่วนไม่เป็นอิสระต่อกันหรือที่เรียกว่าเทสต์เลท มีแนวโน้มทำให้การประมาณค่าพารามิเตอร์คลาดเคลื่อน ดังนั้น ข้อสอบปรนัยแบบเลือกตอบเมื่อมีคำถามต่อเนื่องจากสถานการณ์เดียวกัน ที่คาดว่าจะมีปัญหาข้อตกลงเบื้องต้นเรื่องความเป็นอิสระต่อกัน (local independence) ยังคาดได้อีกว่าจะส่งผลต่อความถูกต้องแม่นยำของการประมาณค่าพารามิเตอร์ ดังนั้น ในการศึกษาครั้งนี้จึงมุ่งศึกษาผลการเปรียบเทียบผลการประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสองข้อสอบและโมเดลทฤษฎีการตอบสองข้อสอบแบบเทสต์เลทเมื่อแบบสอบมีข้อสอบบางส่วนไม่เป็นอิสระต่อกัน เพื่อให้ได้สารสนเทศที่เป็นประโยชน์ต่อการวางแผนการสร้างข้อสอบภายใต้บริบทของทฤษฎีการตอบสองข้อสอบในอนาคต และเป็นการขยายองค์ความรู้และเป็นประโยชน์ต่อวงการศึกษาระดับชาติต่อไป

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์หลักเพื่อเปรียบเทียบผลการประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลทเมื่อแบบสอบมีข้อสอบบางส่วนไม่เป็นอิสระต่อกัน โดยมีจุดมุ่งหมายเฉพาะ ดังนี้

1. เพื่อวิเคราะห์ค่าพารามิเตอร์ของแบบสอบเมื่อมีข้อสอบบางส่วนไม่เป็นอิสระต่อกันตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท
2. เพื่อศึกษาความสัมพันธ์ของผลการประมาณค่าพารามิเตอร์ของแบบสอบเมื่อมีข้อสอบบางส่วนไม่เป็นอิสระต่อกันตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท
3. เพื่อเปรียบเทียบผลการประมาณค่าพารามิเตอร์ของแบบสอบเมื่อมีข้อสอบบางส่วนไม่เป็นอิสระต่อกันตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท
4. เพื่อศึกษาการเปลี่ยนแปลงฟังก์ชันสารสนเทศของแบบสอบเมื่อเพิ่มจำนวนกลุ่มข้อสอบที่ไม่เป็นอิสระต่อกันในแบบสอบ

ขอบเขตของการวิจัย

1. การวิจัยครั้งนี้ใช้วิธีการศึกษาโดยการจำลองข้อมูล เพื่อเปรียบเทียบผลการประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลทเมื่อแบบสอบมีข้อสอบบางส่วนไม่เป็นอิสระต่อกัน โดยมีตัวแปรและเงื่อนไขในการศึกษาดังนี้

ตัวแปรอิสระ คือ จำนวนกลุ่มข้อสอบที่มีคำถามต่อเนื่องจากสถานการณ์เดียวกัน

ตัวแปรตาม คือ ผลเปรียบเทียบค่าพารามิเตอร์และค่าฟังก์ชันสารสนเทศของแบบสอบ

เงื่อนไขที่ทำการศึกษา ประกอบด้วย (1) ลักษณะ

ของแบบสอบ 2 แบบ ได้แก่ แบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันหรือแบบสอบที่ไม่มีคำถามต่อเนื่องจากสถานการณ์เดียวกัน จำนวน 1 ชุด และแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันหรือแบบสอบที่มีจำนวนกลุ่มข้อสอบที่มีคำถามต่อเนื่องจากสถานการณ์เดียวกัน ซึ่งในที่นี้คือ เทสต์เลท จำนวน 20 ชุด โดยรูปแบบการตอบสนองข้อสอบ มี 2 แบบ คือ การตอบถูก (Correct--CR) ตอบผิด (Incorrect--IN) ภายใต้แบบสอบ 21 ชุด ที่ต่างกัน (2) ขนาดตัวอย่าง มี 1 ขนาด คือ 1,500 คน (3) ความยาวแบบสอบ มี 1 ระดับคือ 100 ข้อ (4) อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน มี 6 ระดับ คือ 5%, 10%, 15%, 20%, 25% และ 100% และ (5) วิธีการประมาณค่าพารามิเตอร์ของข้อสอบ มี 2 แบบ คือ การประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท

2. โมเดลการวัดที่ใช้ในการวิจัยครั้งนี้เป็นโมเดลการตอบสนองข้อสอบแบบโลจิสติก ชนิด 3 พารามิเตอร์ (Three-Parameter Logistic model--3PL) โดยมีการตรวจให้คะแนนแบบ 2 ค่า

3. แบบสอบที่ใช้ในการศึกษาครั้งนี้ เป็นแบบสอบที่มีการตรวจให้คะแนนแบบ 2 ค่า (dichotomous) คือ ผู้สอบได้ 1 คะแนนเมื่อตอบข้อสอบถูก และได้ 0 คะแนนเมื่อตอบข้อสอบผิด โดยไม่มีการลงโทษเมื่อบุคคลตอบข้อสอบผิดหรือละเว้นการตอบสนองข้อสอบ นั่นคือ ไม่มีการให้คะแนนติดลบกรณีผู้สอบตอบข้อสอบผิดหรือละเว้นการตอบสนองข้อสอบ

