

การวิเคราะห์ความเป็นเอกมิติของแบบสอบ
และความเป็นอิสระต่อกันเมื่อมีข้อสอบบางส่วนไม่เป็นอิสระต่อกัน

An Analysis of Unidimensionality and Local Independence with some Items
Lacking of Local Independence

กมลทิพย์ ศรีหาเศษ^{1*} และ สุวิมล ทิรกานนท์²

Kamontip Srihaset^{1*} and Suwimon Tirakanant²

(Received: May 28, 2020 ; Revised: July 8, 2020 ; Accepted: July 16, 2020)

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์หลักเพื่อศึกษาการเปลี่ยนแปลงของดัชนีวัดความสอดคล้องความเป็นเอกมิติเมื่อเพิ่มจำนวนกลุ่มข้อสอบที่ไม่เป็นอิสระต่อกันในแบบสอบ ภายใต้ลักษณะของแบบสอบ ขนาดกลุ่มตัวอย่าง ความยาวแบบสอบ และอัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน ที่ต่างกัน ผลการวิจัยพบว่า ดัชนีวัดความสอดคล้องความเป็นเอกมิติ ประกอบด้วย CMIN/DF RMR GFI AGFI IFI และ RMSEA ให้ผลสอดคล้องกันว่าแบบสอบที่เพิ่มจำนวนข้อสอบทดสอบเลขที่ไม่เป็นอิสระจากกันทั้ง 20 ชุด มีความเป็นเอกมิติ ทั้งนี้ เมื่อแบบสอบมีจำนวนข้อสอบทดสอบเลขที่สัมพันธ์กันเพิ่มมากขึ้นแล้วค่าดัชนีวัดความสอดคล้องความเป็นเอกมิติมีแนวโน้มลดลง

คำสำคัญ ความเป็นอิสระของข้อสอบ ความเป็นเอกมิติ เทสต์เลข

Abstract

The purpose of this study was to investigate change of goodness of fit index of unidimensionality when the number of item interaction in tests was increased. In this research, the conditions included the difference of type of test, the sample size, the test length, and the rate of mix of some items lacking local independence. The research findings, as the results of the analysis of the goodness of fit index of unidimensionality, revealed that the goodness of fit index consisted of CMIN/DF, RMR, GFI, AGFI, IFI and RMSEA, and they yielded similar goodness of fit index that all of

¹ อาจารย์ ภาควิชาการประเมินและการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

² รองศาสตราจารย์ ภาควิชาการประเมินและการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง

¹ Lecturer, Department of Evaluation and Research, Faculty of Education, Ramkhamhaeng University

² Associate Professor, Department of Evaluation and Research, Faculty of Education,
Ramkhamhaeng University

* Corresponding Author E-mail: ksrihaset@gmail.com

the 20 tests with the mixed testlet had unidimensionality. After all, when the number of item interaction in tests increased, the goodness of fit index of unidimensionality decreased.

Keywords: local independence, unidimensionality, testlet

บทนำ

การใช้แบบสอบวัดผลทางการศึกษามีจุดมุ่งหมายที่สำคัญ คือ การวัดความรู้ความสามารถเฉพาะด้าน (single ability) ของแต่ละวิชา ซึ่งเป็นคุณลักษณะแฝง (latent trait) ที่มีอยู่ในตัวของผู้สอบ ชุดของข้อสอบจะทำหน้าที่เป็นสิ่งที่ช่วยให้ผู้สอบแสดงคุณลักษณะแฝงออกมา การที่แบบสอบจะสามารถวัดความสามารถเฉพาะด้านได้แสดงว่า แบบสอบนั้นจะต้องประกอบด้วยข้อสอบที่วัดความรู้ความสามารถในลักษณะเดียวกัน หรือมีความเป็นเอกมิติ (unidimensionality) และการที่แบบสอบมีความเป็นเอกมิตินี้เองที่จะทำให้นักวัดผลสามารถแปลความหมายจากคะแนนได้ง่าย ชัดเจน และตรงไปตรงมา ซึ่งในปัจจุบันทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) นิยมใช้ในการวิเคราะห์ข้อสอบในปัจจุบันแทนที่ทฤษฎีการทดสอบดั้งเดิม (Classical Test Theory: CTT) ที่ใช้กันมานาน เนื่องจากทฤษฎีการทดสอบดั้งเดิม (CTT) มีข้อบกพร่องในเรื่องของค่าพารามิเตอร์ (parameter) ของข้อสอบ ได้แก่ ค่าความยาก และค่าอำนาจจำแนกของข้อสอบไม่คงที่ แต่จะแปรเปลี่ยนไปตามกลุ่มตัวอย่างที่ทำข้อสอบนั้น (Hambleton & Swaminathan, 1985) และคะแนนของผู้สอบก็แปรเปลี่ยนไปตามค่าความยากของข้อสอบ

เมื่อนักวัดผลได้เสนอทฤษฎีการตอบสนองข้อสอบ (IRT) ในปี 1952 ทฤษฎีนี้ก็ได้รับความสนใจจากนักวิชาการอย่างกว้างขวางแทนที่ทฤษฎีแบบดั้งเดิม (CTT) อย่างรวดเร็ว เพราะทฤษฎีแบบดั้งเดิม (CTT) เป็นทฤษฎีที่ตั้งอยู่บนข้อตกลงที่อ่อนและไม่สมเหตุสมผล (Lord, 1980 ; Allen & Yen, 1979) ส่วน ทฤษฎีการตอบสนองข้อสอบ (IRT) เป็นทฤษฎีที่กล่าวถึงพฤติกรรมที่บุคคลตอบสนองต่อข้อสอบว่าถูกกำหนดขึ้นมาจากลักษณะบางอย่างภายในตัวของบุคคล (latent trait) ซึ่งไม่สามารถสังเกตได้โดยตรง จึงพยายามที่จะอธิบายความสัมพันธ์ระหว่าง ลักษณะภายในตัวบุคคล (trait) กับพฤติกรรมการตอบข้อสอบ และอธิบายความสัมพันธ์ดังกล่าวด้วยฟังก์ชันทางคณิตศาสตร์ ในการวัดผลทางการศึกษา ลักษณะของทฤษฎีการตอบสนองข้อสอบ (IRT) จะบอกถึงความสัมพันธ์ระหว่างความสามารถที่เป็นผลมาจากการเรียนในวิชาใดวิชาหนึ่งกับโอกาสที่จะตอบข้อสอบแต่ละข้อในวิชานั้นได้ถูกต้อง

แม้ว่าทฤษฎีการตอบสนองข้อสอบ (IRT) จะมีจุดเด่นหลายประการก็ตาม แต่ก็มีข้อตกลงเบื้องต้นที่เข้มงวด (strong assumption) ในการใช้ทฤษฎีที่สำคัญ 4 ข้อ คือ 1) ความเป็นเอกมิติของแบบสอบ (unidimensionality) 2) ความเป็นอิสระต่อกัน (local independence) 3) โค้งแสดงความสัมพันธ์ระหว่างความสามารถของผู้สอบและโอกาสที่จะตอบข้อสอบได้ถูกต้อง (item characteristic curve) ซึ่งเป็น logistic function มีลักษณะเป็น monotonic increasing function และ 4) ต้องไม่ใช่แบบสอบที่เป็น speed test