ประโยชน์ที่ได้รับจากการวิจัย

1. ทำให้ทราบผลการประมาณค่าพารามิเตอร์ของข้อสอบไม่เป็นอิสระต่อกันตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท

2. ทำใหทราบถึงการเปลี่ยนแปลงดัชนทที่ใชในการเปรยบเทยผลการประมาณค่าพารามมิตอร และค่าฟังก์ชันสารสนเทศของแบบสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลทเมือเพิ่มจำนวนกลุ่มข้อสอบที่ไม่เป็นอิสระต่อกันในแบบสอบ

3. ผลการวิจัยสามารถนำไปใช้เป็นข้อพิจารณาเงื่อนงายในการวิเคราะห์ข้อสอบภายใต้บริบทของทฤษฎีการตอบสนองข้อสอบ (IRT) และโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท (TRT) โดยให้มีความคลาดเคลื่อนในการวิเคราะห์น้อยที่สุด

แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง หลักการของทฤษฎีการตอบข้อสอบ (Item Response Theory–IRT)

ทฤษฎีการตอบสนองข้อสอบ เป็นโมเดลแสดงความสัมพันธ์ระหว่งการตอบสนองข้อสอบของผู้สอบและตัวแปรแฝงภายใน บ่อยครั้ง ที่ใช้คำว่า “ความสามารถ” ซึ่งวัดโดยเครื่องมือหรือแบบสอบ โดยตัวแปรแฝงภายในใช้สัญลักษณ์คือ θ เป็นโครงสร้างที่มีความเป็นเอกมิติแบบต่อเนื่อง ซึ่งอธิบายได้ด้วยความแปรปรวนร่วมระหว่งการตอบสนองข้อสอบ (Steinberg & Thissen, 1995) โดยผู้สอบที่มีระดับความสามารถ (θ) สูง จะมีความน่าจะเป็นสูงในการตอบข้อสอบถูกหรือเห็นด้วยกับข้อคำถามของแบบสอบถาม (Baker, 2001; Hambleton & Swaminathan, 1985; Reeve, 2003)

โมเดลความสัมพันธ์ของความน่าจะเป็นของการตอบสนองข้อสอบถูกในเงื่อนงายข้อสอบบนตัวแปรระดับความสามารถ (θ) นั้น ไค่งลักษณะข้อสอบ (Item Characteristic Curve–ICC) ในการประมาณค่าจากพารามมิตอรข้อสอบส่วนใหญ่จะใช้การวาดกราฟโมเดลทฤษฎีการตอบสนองข้อสอบ ในการวิจัยจะสมมุติว่าฟังก์ชันปกติสะสม (normal ogive function) หรือ

ฟังก์ชันโลจิสติก (logistic function) จะอธิบายความสัมพันธ์นี้ถูกต้องแม่นยำและสอดคล้องกับข้อมูลฟังก์ชันโลจิสติก (logistic function) ไกล่เคียงกับฟังก์ชันปกติสะสม (normal ogive function) และเป็นฟังก์ชันทางคณิตศาสตร์ที่ง่ายต่อการนำไปใช้ โดยผลที่ได้มีอิทธิพลมากในการนำไปใช้ในงานวิจัย ซึ่งไค่งลักษณะข้อสอบ (Item Characteristic Curve–ICC) สามารถให้ภาพ เช่น การวิเคราะห์ถดถอย (regression) ของคะแนนข้อสอบบนตัวแปรแฝงภายใน (Lord, 1980)

ความน่าจะเป็นในการตอบข้อสอบถูกตามเงื่อนงายข้อสอบ บนระดับความสามารถของผู้สอบ กลุ่มที่มีระดับความสามารถสูง จะมีความน่าจะเป็นของการตอบข้อสอบถูกมากกว่าตอบผิด สำหรับข้อสอบ ที่มีการตรวจให้คะแนนแบบ 2 ค่า (dichotomous) ความน่าจะเป็นของการตอบข้อสอบผิดจะสูงสำหรับกลุ่มที่มีระดับความสามารถต่ำ และจะลดลงเมือผู้สอบมีระดับความสามารถสูง ซึ่งการวัดข้อสอบมีความเชื่อในความไม่แปรเปลี่ยนถ้ากลุ่มผู้สอบแตกต่างกัน ถ้าผู้ตอบมีระดับตัวแปรแฝงภายใน (θ) เท่ากันแล้ว ความน่าจะเป็นของการเห็นด้วยในข้อคำถามควรเท่ากัน (Andrich, 1988) ในทฤษฎีการตอบสนองข้อสอบ (item response theory: IRT) ได้เสนอวิธีการสำหรับเปรยบเทียบการวัดข้อสอบข้ามกลุ่มผู้สอบ โดยโมเดลการตอบสนองข้อสอบเป็นฟังก์ชันของระดับความสามารถของผู้สอบและคุณลักษณะของข้อสอบในแบบสอบ (Andrich, 1988; Embretson & Reise, 2000) และยอมให้ข้อสอบที่ใช้ในการสอบมีความแตกต่างของฟังก์ชันข้อสอบและความไม่เท่าเทียมกันในการวัด (Embretson & Reise, 2000) ทฤษฎีการตอบสนองข้อสอบถูกนำไปประยุกต์ใช้อย่างกว้างขวางในงานวิจัยทางการศึกษาและจิตวิทยา และในปัจจุบันยังมีการนำไปประยุกต์ใช้ในสาขาสังคมวิทยาด้วย โมเดลทฤษฎีการตอบสนองข้อสอบ เป็นโมเดลทางคณิตศาสตร์ที่มีความ

สัมพันธ์กับความน่าจะเป็นของการตอบที่แสดงถึงความสามารถของผู้สอบและคุณลักษณะของข้อสอบ มาตราวัดทั่วไปมีจุดกึ่งกลางการตัดสินใจคือศูนย์ หน่วยของการวัดของค่าในหนึ่งหน่วยอยู่ในช่วง -4 ถึง +4 ซึ่งเป็นคะแนนที่เรียกว่า Z Score (Camilli & Shepard, 1994) แต่ละค่าของตัวแปรแฝงในมาตราวัดการวัดเป็นความน่าจะเป็นในการตอบข้อสอบของผู้สอบซึ่งเป็นการตรวจให้คะแนนแบบ 2 ค่า (dichotomous) นั้น ความน่าจะเป็นจะต่ำในกลุ่มผู้สอบที่มีระดับความสามารถต่ำ และสูงสำหรับกลุ่มผู้สอบที่มีระดับความสามารถสูง ซึ่งฟังก์ชันของพารามิเตอร์ความสามารถของผู้สอบ จะเป็นโค้งความน่าจะเป็นที่มีรูปร่างแบบ S-Shaped แต่มีข้อยกเว้นสำหรับข้อสอบที่มีค่าอำนาจจำแนกต่ำ ซึ่งโค้งนี้เป็นที่รู้จักกันดีในชื่อของโค้งลักษณะข้อสอบ (Item Characteristic Curve--ICC) และเป็นโค้งลักษณะ เฉพาะของข้อสอบ (Baker, 2001; Embretson & Reise, 2000)

เนื่องจากมีแนวคิดหลายแนวคิดเกี่ยวกับความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบได้ถูกต้องกับระดับความสามารถที่วัดได้โดยใช้ชุดของข้อสอบ หรือแบบสอบนั้น จึงทำให้เกิดโมเดลของโค้งแสดงลักษณะข้อสอบ (item response model) ขึ้นมาหลายโมเดล และนอกจากนี้ โมเดลที่สัมพันธ์กับทฤษฎีการตอบสนองข้อสอบโมเดลการตอบสนองข้อสอบแบบโลจิสติก (logistic model) ที่สำคัญมี 3 โมเดล คือ โมเดลการตอบสนองข้อสอบแบบโลจิสติก ชนิด 1 พารามิเตอร์ (One-Parameter Logistic model--1PL), โมเดลการตอบสนองข้อสอบแบบโลจิสติก ชนิด 2 พารามิเตอร์ (Two-Parameter Logistic model--2PL) และ โมเดลการตอบสนองข้อสอบแบบโลจิสติก ชนิด 3 พารามิเตอร์ (Three-Parameter Logistic model--3PL) (Baker, 2001; Hambleton & Swaminathan, 1985; Reeve, 2003)

หลักการของทฤษฎีการตอบข้อสอบแบบเทสต์เลท (Testlets Item Response Theory--IRT)