สิ่งที่น่าสนใจสำหรับการใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) วิเคราะห์ข้อสอบปรนัยแบบเลือกตอบเมื่อมีคำถามต่อเนื่องจากสถานการณ์เดียวกัน คือ ลักษณะของข้อสอบไม่สอดคล้องกับข้อตกลงเบื้องต้นเรื่อง

ความเป็นอิสระต่อกัน (local independence) นอกจากนี้ นักวิชาการยังพบความสัมพันธ์ระหว่างข้อตกลงของทฤษฎีการตอบสนองข้อสอบ (IRT) เรื่องความเป็นเอกมิติ (unidimensionality) และความเป็นอิสระต่อกันในการตอบข้อสอบ (local independence) แม้ว่าในปี 1978 Warm (1978) เสนอความเห็นไว้ว่า ถ้าแบบสอบมีความเป็นเอกมิติแล้ว ก็เป็นเงื่อนไขที่เพียงพอว่าแบบสอบนั้นมีความเป็นอิสระต่อกันในการตอบข้อสอบ แต่ถ้าแบบสอบมีความเป็นอิสระต่อกันในการตอบข้อสอบแล้ว ยังไม่เพียงพอที่จะบอกได้ว่าแบบสอบนั้นมีความเป็นเอกมิติหรือไม่ ซึ่งไม่เห็นด้วยกับการนำวิธีการตรวจสอบความเป็นอิสระต่อกันมาใช้ตรวจสอบความเป็นเอกมิติ แต่ต่อมาในปี 1985 Hambleton & Swaminathan (1985) ได้กล่าวถึงความเป็นอิสระต่อกันในการตอบข้อสอบว่า สามารถใช้แสดงถึงความเป็นเอกมิติของแบบสอบได้ จึงมีนักวัดผลหลายท่านสนใจศึกษาถึงความเป็นอิสระต่อกันในการตอบข้อสอบอีกครั้งหนึ่งดังจะเห็นได้จากการศึกษา (Roznowski, Tucker & Humphreys, 1991) ที่ใช้แนวคิดของการใช้ความเป็นอิสระต่อกันในการตอบข้อสอบเพื่อบ่งชี้ความเป็นเอกมิติ

ทั้งนี้ ในบริบทของสถานการณ์การทดสอบ ในกรณีที่แบบสอบที่ใช้มีชุดข้อสอบที่มีเนื้อหาสัมพันธ์กันภายใต้เนื้อเรื่องหรือสถานการณ์ย่อยใด ๆ หรือที่รู้จักกันทั่วไปว่า “เทสต์เลท (testlets)” ซึ่งถือว่าเป็นการฝ่าฝืนข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ ซึ่งการประมาณค่าพารามิเตอร์ตามแนวคิดทฤษฎีการตอบสนองข้อสอบ การประมาณค่าจะมีความถูกต้องแม่นยำมากที่สุดเมื่อข้อสอบแต่ละข้อมีความเป็นอิสระต่อกัน แต่สำหรับข้อสอบที่สัมพันธ์กันนั้น มีแนวโน้มทำให้การประมาณค่าพารามิเตอร์มีความสูงขึ้น ซึ่งการฝ่าฝืนข้อตกลงของทฤษฎีการตอบสนองข้อสอบกรณีที่มีข้อมูลแบบเทสต์เลทจะนำไปสู่ความลำเอียงในการประมาณค่าพารามิเตอร์ (Sireci, Thissen, & Wainer, 1991)

จากผลการวิจัยที่ผ่านมาส่วนใหญ่มุ่งศึกษาแนวทางการประมาณค่าพารามิเตอร์ของแบบสอบเทสต์เลท โดยโมเดล และวิธีการต่าง ๆ หรือการประมาณค่าพารามิเตอร์ของข้อสอบเทสต์เลทที่มีรูปแบบการตอบที่ต่างกัน การเปรียบเทียบคะแนนจากการสอบด้วยแบบสอบเทสต์เลท เป็นต้น (จกมล บัวแก้ว, ไพรัตน์ วงษ์นาม และสมพงษ์ ปั่นหุ่น, 2561; Baghaei & Ravand, 2016; Cao, Lu, & Tao, 2014; Eckes. & Baghaeib, 2015; Edwards & Li Cai, 2011; Jiao, Wang, & He, 2013; Brandt, 2012; Li, 2017; Liu & Maydeu-Olivares, 2012; Rajlic, 2019; Ravand, 2015; Rubright, 2018; Setiawati, Izzaty, & Hidayat, 2018; Tao & Cao, 2016; Zhan, Liao & Bian, 2018) โดยการวิจัยดังกล่าว ไม่ได้ศึกษาผลของการเพิ่มจำนวนข้อสอบที่ไม่เป็นอิสระต่อกันว่าส่งผลต่อการเปลี่ยนแปลงของดัชนีวัดความสอดคล้องความเป็นเอกมิติหรือไม่อย่างไร การตรวจสอบความเป็นเอกมิติเมื่อมีการฝ่าฝืนข้อตกลงของทฤษฎีการตอบสนองข้อสอบจึงเป็นประเด็นสำคัญประการหนึ่งที่ผู้วิจัยมุ่งศึกษาในครั้งนี้ โดยดัชนีวัดความสอดคล้องความเป็นเอกมิติ ประกอบด้วย 3 กลุ่มสำคัญคือ 1) ค่าสถิติทดสอบไคสแควร์หรือค่าไคสแควร์สัมพัทธ์ 2) ค่าดัชนีตรวจสอบความกลมกลืน และ 3) ค่าความคลาดเคลื่อนของการประมาณค่า ซึ่งยังไม่มีงานวิจัยที่ระบุได้แน่ชัดว่าดัชนีตัวใดให้ผลถูกต้องแม่นยำมากที่สุด ในการวิจัยครั้งนี้ ผู้วิจัยจึงใช้ CMIN/DF RMR GFI AGFI IFI และ RMSEA ซึ่งครอบคลุมดัชนีวัดความสอดคล้องความเป็นเอกมิติทั้ง 3 กลุ่ม ในการตรวจสอบความเป็นเอกมิติเมื่อแบบสอบที่ใช้มีข้อสอบบางส่วนไม่เป็นอิสระต่อกัน

ในส่วนของการเป็นอิสระต่อกันในการตอบข้อสอบ (local independence) ซึ่งหมายถึง การตอบข้อสอบของผู้สอบในแต่ละข้อเป็นอิสระต่อกัน ตัวอย่างเช่น เนื้อหาของแบบสอบข้อหนึ่งต้องไม่ให้เงื่อนไขในการตอบข้ออื่น ในทางสถิติ เมื่อมีความเป็นอิสระต่อกัน ความน่าจะเป็นของรูปแบบการตอบข้อสอบจะมีค่าเท่ากับผลคูณของความน่าจะเป็นของข้อสอบแต่ละข้อ ซึ่งการที่ลำดับของข้อสอบไม่มีผลต่อการทำข้อสอบ ทำให้แบบสอบที่ได้มีความเป็นเอกมิติไปด้วย (Hambleton & Traub, 1974; Yen, 1984) วิธีการตรวจสอบความเป็นอิสระต่อกันในการตอบข้อสอบมีหลายวิธีด้วยกัน โดยส่วนใหญ่ตรวจสอบโดยใช้ Chi-Square test (Nandakumar & Stout, 1993; Van den Wollenberg, 1982; Yen, 1984) ในการวิจัยครั้งนี้ ผู้วิจัยจึงใช้ χ^2 -test ในการตรวจสอบความเป็นอิสระระหว่างข้อสอบแต่ละข้อในแบบสอบ

จากเหตุผลที่กล่าวมาข้างต้น ผลของการฝ่าฝืนข้อตกลงเรื่องความเป็นอิสระต่อกันของข้อสอบตามแนวคิดของทฤษฎีการตอบสนองข้อสอบเมื่อแบบสอบที่ใช้มีข้อสอบบางส่วนไม่เป็นอิสระต่อกันหรือที่เรียกว่า เทสต์เลท มีแนวโน้มทำให้การประมาณค่าพารามิเตอร์คลาดเคลื่อน ดังนั้น ข้อสอบปรนัยแบบเลือกตอบเมื่อมีคำถามต่อเนื่องจากสถานการณ์เดียวกัน ที่คาดว่าจะมีปัญหาข้อตกลงเบื้องต้นเรื่องความเป็นอิสระต่อกัน (local independence) ยังคาดได้อีกว่าจะส่งผลต่อข้อตกลงเบื้องต้นเรื่อง ความเป็นเอกมิติของแบบสอบ (unidimensionality) คณะผู้วิจัยจึงเห็นว่า หากข้อสอบที่มีคำถามต่อเนื่องจากสถานการณ์เดียวกัน มีความจำเป็นต่อการวัดความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ การแก้ปัญหา แล้ว ก็ควรจะได้มีการวิเคราะห์ถึงผลของคำถามในลักษณะนี้ต่อข้อตกลงเบื้องต้นทั้งสองข้อของการใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) ดังนั้น ในการศึกษานี้จึงมุ่งศึกษาการเปลี่ยนแปลงดัชนีวัดความสอดคล้องความเป็นเอกมิติเมื่อเพิ่มจำนวนกลุ่มข้อสอบที่ไม่เป็นอิสระต่อกันในแบบสอบ เพื่อให้ได้สารสนเทศที่เป็นประโยชน์ต่อการวางแผนการสร้างข้อสอบภายใต้บริบทของทฤษฎีการตอบสนองข้อสอบในอนาคต และเป็นการขยายองค์ความรู้และเป็นประโยชน์ต่อวงการการศึกษาของประเทศไทยต่อไป

ความมุ่งหมายของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์หลักเพื่อศึกษาการเปลี่ยนแปลงดัชนีวัดความสอดคล้องความเป็นเอกมิติเมื่อเพิ่มจำนวนกลุ่มข้อสอบที่ไม่เป็นอิสระต่อกันในแบบสอบ โดยมีจุดมุ่งหมายเฉพาะดังนี้

1. เพื่อประมาณค่าพารามิเตอร์ข้อสอบและฟังก์ชันสารสนเทศของแบบสอบเมื่อเพิ่มจำนวนกลุ่มข้อสอบที่ไม่เป็นอิสระต่อกันในแบบสอบ
2. เพื่อศึกษาการเปลี่ยนแปลงดัชนีวัดความสอดคล้องความเป็นเอกมิติเมื่อเพิ่มจำนวนกลุ่มข้อสอบที่ไม่เป็นอิสระต่อกันในแบบสอบ

วิธีดำเนินการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

1.1 ตัวแปรอิสระ คือ จำนวนกลุ่มข้อสอบที่มีคำถามต่อเนื่องจากสถานการณ์เดียวกัน ซึ่งในที่นี้คือ ลักษณะของแบบสอบ 2 แบบ ได้แก่

1) แบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันหรือแบบสอบที่ไม่มีคำถามต่อเนื่องจากสถานการณ์เดียวกัน จำนวน 1 ชุด ซึ่งจำลองข้อมูลขึ้นภายใต้ทฤษฎีการตอบสนองข้อสอบ

2) แบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันหรือแบบสอบที่มีจำนวนกลุ่มข้อสอบที่มีคำถามต่อเนื่องจากสถานการณ์เดียวกัน จำนวน 20 ชุด ชุดละ 5 ข้อ ซึ่งจำลองข้อมูลขึ้นภายใต้ทฤษฎีการตอบสนองข้อสอบแบบเทสต์เลท

โดยรูปแบบการตอบสนองข้อสอบ มี 2 แบบ คือ การตอบถูก (correct: CR) ตอบผิด (incorrect: IN) ภายใต้แบบสอบ 21 ชุด ที่ต่างกัน

1.2 ตัวแปรตาม คือ ดัชนีวัดความสอดคล้องความเป็นเอกมิติ และผลเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบสอบ

1.3 เงื่อนไขที่ทำการศึกษา ประกอบด้วย

- 1) ขนาดกลุ่มตัวอย่าง มี 1 ขนาด คือ 1,500 คน
- 2) ความยาวแบบสอบ มี 1 ระดับ คือ 100 ข้อ
- 3) อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน มี 6 ระดับ คือ 5% 10% 15% 20% 25% และ 100%

2. เครื่องมือที่ใช้ในการวิจัย

การจำลองข้อมูลเพื่อใช้ศึกษาการเปลี่ยนแปลงดัชนีวัดความสอดคล้องความเป็นเอกมิติเมื่อเพิ่มจำนวนกลุ่มข้อสอบที่ไม่เป็นอิสระต่อกันในแบบสอบครั้งนี้ ใช้การจำลองข้อมูลในการศึกษาการเปลี่ยนแปลงดัชนีที่ใช้ในการตรวจสอบความเป็นเอกมิติเมื่อเพิ่มจำนวนกลุ่มข้อสอบที่ไม่เป็นอิสระต่อกันในแบบสอบที่ต่างกันทั้งสิ้น 21 ชุด ขนาดกลุ่มตัวอย่าง จำนวน 1 ขนาด ความยาวแบบสอบ จำนวน 1 ระดับ และอัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน จำนวน 6 ระดับโดยใช้โปรแกรม R โดยขั้นตอนของการจำลองข้อมูลประกอบไปด้วย 2 ขั้นตอน คือ