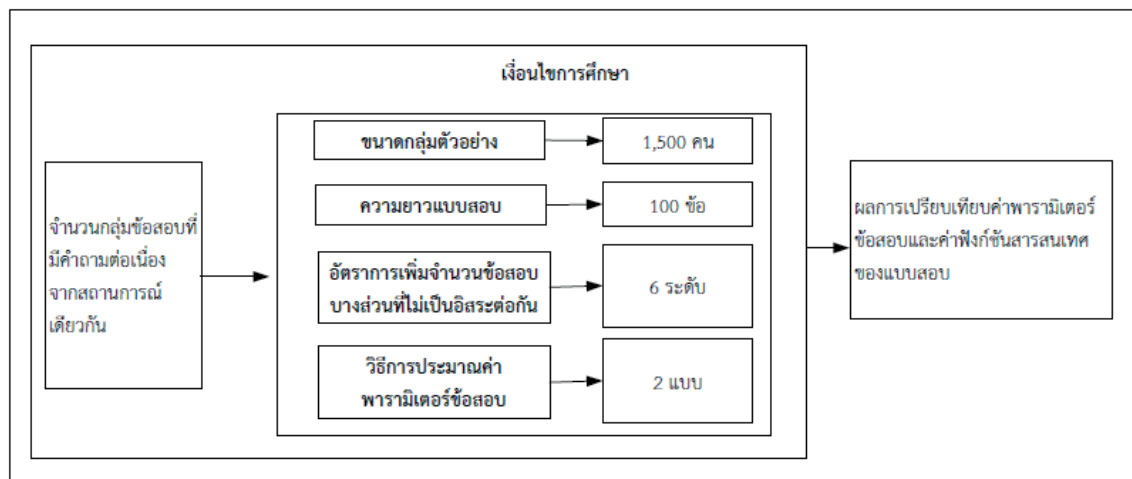
ในบริบทของสถานการณ์การทดสอบด้วยแบบสอบ ถ้าพบความสัมพันธ์ระหว่างข้อสอบจากผลการวิเคราะห์ข้อมูลแบบสอบที่ใช้ชุดข้อสอบที่มีเนื้อหาสัมพันธ์กันภายในสถานการณ์ เนื้อเรื่อง หรือเนื้อหาเดียวกัน วิธีการหนึ่งที่ใช้ในการตรวจให้คะแนนของแบบสอบประเภทนี้คือ การใช้โมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท (testlets item response theory models) และทฤษฎีการตอบสนองข้อสอบแบบพหุ (polytomous item response theory models) (Thissen, Steinberg, & Mooney, 1989; Thissen, & Orlando, 2001; Yen, 1993) ซึ่งแบบสอบเทสต์เลทเป็นกลุ่มข้อสอบที่หน่วยการให้คะแนนเป็นส่วนหนึ่งของแบบสอบทั้งฉบับ ภายในชุดข้อสอบแบบเทสต์เลท ข้อสอบแต่ละข้ออาจจะมีความเป็นอิสระต่อกันหรือสัมพันธ์กันก็ได้ (Wainer & Kiely, 1987) ตัวอย่างเช่น ในการทดสอบความสามารถในการอ่าน หัวข้อย่อยเกี่ยวกับเรื่องคำ ด้วยแบบสอบการประเมินผลทางการศึกษา (the Scholastic Assessment Test--SAT) ชุดข้อสอบแต่ละข้อที่วัดเรื่องคำจะถือเป็น 1 เทสต์เลท ซึ่งแบบสอบทั่วไปมักมีข้อสอบส่วนหนึ่งที่เป็นข้อสอบแบบเทสต์เลท การใช้โมเดลทฤษฎีการตอบสนองข้อสอบแบบพหุในการให้คะแนนแบบเทสต์เลท ข้อมูลจะถูกวิเคราะห์เรื่องความเป็นอิสระต่อกันของข้อสอบข้ามชุดแบบสอบเทสต์เลทที่แตกต่างกัน

ในการประมาณค่าความเที่ยงตามแนวคิดทฤษฎีการตอบสนองข้อสอบ การประมาณค่าจะมีความถูกต้องแม่นยำมากที่สุดเมื่อข้อสอบแต่ละข้อมีความเป็นอิสระต่อกัน แต่สำหรับข้อสอบที่สัมพันธ์กันนั้น มีแนวโน้มทำให้การประมาณค่าความเที่ยงมีค่าสูงขึ้น (Sireci et al., 1991) เมื่อข้อสอบที่แตกต่างกันสัมพันธ์กันภายใต้สถานการณ์หรือเนื้อหาเดียวกัน การรวมกลุ่มข้อสอบเข้าด้วยกันเป็นชุดแบบสอบเทสต์เลทเป็น

โมเดลที่มีโครงสร้างของแบบสอบถูกต้องและเหมาะสม การใช้วิธีการรวมข้อสอบที่สัมพันธ์กันเป็นกลุ่มชุดข้อสอบเทสต์เลท โดยชุดข้อสอบเทสต์เลทเป็นโมเดลย่อยของทฤษฎีการตอบสนองข้อสอบ ดังนั้น การศึกษาความเหมาะสมของชุดข้อสอบที่สัมพันธ์กันของโมเดลเทสต์เลทและโครงสร้างแบบสอบที่อิงแบบสอบเทสต์เลท จึงเป็นวิธีที่มักใช้ในการศึกษาข้อตกลงเรื่องความเป็นอิสระต่อกันของข้อสอบของทฤษฎีการตอบสนองข้อสอบ (Zenisky, Hambleton & Sireci, 2003) กรณีที่มีข้อมูลแบบเทสต์เลทจะนำไปสู่ความลำเอียงในการประมาณค่าพารามิเตอร์

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบผลการประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบ

สนองข้อสอบแบบเทสต์เลทเมื่อแบบสอบมีข้อสอบบางส่วนไม่เป็นอิสระต่อกัน โดยมีเงื่อนไขในการศึกษาคือ ลักษณะของแบบสอบ 2 แบบ คือ แบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันหรือแบบสอบที่ไม่มีคำถามต่อเนื่องจากสถานการณ์เดียวกัน จำนวน 1 ชุด และแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันหรือแบบสอบที่มีจำนวนกลุ่มข้อสอบที่มีคำถามต่อเนื่องจากสถานการณ์เดียวกัน จำนวน 20 ชุด รูปแบบการตอบสนองข้อสอบ มี 2 แบบ คือ การตอบถูก (Correct--CR) ตอบผิด (Incorrect--IN) ภายใต้อะไรแบบสอบ 21 ชุด ที่ต่างกัน โดยมีเกณฑ์การเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบสอบ คือ ดัชนีค่ารากที่สองของความคลาดเคลื่อนยกกำลังสองเฉลี่ย (RMSE) ดังแสดงในภาพ 1



ภาพ 1 กรอบแนวความคิดที่ใช้ในการวิจัย

ระเบียบวิธีวิจัย

การวิจัยครั้งนี้มีเงื่อนไขในการศึกษาคือ ลักษณะของแบบสอบ 2 แบบ ได้แก่ แบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันหรือแบบสอบที่ไม่มีคำถามต่อเนื่องจากสถานการณ์เดียวกัน จำนวน 1 ชุด และแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันหรือแบบสอบที่มี