2.1 สร้างข้อสอบที่มีความเป็นอิสระต่อกันและชุดข้อสอบเทสต์เลทที่ไม่เป็นอิสระต่อกันค่าพารามิเตอร์ที่แท้จริงของข้อสอบ และแบบแผนการตอบของผู้สอบแต่ละคน ภายใต้ทฤษฎีการตอบสนองข้อสอบ ตามโมเดลการตอบสนองข้อสอบแบบโลจิสติก ชนิด 3 พารามิเตอร์ (3PL) โดยมีการตรวจให้คะแนนแบบ 2 ค่า

2.2 สร้างตัวแปรทำนายและจัดชุดข้อสอบที่ไม่เป็นอิสระต่อกันแทรกในแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันภายใต้เงื่อนไขการศึกษา

ตาราง 1 ค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบที่ใช้ในการจำลองข้อมูลเพื่อการศึกษา

ค่าพารามิเตอร์	ลักษณะการแจกแจง
ค่าพารามิเตอร์ของข้อสอบภายใต้โมเดล IRT	
- ค่าอำนาจจำแนกของข้อสอบ (a)	lognormal(0,0.5)
- ค่าความยากของข้อสอบ (b)	N(0,1)
- ค่าโอกาสในการเดา (c)	uniform(0,0.3)

ค่าพารามิเตอร์	ลักษณะการแจกแจง
ค่าพารามิเตอร์ของข้อสอบทดสอบทดสอบได้โมเดล TRT	
- ค่าอำนาจจำแนกของข้อสอบ (a)	lognormal(0,0.5)
- ค่าความยากของข้อสอบ (b)	N(0,1)
- ค่าโอกาสในการเดา (c)	uniform(0,0.3)
- อิทธิพลสุ่มของข้อสอบ (γ)	N(0,1)
ค่าพารามิเตอร์ความสามารถของผู้สอบ (θ)	N(0,1)

3. การเก็บรวบรวมข้อมูล

การตรวจสอบความถูกต้องของข้อมูลและเงื่อนไขที่จำลองขึ้น มีรายละเอียดดังนี้

3.1 การตรวจสอบความเป็นเอกมิติ ผู้วิจัยตรวจสอบความเป็นเอกมิติซึ่งเป็นข้อตกลงเบื้องต้นที่สำคัญของทฤษฎีการตอบสนองข้อสอบที่มีการตรวจให้คะแนนแบบ 2 ค่า โดยการคำนวณอัตราส่วนของค่าไอเกน (Eigen ratio: ER) ขององค์ประกอบที่ 1 กับองค์ประกอบที่ 2 ที่ได้จากการวิเคราะห์องค์ประกอบमुखสำคัญ (Principal factor analysis) ค่าที่ได้ไม่ควรต่ำกว่า 3 (วรณูช แหยมแสง, 2537) หากอัตราส่วนที่ได้มีค่าต่ำกว่า 3 ผู้วิจัยจะทำการจำลองสถานการณ์นั้นใหม่

3.2 แบบสอบที่ข้อสอบเป็นอิสระต่อกัน ตรวจสอบความเป็นอิสระต่อกันของข้อสอบ ด้วย χ^2 -test ระหว่างข้อสอบแต่ละคู่ในแบบสอบ 1 ชุด จำนวน 100 ข้อ จะมีการทดสอบทั้งสิ้น $\frac{n(n-1)}{2}$ คู่ หากพบความสัมพันธ์ระหว่างข้อสอบ ผู้วิจัยจะทำการจำลองสถานการณ์นั้นใหม่

3.3 แบบสอบที่ข้อสอบสัมพันธ์กัน หรือทดสอบทดสอบได้ ตรวจสอบความเป็นอิสระต่อกันของข้อสอบ ด้วย χ^2 -test ระหว่างข้อสอบแต่ละคู่ในแบบสอบทดสอบได้ 1 ชุด ๆ ละ 5 ข้อ จะมีการทดสอบทั้งสิ้นแบบสอบทดสอบได้ละ $\frac{n(n-1)}{2}$ คู่ หากไม่พบความสัมพันธ์ระหว่างข้อสอบในแบบสอบทดสอบได้ ผู้วิจัยจะทำการจำลองสถานการณ์นั้นใหม่

3.4 การตรวจสอบความถูกต้องของโปรแกรมที่พัฒนาขึ้นเพื่อใช้ในการจำลองข้อมูล ผู้วิจัยตรวจสอบความถูกต้องของโปรแกรมที่พัฒนาขึ้นโดยผู้เชี่ยวชาญที่มีความรู้และเชี่ยวชาญเกี่ยวกับการจำลองข้อมูลโดยใช้โปรแกรม R ร่วมกับการตรวจสอบด้วย package 'Rcmdr' และโปรแกรมสำเร็จรูปทางสถิติเพื่อตรวจสอบความถูกต้องของค่าที่ได้ ซึ่งโปรแกรมที่พัฒนาขึ้นมีความถูกต้องก็ต่อเมื่อได้ผลการคำนวณที่มีค่าเท่ากันทั้งสามโปรแกรม หากผลการคำนวณที่ได้มีค่าเท่ากันแสดงว่าโปรแกรมที่ผู้วิจัยพัฒนาขึ้นมีความถูกต้อง แต่ถ้าผลการคำนวณที่ได้มีค่าไม่เท่ากัน ผู้วิจัยจะตรวจหาที่ผิดพลาดของโปรแกรมที่สร้างขึ้น แล้วดำเนินการแก้ไขจนได้ผลการคำนวณจากทั้งสามโปรแกรมเท่ากัน ซึ่งผลการตรวจสอบที่ได้พบว่าทั้งสามโปรแกรมให้ผลการคำนวณที่เท่ากัน

3.5 การตรวจสอบความถูกต้องของโปรแกรมที่พัฒนาขึ้นสำหรับตรวจสอบดัชนีวัดความสอดคล้องความเป็นเอกมิติและผลเปรียบเทียบกับค่าฟังก์ชันสารสนเทศของแบบสอบ ผู้วิจัยใช้หลักการตรวจสอบเช่นเดียวกับการ

การตรวจสอบความถูกต้องของโปรแกรมที่ใช้ในการจำลองข้อมูล นั่นคือ ผู้วิจัยตรวจสอบความถูกต้องของโปรแกรมที่พัฒนาขึ้นภายใต้โปรแกรม R ร่วมกับการตรวจสอบด้วย package ‘Rcmdr’ และโปรแกรมสำเร็จรูปทางสถิติเพื่อตรวจสอบความถูกต้องของค่าที่ได้ ซึ่งโปรแกรมที่พัฒนาขึ้นมีความถูกต้องก็ต่อเมื่อได้ผลการคำนวณที่มีค่าเท่ากันทั้งสามโปรแกรม หากผลการคำนวณที่ได้มีค่าเท่ากันแสดงว่าโปรแกรมที่ผู้วิจัยพัฒนาขึ้นมีความถูกต้อง แต่ถ้าผลการคำนวณที่ได้มีค่าไม่เท่ากัน ผู้วิจัยจะตรวจหาที่ผิดพลาดของโปรแกรมที่สร้างขึ้นแล้วดำเนินการแก้ไขจนได้ผลการคำนวณจากทั้งสามโปรแกรมเท่ากัน ซึ่งผลการตรวจสอบที่ได้พบว่าทั้งสามโปรแกรมให้ผลการคำนวณที่เท่ากัน

4. การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลที่ได้จากการจำลองข้อมูลและเกณฑ์ที่ใช้ในการตรวจสอบดัชนีวัดความสอดคล้องความเป็นเอกมิติ และผลเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบสอบ มีรายละเอียดดังนี้

4.1 ความเป็นเอกมิติของแบบสอบ ตรวจสอบด้วย eigen value test และพิจารณาอัตราส่วนของค่า eigen ที่ 1 กับค่า eigen ที่ 2 มากกว่า 3 ถือว่าแบบสอบนั้นมีความเป็นเอกมิติซึ่งถือว่าเป็นข้อตกลงเบื้องต้นที่สำคัญของทฤษฎีการตอบสนองข้อสอบ

4.2 ความเป็นอิสระต่อกันของข้อสอบ ตรวจสอบด้วย χ^2 -test ระหว่างข้อสอบแต่ละคู่ในแบบสอบ 1 ชุด จำนวน 100 ข้อ

4.3 ฟังก์ชันสารสนเทศของข้อสอบ วิเคราะห์ฟังก์ชันสารสนเทศของข้อสอบเพื่อใช้วิเคราะห์ฟังก์ชันสารสนเทศของแบบสอบ ด้วยสูตร

$$I_i(\theta) = \frac{(P'_i(\theta))^2}{P_i(\theta)Q_i(\theta)}, i=1, 2, \dots, k$$

เมื่อ $I_i(\theta)$ แทนค่าฟังก์ชันสารสนเทศหรือค่าสารสนเทศที่ได้รับจากข้อสอบข้อที่ i สำหรับผู้สอบที่มีความสามารถ θ

$P'_i(\theta) = P'_i$ แทนความชันของฟังก์ชันการตอบสนองข้อสอบข้อที่ i ณ ตำแหน่งความสามารถ θ

$P_i(\theta) = P_i$ แทนความน่าจะเป็นที่ผู้ตอบซึ่งมีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

$$Q_i(\theta) = Q_i = 1 - P_i(\theta)$$

4.4 ฟังก์ชันสารสนเทศของแบบสอบ วิเคราะห์ฟังก์ชันสารสนเทศของแบบสอบเพื่อใช้พิจารณาความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบ ด้วยสูตร

$$I(\theta) = \sum_{i=1}^k I_i(\theta), i=1, 2, \dots, k$$

เมื่อ $I(\theta)$ แทนค่าฟังก์ชันสารสนเทศหรือค่าสารสนเทศที่ได้รับจากแบบสอบสำหรับ
ผู้สอบที่มีความสามารถ θ

4.5 การประมาณค่าความคลาดเคลื่อน ความถูกต้องแม่นยำของการประมาณค่าฟังก์ชัน
สารสนเทศของแบบสอบใช้เกณฑ์ดัชนีค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (root-mean square
error: RMSE) ด้วยสูตร

$$RMSE(\theta_k) = \sqrt{\frac{\sum (\theta - \hat{\theta})^2}{n_k}}$$

เมื่อ θ แทน ค่าพารามิเตอร์ที่แท้จริง ณ ระดับความสามารถที่ k (-4.0, -3.9, ..., 4.0)

$\hat{\theta}$ แทน ค่าพารามิเตอร์ที่ได้จากการประมาณค่า

n_k แทน จำนวนข้อมูลจำลอง ณ ระดับความสามารถที่ k

4.6 การตรวจสอบดัชนีวัดความสอดคล้องความเป็นเอกมิติและการเปรียบเทียบค่าสถิติที่ได้จาก
การเพิ่มจำนวนชุดข้อสอบทดสอบที่มีความสัมพันธ์กันที่ละชุด ๆ ละ 5 ข้อ ในแบบสอบและความเป็นอิสระต่อกัน
กัน ผู้วิจัยดำเนินการประมาณค่าสถิติต่าง ๆ ได้แก่ ดัชนีวัดความสอดคล้อง (goodness – of – fit index)
CMIN/DF RMR GFI AGFI IFI และ RMSEA ค่าฟังก์ชันสารสนเทศของแบบสอบ และค่าความคลาดเคลื่อน
(RMSE) โดยใช้ค่าสถิติที่ได้จากเงื่อนไขการศึกษาทั้ง 126 เงื่อนไข

4.7 เปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบสอบเมื่อเพิ่มจำนวนชุดข้อสอบทดสอบที่มีความสัมพันธ์กันด้วยระดับอัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันต่างกัน 6 ระดับ คือ ระดับ
คือ 5% 10% 15% 20% 25% และ 100% ในแบบสอบและความเป็นอิสระต่อกัน ด้วยดัชนีค่ารากที่สองของ
ความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) โดยมีหลักเกณฑ์ในการพิจารณาจากการเปรียบเทียบค่าดัชนี RMSE
ที่ได้โดยการเขียนกราฟเส้นเปรียบเทียบ ถ้าแบบสอบที่เมื่อเพิ่มจำนวนชุดข้อสอบทดสอบที่มีความสัมพันธ์กัน
ชุดใด มีค่าดัชนี RMSE น้อยกว่า แสดงว่าแบบสอบที่เมื่อเพิ่มจำนวนชุดข้อสอบทดสอบที่มีความสัมพันธ์กันนั้น
มีคุณภาพในการประมาณค่าพารามิเตอร์ข้อสอบใกล้เคียงความเป็นจริง สะท้อนให้เห็นถึงความคลาดเคลื่อน
ของการประมาณค่าพารามิเตอร์ข้อสอบมีน้อย แต่ถ้าแบบสอบที่เมื่อเพิ่มจำนวนชุดข้อสอบทดสอบที่มีความสัมพันธ์กันชุดใดมีค่าดัชนี RMSE มากกว่าแสดงว่าแบบสอบที่เมื่อเพิ่มจำนวนชุดข้อสอบทดสอบที่มีความสัมพันธ์กันนั้นมีคุณภาพในการประมาณค่าพารามิเตอร์ข้อสอบต่ำ สะท้อนให้เห็นถึงความคลาดเคลื่อน
ของการประมาณค่าพารามิเตอร์ข้อสอบมีมาก

4.8 สรุปผลวิเคราะห์ดัชนีวัดความสอดคล้องความเป็นเอกมิติ และผลเปรียบเทียบค่าฟังก์ชัน
สารสนเทศของแบบสอบ