จำนวนกลุ่มข้อสอบที่มีคำถามต่อเนื่องจากสถานการณ์เดียวกัน ซึ่งในที่นี้คือ เทสต์เลท จำนวน 20 ชุด ขนาดตัวอย่าง มี 1 ขนาด คือ 1,500 คน ความยาวแบบสอบ มี 1 ระดับคือ 100 ข้อ อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน มี 6 ระดับ คือ 5%, 10%, 15%, 20%, 25% และ 100% วิธีการประมาณ

ค่าพารามิเตอร์ของข้อสอบ 2 แบบ คือ การประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท โดยมีขั้นตอนการวิจัยตามลำดับ ดังนี้

1. ศึกษาค้นคว้าเกี่ยวกับทฤษฎีและหลักการ พร้อมทั้งวิธีการที่ใช้ในการประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท และงานวิจัยที่เกี่ยวข้อง จากเอกสาร หนังสือ วารสารและรายงานการวิจัยด้านการวัดผลการศึกษา ทั้งในประเทศและต่างประเทศ

2. ศึกษารายละเอียดการใช้โปรแกรมสำเร็จรูปเพื่อใช้ในการจำลองข้อมูลตามเงื่อนไขต่างๆ ที่ต้องการศึกษา และเพื่อใช้ในการประมาณค่าพารามิเตอร์

3. จำลองข้อมูลสำหรับใช้ศึกษาการประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท พร้อมตรวจสอบความถูกต้องของข้อมูลที่จำลองขึ้นตามเงื่อนไขความเป็นเอกมิติและความเป็นอิสระต่อกันของข้อสอบ

4. วิเคราะห์ข้อมูล การประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลทเมื่อแบบสอบมีข้อสอบบางส่วนไม่เป็นอิสระต่อกัน และเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบสอบที่ได้จากการประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลทด้วยดัชนีค่ารากที่สองของความคลาดเคลื่อนยกกำลังสองเฉลี่ย (Root Mean Square Error of Approximation--RMSE)

5. สรุปผลการวิเคราะห์ข้อมูล เพื่อศึกษาการประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลทเมื่อแบบสอบมีข้อสอบบางส่วน

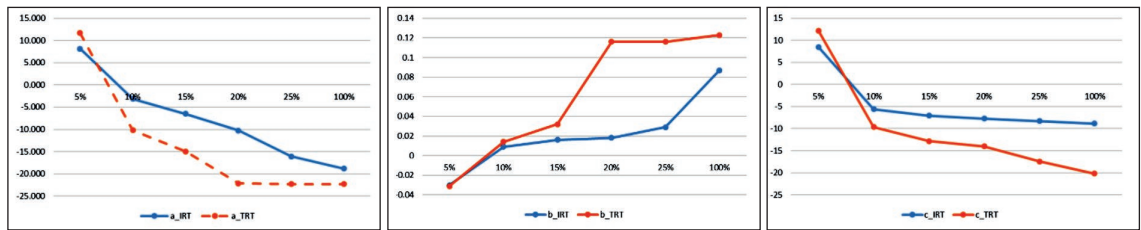
ไม่เป็นอิสระต่อกัน และเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบสอบที่ได้จากแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันหรือแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันหรือแบบสอบที่มีจำนวนกลุ่มข้อสอบที่มีคำถามต่อเนื่องจากสถานการณ์เดียวกัน

ผลการวิจัย

1. ผลการเปรียบเทียบร้อยละของการเปลี่ยนแปลงค่าพารามิเตอร์ของแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันเทียบกับแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลทที่อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันต่างกัน 6 ระดับ คือ 5%, 10%, 15%, 20%, 25% และ 100% พบว่า ร้อยละของการเปลี่ยนแปลงค่าพารามิเตอร์ของแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันเทียบกับแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลทเมื่ออัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันเพิ่มขึ้น ภายใต้การประมาณค่าตามโมเดลทฤษฎีการตอบสนองข้อสอบร้อยละของการเปลี่ยนแปลงค่าพารามิเตอร์ของแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันเทียบกับแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันมีแนวโน้มให้ค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าโอกาสในการเดาใกล้เคียงค่าพารามิเตอร์ข้อสอบที่เป็นอิสระจากกันมากกว่าโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท โดยค่าอำนาจจำแนกของข้อสอบและค่าโอกาสในการเดาภายใต้การประมาณค่าพารามิเตอร์ของแบบสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลทให้ผลสอดคล้องไปในทิศทางเดียวกัน คือเมื่ออัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันเพิ่มขึ้น

การประมาณค่าพารามิเตอร์ของแบบสอบมีแนวโน้มต่ำกว่าค่าพารามิเตอร์ของแบบสอบที่ข้อสอบที่ข้อสอบมีความเป็นอิสระต่อกัน ส่วนค่าความยากของข้อสอบเมื่ออัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อ