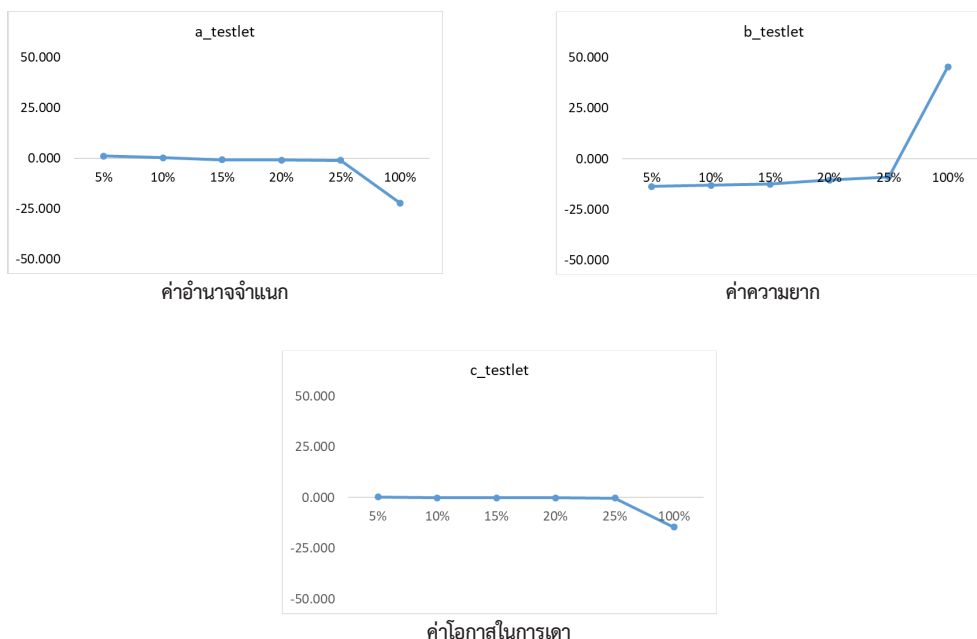
ผลการวิจัย

1. ผลการเปรียบเทียบร้อยละของการเปลี่ยนแปลงค่าพารามิเตอร์ของแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันเทียบกับแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกัน

ผลการเปรียบเทียบร้อยละของการเปลี่ยนแปลงค่าพารามิเตอร์ของแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันเทียบกับแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกัน พบว่า เมื่ออัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันเพิ่มขึ้น ร้อยละของการเปลี่ยนแปลงค่าพารามิเตอร์ของแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันเทียบกับแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันมีแนวโน้มให้ค่าอำนาจจำแนกของข้อสอบและค่าโอกาสในการเดาคำกว่าแบบสอบที่ข้อสอบเป็นอิสระต่อกัน ส่วนค่าความยากของข้อสอบมีแนวโน้มสูงกว่าแบบสอบที่ข้อสอบเป็นอิสระต่อกัน โดยช่วงของร้อยละของการเปลี่ยนแปลงค่าพารามิเตอร์ของแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันเทียบกับแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันมีช่วงกว้างขึ้น รายละเอียดปรากฏดังตาราง 2 และภาพ 1 ตามลำดับ

ตาราง 2 ภาพรวมของผลการเปรียบเทียบร้อยละของการเปลี่ยนแปลงค่าพารามิเตอร์ของแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันเทียบกับแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันที่อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันต่างกัน 6 ระดับ

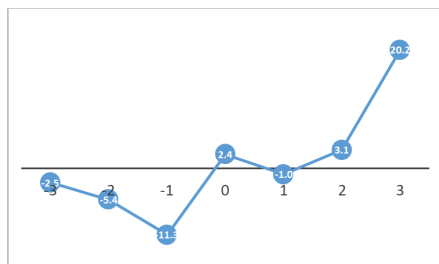
ค่าพารามิเตอร์ของแบบสอบ	อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันต่างกัน	จำนวนข้อสอบที่เป็นอิสระต่อเทสต์เลต	ร้อยละของการเปลี่ยนแปลงค่าพารามิเตอร์ของแบบสอบ			
			M	SD	Max	Min
ค่าอำนาจจำแนกเมื่อมีข้อสอบบางส่วนไม่เป็นอิสระต่อกัน (a_testlet)	5%	95:5	1.227	0.664	4.781	-0.607
	10%	90:10	0.269	0.894	5.272	-3.280
	15%	85:15	-0.682	0.888	3.346	-5.206
	20%	80:20	-0.818	0.978	5.404	-5.407
	25%	75:25	-1.022	1.074	5.747	-6.027
	100%	0:100	-22.088	1.421	4.056	-24.468
ค่าความยากง่ายมีข้อสอบบางส่วนไม่เป็นอิสระต่อกัน (b_testlet)	5%	95:5	-13.602	0.927	-10.754	-16.362
	10%	90:10	-13.082	1.247	-7.440	-18.364
	15%	85:15	-12.486	1.090	-6.248	-18.510
	20%	80:20	-10.434	1.089	-6.019	-18.700
	25%	75:25	-8.929	1.138	-4.031	19.654
	100%	0:100	45.409	1.530	60.363	-4.848
ค่าโอกาสในการเดาเมื่อมีข้อสอบบางส่วนไม่เป็นอิสระต่อกัน (c_testlet)	5%	95:5	0.278	0.103	0.592	-0.214
	10%	90:10	-0.037	0.139	0.590	-0.821
	15%	85:15	-0.076	0.148	0.629	-0.848
	20%	80:20	-0.104	0.146	0.918	-0.893
	25%	75:25	-0.153	0.161	0.961	-0.894
	100%	0:100	-14.600	0.130	-1.645	-17.914



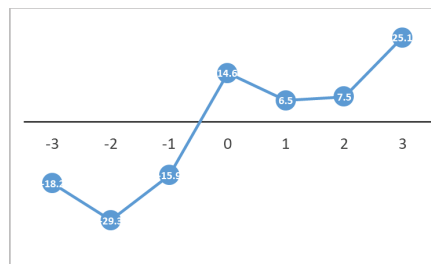
ภาพ 1 ภาพรวมของผลการเปรียบเทียบร้อยละของการเปลี่ยนแปลงค่าพารามิเตอร์ของแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกันเทียบกับแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันที่อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันต่างกัน 6 ระดับ

2. ผลการเปรียบเทียบร้อยละของการเปลี่ยนแปลงฟังก์ชันสารสนเทศของแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันกับแบบสอบที่มีข้อสอบทดสอบเลขที่ไม่เป็นอิสระต่อกัน

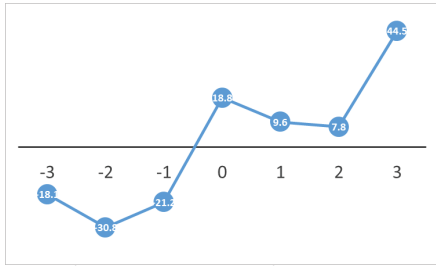
ผลการเปรียบเทียบร้อยละของการเปลี่ยนแปลงฟังก์ชันสารสนเทศของแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกัน พบว่า เมื่อแบบสอบมีจำนวนข้อสอบที่ไม่เป็นอิสระต่อกันตั้งแต่ 20 ข้อขึ้นไป ร้อยละของการเปลี่ยนแปลงค่าฟังก์ชันสารสนเทศของแบบสอบที่เพิ่มจำนวนข้อสอบทดสอบเลขที่ระดับความสามารถ 0 ถึง 3 สูงกว่าแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกัน ส่วนที่ระดับความสามารถ -3 ถึง -1 ร้อยละของการเปลี่ยนแปลงค่าฟังก์ชันสารสนเทศของแบบสอบที่เพิ่มจำนวนข้อสอบทดสอบเลขต่ำกว่าแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกัน รายละเอียดปรากฏดังภาพ 2



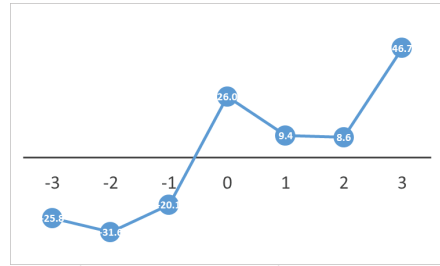
อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน 5%



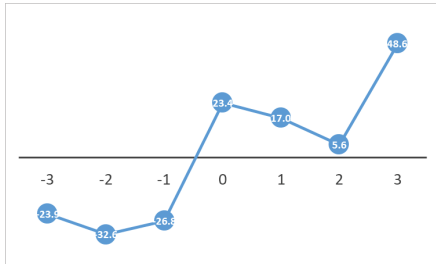
อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน 10%



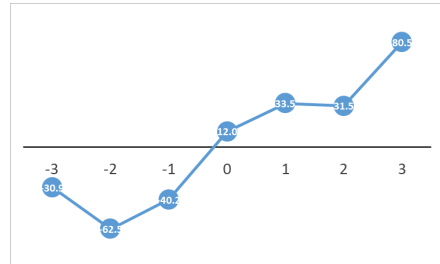
อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน 15%



อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน 20%



อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน 25%



อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน 100%

ภาพ 2 ผลการเปรียบเทียบร้อยละของการเปลี่ยนแปลงฟังก์ชันสารสนเทศของแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันกับแบบสอบที่เพิ่มจำนวนข้อสอบทดสอบทดสอบที่ไม่เป็นอิสระต่อกัน

3. ผลการเปรียบเทียบค่า RMSE ของค่าฟังก์ชันสารสนเทศของแบบสอบ

ในการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบสอบที่ข้อสอบมีความเป็นอิสระต่อกันกับแบบสอบที่ข้อสอบบางส่วนไม่เป็นอิสระต่อกัน ผู้วิจัยใช้ดัชนี RMSE เพื่อสะท้อนถึงความคงที่ (stability) ของค่าฟังก์ชันสารสนเทศของแบบสอบที่ประมาณค่าได้ โดยมีหลักเกณฑ์ในการพิจารณาจากการเปรียบเทียบค่าที่ได้คือ ถ้ามีค่าดัชนี RMSE น้อยกว่า แสดงว่ามีความคงที่ของการประมาณค่าพารามิเตอร์สูงกว่า สะท้อนให้เห็นถึงความคลาดเคลื่อนของการประมาณค่าพารามิเตอร์มีน้อย แต่ถ้ามีค่าดัชนี RMSE มากกว่า แสดงว่ามีความคงที่ของการประมาณค่าพารามิเตอร์ต่ำ สะท้อนให้เห็นถึงความคลาดเคลื่อนของการประมาณค่าพารามิเตอร์มีมาก โดยผลการเปรียบเทียบค่า RMSE ของค่าฟังก์ชันสารสนเทศของแบบสอบ พบว่า ค่า RMSE ของค่าฟังก์ชันสารสนเทศของแบบสอบที่มีข้อสอบทดสอบที่ไม่เป็นอิสระต่อกันด้วยอัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน ทั้ง 6 ระดับมีค่ามากกว่าศูนย์ ซึ่งผลการเปรียบเทียบค่า RMSE ของค่าฟังก์ชันสารสนเทศของแบบสอบที่เพิ่มจำนวนข้อสอบทดสอบที่ไม่เป็นอิสระต่อกันด้วยอัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน ทั้ง 6 ระดับ ให้ผลสอดคล้องไปในทิศทางเดียวกันคือ เมื่อข้อสอบมีจำนวนข้อสอบทดสอบที่สัมพันธ์กันจะให้ผลการประมาณค่าสูงเกินความเป็นจริง และเมื่ออัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันเพิ่มขึ้น หรือข้อสอบมีจำนวนข้อสอบทดสอบที่สัมพันธ์กันเพิ่มมากขึ้นแล้ว ค่าฟังก์ชันสารสนเทศของแบบสอบมีแนวโน้มให้ผลการประมาณค่าสูงเกินความเป็นจริงเพิ่มขึ้นด้วย รายละเอียดปรากฏดังตาราง 3

ตาราง 3 ผลการเปรียบเทียบค่า RMSE ของค่าฟังก์ชันสารสนเทศของแบบสอบด้วยอัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน 6 ระดับ

อัตราการเพิ่มจำนวนข้อสอบบางส่วน ที่ไม่เป็นอิสระต่อกันต่างกัน	จำนวนข้อสอบที่เป็นอิสระต่อแต่ละเลข	RMSE
5%	95:5	0.706
10%	90:10	1.526
15%	85:15	2.411
20%	80:20	2.612
25%	75:25	3.685
100%	0:100	7.512

4. ผลการตรวจสอบดัชนีวัดความสอดคล้องความเป็นเอกมิติ

ผลการตรวจสอบดัชนีวัดความสอดคล้องความเป็นเอกมิติของแบบสอบที่เพิ่มจำนวนข้อสอบแต่ละเลขที่ไม่เป็นอิสระต่อกันด้วยอัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน ทั้ง 6 ระดับ พบว่า ดัชนีวัดความสอดคล้องความเป็นเอกมิติที่บ่งชี้ความสอดคล้องและความกลมกลืนของตัวแบบกับข้อมูลเชิงประจักษ์ทั้ง 8 ค่า สอดคล้องตามเกณฑ์ให้ผลสอดคล้องกัน ว่าแบบสอบที่เพิ่มจำนวนข้อสอบแต่ละเลขที่ไม่เป็นอิสระต่อกันด้วยอัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน ทั้ง 6 ระดับ มีความเป็นเอกมิติทุกเงื่อนไขการศึกษา ทั้งนี้ ผลการตรวจสอบการเปลี่ยนแปลงค่าดัชนีวัดความสอดคล้องความเป็นเอกมิติพบว่าเมื่อด้วยอัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันเพิ่มขึ้น หรือข้อสอบมีจำนวนข้อสอบแต่ละเลขที่สัมพันธ์กันเพิ่มมากขึ้นแล้ว ค่าดัชนีวัดความสอดคล้องความเป็นเอกมิติมีค่าลดลง รายละเอียดปรากฏดังตาราง 4