กันเพิ่มขึ้นการประมาณค่าพารามิเตอร์ของแบบสอบมีแนวโน้มสูงกว่าค่าพารามิเตอร์ของแบบสอบที่ข้อสอบที่ข้อสอบมีความเป็นอิสระต่อกัน ดังแสดงในภาพ 2 (ดูตาราง 1)



ภาพ 2 ผลการเปรียบเทียบร้อยละของการเปลี่ยนแปลงค่าพารามิเตอร์ของแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันเทียบกับแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบทดสอบเลขที่อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันต่างกัน 6 ระดับ

ตาราง 1

ผลการเปรียบเทียบร้อยละของการเปลี่ยนแปลงค่าพารามิเตอร์ของแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันเทียบกับแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบทดสอบเลขที่อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันต่างกัน 6 ระดับ

ค่าพารามิเตอร์ ของแบบสอบ	อัตราการเพิ่ม จำนวนข้อสอบ บางส่วนที่ไม่ เป็นอิสระต่อกัน ต่างกัน	จำนวน ข้อสอบที่ เป็นอิสระ ต่อทดสอบเลข	ร้อยละของการเปลี่ยนแปลงค่า พารามิเตอร์ของแบบสอบ IRT				ร้อยละของการเปลี่ยนแปลงค่า พารามิเตอร์ของแบบสอบ TRT			
			M	SD	Max	Min	M	SD	Max	Min
ค่าอำนาจจำแนกเมื่อมีข้อสอบ บางส่วนไม่เป็นอิสระต่อกัน (a_testlet)	5%	95:5	8.140	0.522	10.638	0.151	11.665	0.021	12.559	0.235
	10%	90:10	-3.144	0.549	8.012	-5.212	-10.188	0.489	6.933	-10.364
	15%	85:15	-6.505	0.605	5.012	-10.212	-14.961	0.489	6.933	-15.364
	20%	80:20	-10.274	0.622	4.597	-15.199	-22.126	0.691	3.517	-25.351
	25%	75:25	-16.075	0.622	2.060	-20.324	-22.326	0.683	1.981	-25.475
	100%	0:100	-18.817	0.761	0.604	-27.408	-22.326	0.720	1.182	-27.559
ค่าความยากง่ายมีข้อสอบ บางส่วนไม่เป็นอิสระต่อกัน (b_testlet)	5%	95:5	-0.030	1.012	2.308	-2.634	-0.031	0.057	0.077	-0.121
	10%	90:10	0.009	1.013	2.308	-2.634	0.014	1.090	2.946	-2.634
	15%	85:15	0.016	1.058	2.308	-2.634	0.032	1.090	2.946	-2.634
	20%	80:20	0.018	1.050	2.308	-2.634	0.116	1.110	2.967	-2.605
	25%	75:25	0.029	1.051	2.308	-2.634	0.116	1.164	2.966	-2.733
	100%	0:100	0.087	1.450	3.106	-3.066	0.123	2.254	3.374	-2.672

ตาราง 1 (ต่อ)

ค่าพารามิเตอร์ ของแบบสอบ	อัตราการเพิ่ม จำนวนข้อสอบ บางส่วนที่ไม่ เป็นอิสระต่อกัน ต่อทดสอบเลข ต่างกัน	จำนวน ข้อสอบที่ เป็นอิสระ	ร้อยละของการเปลี่ยนแปลงค่า พารามิเตอร์ของแบบสอบ IRT				ร้อยละของการเปลี่ยนแปลงค่า พารามิเตอร์ของแบบสอบ TRT			
			M	SD	Max	Min	M	SD	Max	Min
			ค่าโอกาสในการเดา	5%	95:5	8.496	0.097	9.825	-0.007	12.188
เมื่อมีข้อสอบบางส่วน	10%	90:10	-5.590	0.098	8.082	-6.062	-9.625	0.106	-4.986	-10.249
ไม่เป็นอิสระต่อกัน	15%	85:15	-7.066	0.097	6.082	-8.062	-12.835	0.106	-8.986	-14.249
(c_testlet)	20%	80:20	-7.752	0.095	3.898	-9.069	-14.046	0.115	-13.991	-20.077
	25%	75:25	-8.293	0.093	0.475	-10.070	-17.436	0.121	-16.991	-21.289
	100%	0:100	-8.846	0.088	0.685	-10.082	-20.169	0.105	-15.000	-20.250

2. ผลการวิเคราะห์ความสัมพันธ์ระหว่างค่าพารามิเตอร์ของแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันภายใต้การประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบทดสอบเลขที่อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันต่างกัน 6 ระดับ คือ 5%, 10%, 15%, 20%, 25% และ 100% พบว่าค่าพารามิเตอร์ของแบบสอบ

ที่มีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ทุกคู่ โดยค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง a_{IRT} และ a_{TRT} อยู่ในช่วง 0.867-0.949 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง b_{IRT} และ b_{TRT} อยู่ในช่วง 0.736-0.889 และ ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง c_{IRT} และ c_{TRT} อยู่ในช่วง 0.812-0.916 ตามลำดับ (ดูตาราง 2)

ตาราง 2

ผลการวิเคราะห์ความสัมพันธ์ระหว่างค่าพารามิเตอร์ของแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันภายใต้การประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบทดสอบเลขที่อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันต่างกัน 6 ระดับ

%		a_TRT		b_TRT		c_TRT		
5	a_IRT	.949*	5	b_IRT	.889*	5	c_IRT	.916*
10		.874*	10		.853*	10		.923*
15		.887*	15		.875*	15		.873*
20		.877*	20		.736*	20		.867*
25		.867*	25		.796*	25		.812*
100		.882*	100		.785*	100		.865*

สอบที่เพิ่มจำนวนข้อสอบทดสอบที่ไม่นับเป็นอิสระต่อกันภายใต้การประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบมีแนวโน้มสูง

กว่าโมเดลทฤษฎีการตอบสนองข้อสอบแบบทดสอบแบบทดสอบรายละเอียดยังคงตาราง 4

ตาราง 4

ผลการเปรียบเทียบค่า RMSE ของค่าฟังก์ชันสารสนเทศของแบบสอบด้วยอัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน 6 ระดับภายใต้การประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบทดสอบ

อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันต่างกัน	จำนวนข้อสอบที่เป็นอิสระต่อทดสอบแต่ละ	RMSE_IRT	RMSE_TRT
5%	95:5	1.704	1.673
10%	90:10	2.927	1.933
15%	85:15	4.316	4.273
20%	80:20	4.700	4.412
25%	75:25	5.705	5.659
100%	0:100	7.143	6.374

การอภิปรายผล

จากผลการวิจัยที่นำเสนอข้างต้นนั้น ผลการวิจัยดังกล่าวยังมีประเด็นที่น่าสนใจคือ ผลการวิเคราะห์ความสัมพันธ์ระหว่างค่าพารามิเตอร์ของแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันเทียบกับแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันภายใต้การประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบทดสอบที่อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันต่างกัน พบว่า มีความสัมพันธ์การสูง และผลการเปรียบเทียบค่าเฉลี่ยของค่าพารามิเตอร์ข้อสอบภายใต้การประมาณค่าตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลการตอบสนองข้อสอบแต่ละข้อสอบทุกคู่ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ทั้งนี้เป็นผลเนื่องมาจากโมเดลทฤษฎีการตอบสนองข้อสอบ เป็นโมเดลทางคณิตศาสตร์ที่มีความสัมพันธ์กับความน่าจะเป็น

ของการตอบที่แสดงถึงความสามารถของผู้สอบ และคุณลักษณะของข้อสอบ (Camilli & Shepard, 1994) โดยในการประมาณค่าความเที่ยงตามแนวคิดทฤษฎีการตอบสนองข้อสอบ การประมาณค่าจะมีความถูกต้องแม่นยำมากที่สุดเมื่อข้อสอบแต่ละข้อมีความเป็นอิสระต่อกัน แต่สำหรับข้อสอบที่สัมพันธ์กันนั้น มีแนวโน้มทำให้การประมาณค่าความเที่ยงมีค่าสูงขึ้น (Sireci et al., 1991) ส่วนทฤษฎีการตอบสนองข้อสอบแบบทดสอบเป็นแนวคิด ซึ่งขยายผลมาจากโมเดลทฤษฎีการตอบสนองข้อสอบโดย Bradlow, Wainer & Wang (1999) นอกจากนี้ ผลการวิจัยยังสอดคล้องกับงานวิจัยของ Buakaew, Wongnam, and Panhoon (2018) ที่พบว่า ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ ค่าพารามิเตอร์ความยากของข้อสอบ ค่าพารามิเตอร์โอกาสในการเดา และค่าพารามิเตอร์ความสามารถของผู้สอบ ระหว่างวิธีการ

ประมาณค่าตามโมเดลทฤษฎีการตอบสนองข้อสอบ และโมเดลทฤษฎีการตอบสนองแบบเทสต์เลท มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และมีความสัมพันธ์ที่ระดับสูงมาก

ทั้งนี้ เมื่อพิจารณาค่า RMSE ของค่าฟังก์ชันสารสนเทศของแบบสอบที่เพิ่มจำนวนข้อสอบเทสต์เลทที่ไม่เป็นอิสระต่อกันภายใต้การประมาณค่าพารามิเตอร์ของข้อสอบตามโมเดลทฤษฎีการตอบสนองข้อสอบมีแนวโน้มสูงกว่าโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท ซึ่งสอดคล้องกับผลการวิจัยของ Kim and Oslund (2015) ศึกษาผลการเปรียบเทียบโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลทในแบบสอบแบบเทสต์เลทวัดความสามารถด้านการอ่านของนักเรียน ผลการวิจัยพบว่า การฝ่าฝืนข้อตกลงเบื้องต้นเกี่ยวกับความเป็นอิสระของโมเดลการตอบสนองข้อสอบ สำหรับแบบสอบแบบเทสต์เลทโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลทให้ผลการการประมาณค่าความคลาดเคลื่อนมาตรฐานต่ำกว่าความเป็นจริง ในขณะที่โมเดลการตอบสนองข้อสอบให้ผลการประมาณค่าพารามิเตอร์ข้อสอบและความสามารถของผู้สอบสูงกว่าความเป็นจริง ดังนั้น โมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลทจึงเป็นทางเลือกที่เหมาะสมสำหรับการประเมินผลแบบสอบแบบเทสต์เลทวัดความสามารถด้านการอ่าน เนื่องจากแม้ว่าโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลทจะให้ผลการประมาณค่าพารามิเตอร์สูงเกินความเป็นจริง แต่ให้ผลการการประมาณค่าความคลาดเคลื่อนมาตรฐานต่ำกว่าความเป็นจริง

ข้อเสนอแนะ

1. การวิจัยในครั้งนี้ แม้ว่าภายใต้บริบทของทฤษฎีการตอบสนองข้อสอบนั้นข้อสอบจะยังมีความแกร่งในการฝ่าฝืนข้อตกลงเรื่องความเป็นอิสระต่อกันของข้อสอบและให้ผลการประมาณค่าพารามิเตอร์ข้อสอบใกล้เคียงข้อสอบที่เป็นอิสระจากกันมากกว่าทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท แต่เมื่อพิจารณาค่า RMSE แล้วพบว่า ค่าพารามิเตอร์ภายใต้โมเดลทฤษฎีการตอบสนองข้อสอบให้ค่าสูงกว่าโมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท ดังนั้น กรณีข้อสอบไม่เป็นอิสระจากกันควรประมาณค่าภายใต้โมเดลทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลทจะให้ผลถูกต้องแม่นยำสูงกว่าการประมาณค่าภายใต้โมเดลทฤษฎีการตอบสนองข้อสอบ

2. การนำผลการศึกษาไปใช้ในสถานการณ์การสอบระดับชาตินั้น ประเด็นที่ควรพิจารณาอย่างยิ่งคือ ขนาดตัวอย่าง การศึกษาครั้งนี้สามารถนำไปใช้ได้กับสถานการณ์ของการสอบที่มีผู้สอบ 1,000 คน ขึ้นไป รวมถึงสถานการณ์ของการสอบระดับชาติที่อาจมีผู้สอบมากกว่าหนึ่งแสนคนขึ้นไป

อย่างไรก็ตาม แม้ในสถานการณ์ของการสอบระดับชาตินั้น กลุ่มผู้สอบจะเปลี่ยนกลุ่มไม่ซ้ำกัน และการสร้างแบบสอบเพื่อใช้ในการสอบจะไม่มีแบบสอบที่ใช้สอบระดับชาติไปแล้วมาใช้ใหม่ แต่ยังมี ความจำเป็นอย่างยิ่งที่ควรพิจารณาถึงการประมาณค่าพารามิเตอร์ของข้อสอบที่มีข้อสอบที่สัมพันธ์กัน เพื่อให้ได้ค่าพารามิเตอร์ที่ใกล้เคียงกับค่าพารามิเตอร์ที่แท้จริงมากที่สุด เพื่อให้สามารถวางแผนการพัฒนาการศึกษาของประเทศมีความถูกต้องต่อไป

References

- Andrich, D. (1988). *Rasch models for measurement*. Sage.
- Baker, F. B. (2001). *The basics of item response theory* (2nd ed.). ERIC Clearinghouse on Assessment and Evaluation.
- Bradlow, E. T., Wainer, H., & Wang, X. (1999). A bayesian random effects model for testlets. *Psychometrika*, 64(2), 153-168.
- Buakaew, J., Wongnam, P., & Panhoon, S. (2018). Estimated parameters of the testlet-test under item response theory and testlet response theory model. *Suratthani Rajabhat Journal*, 5(1), 139-155. [In Thai]
- Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. Sage.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: principles and application*. Kluwer-Nijhoff.
- Kim, W., & Oslund, E. (2015). *Applying testlet response theory models to a reading comprehension test*. Proceeding of Twenty-Second Annual Meeting Society for the Scientific Study of Reading. Newport Beach, California.
- Lord, F. M. (1980). *Application of item response theory to practical testing problems*. Erlbaum.
- Reeve, B. B. (2003). *An introduction to modern measurement theory*. Retrieved from <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.207.3244&rep=rep1&type=pdf>
- Sireci, S. G., Thissen, D., & Wainer, H. (1991). On the reliability of testlet-based tests. *Journal of Educational Measurement*, 28(3), 237-247.
- Steinberg, L., & Thissen, D. (1995). Item response theory in personality research. In P. E. Shrout, & S. T. Fiske (Eds.), *Personality research, methods, and theory: A festschrift honoring Donald W. Fiske* (pp. 161-181). Erlbaum.
- Thissen, D., & Orlando, M. (2001). *Item response theory for items scored in two categories*. In D. Thissen, & H. Wainer (Eds.), *Test scoring* (pp. 73-140). Lawrence Erlbaum Associates.
- Thissen, D., Steinberg, L., & Mooney, J. (1989). Trace lines for testlets: A use of multiple categorical response models. *Journal of Educational Measurement*, 26(3), 247-260.
- Wainer, H., & Kiely, G. L. (1987). Item clusters and computerized-adaptive testing: A case for testlets. *Journal of Educational Measurement*, 24(3), 185-201.
- Yen, W. M. (1993). Scaling performance assessments: Strategies for managing local item dependence. *Journal of Educational Measurement*, 30(3), 187-213.
- Zenisky, A. L., Hambleton, R. K., & Sireci, S. G. (2003). *Effects of local item dependence on the validity of irt item, test, and ability statistics*. Association of American Medical Colleges.