ตาราง 4 ผลการตรวจสอบดัชนีวัดความสอดคล้องความเป็นเอกมิติของแบบสอบที่เพิ่มจำนวนข้อสอบแต่ละเลขที่ไม่เป็นอิสระต่อกันด้วยอัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกัน 6 ระดับ คือ 5% 10% 15% 20% 25% และ 100%

อัตราการเพิ่ม จำนวนข้อสอบ บางส่วนที่ไม่เป็น อิสระต่อกัน ต่างกัน	จำนวนข้อสอบที่ เป็นอิสระต่อ แต่ละเลข	ผลการตรวจสอบความเป็นเอกมิติ							
		chi- square	df	CMIN/df	RMR	GFI	AGFI	IFI	RMSEA
0%	100:0	3803.965	4850	0.784	0.005	0.967	0.964	0.966	0.010
5%	95:5	3836.222	4850	0.791	0.005	0.961	0.958	0.960	0.010
10%	90:10	3698.362	4850	0.763	0.005	0.954	0.951	0.953	0.011
15%	85:15	3802.820	4850	0.784	0.005	0.948	0.945	0.947	0.012
20%	80:20	3816.108	4850	0.787	0.005	0.942	0.939	0.941	0.012
25%	75:25	3923.070	4850	0.809	0.005	0.935	0.932	0.934	0.013
100%	0:100	6044.601	4850	1.246	0.005	0.921	0.918	0.933	0.015

อภิปรายผล

จากผลการวิจัยที่นำเสนอข้างต้นนั้น ผลการวิจัยดังกล่าวยังมีประเด็นที่น่าสนใจดังต่อไปนี้

1. ค่า RMSE ของค่าฟังก์ชันสารสนเทศของแบบสอบที่มีข้อสอบทดสอบเลขที่ไม่เป็นอิสระต่อกันทั้ง 20 ชุด มีค่ามากกว่าศูนย์ ซึ่งสะท้อนให้เห็นว่าแบบสอบที่มีข้อสอบทดสอบเลขที่ไม่เป็นอิสระต่อกันให้ผลการประมาณค่าสูงเกินความเป็นจริง และผลการประมาณค่าที่สูงเกินจริงมีแนวโน้มเพิ่มมากขึ้น เมื่อจำนวนข้อสอบทดสอบเลขที่สัมพันธ์กันเพิ่มมากขึ้น ซึ่งสอดคล้องกับ McDonald (1982) และ Hambleton & Swaminathan (1985) มีความเห็นตรงกันว่า ถ้าแบบทดสอบมีความเป็นเอกมิติ (unidimension) แล้ว จะมีความเป็นอิสระต่อกันในการตอบข้อสอบ (local independent) ด้วย ความเป็นอิสระในการตอบข้อสอบมีความสำคัญต่อการประมาณค่าพารามิเตอร์ต่าง ๆ มาก ถ้าขาดความเป็นอิสระในการตอบข้อสอบแล้ว ค่าพารามิเตอร์ต่าง ๆ ที่ประมาณค่าได้ จะมีความคลาดเคลื่อนจากความเป็นจริง นอกจากนี้ ข้อตกลงที่เกี่ยวข้องกันระหว่างความเป็นเอกมิติและความเป็นอิสระต่อกันในการตอบข้อสอบ เริ่มต้นจากแบบสอบที่วัดความสามารถเดียวกันใช้ผู้สอบที่ระดับ θ เดียวกัน ควรมีการตอบข้อสอบเป็นอิสระต่อกันทางสถิติ (statistically independence) แต่ถ้าที่ระดับ θ เดียวกัน พบว่าการตอบไม่เป็นอิสระต่อกันทางสถิติแล้ว ทำให้ผู้สอบบางคนมีคะแนนมากกว่าผู้สอบบางคน ที่ระดับ θ เดียวกัน แสดงว่าผู้สอบต้องใช้ความสามารถมากกว่าหนึ่งอย่างในการทำข้อสอบ ซึ่งสอดคล้องกับ Sireci, Thissen, & Wainer (1991) ที่กล่าวว่า ในการประมาณค่าความเที่ยงตามแนวคิดทฤษฎีการตอบสนองข้อสอบ การประมาณค่าจะมีความถูกต้องแม่นยำมากที่สุดเมื่อข้อสอบแต่ละข้อมีความเป็นอิสระต่อกัน แต่สำหรับข้อสอบที่สัมพันธ์กันนั้น มีแนวโน้มทำให้การประมาณค่าความเที่ยงมีค่าสูงขึ้น

ทั้งนี้ แม้ว่าแบบสอบที่มีข้อสอบทดสอบเลขที่ไม่เป็นอิสระต่อกันนั้น มีแนวโน้มทำให้การประมาณค่าความเที่ยงมีค่าสูงขึ้น แต่ผลการตรวจสอบดัชนีวัดความสอดคล้องความเป็นเอกมิติของแบบสอบที่เพิ่มจำนวนข้อสอบทดสอบเลขที่ไม่เป็นอิสระต่อกัน พบว่าแบบสอบที่เพิ่มจำนวนข้อสอบทดสอบเลขที่ไม่เป็นอิสระต่อกันทั้ง 20 ชุด ดัชนีวัดความสอดคล้องความเป็นเอกมิติของตัวแบบกับข้อมูลเชิงประจักษ์ทั้ง 8 ค่า ให้ผลสอดคล้องกับแบบสอบที่เพิ่มจำนวนข้อสอบทดสอบเลขที่ไม่เป็นอิสระต่อกันทั้ง 20 ชุด มีความเป็นเอกมิติ ซึ่งสอดคล้องกับผลการวิจัยของ Rajlic (2019) ที่ศึกษาการฝ่าฝืนข้อตกลงเรื่องความเป็นเอกมิติและความเป็นอิสระของข้อสอบ ผลการศึกษาแสดงให้เห็นว่า โมเดลทฤษฎีการตอบสนองข้อสอบแบบเอกมิติมีความแกร่งในการฝ่าฝืนข้อตกลงภายใต้ช่วงของระดับการฝ่าฝืนที่ต่างกัน แต่ให้สารสนเทศที่ได้จากเงื่อนไขของการตรวจสอบความแกร่งในการฝ่าฝืนข้อตกลงสูงกว่าความเป็นจริง

2. จากผลการเปรียบเทียบร้อยละของการเปลี่ยนแปลงฟังก์ชันสารสนเทศของแบบสอบ ซึ่งพบว่าที่อัตราการเพิ่มจำนวนข้อสอบบางส่วนที่ไม่เป็นอิสระต่อกันทั้ง 6 ระดับ ช่วงความสามารถ -3 ถึง -1 ค่าฟังก์ชันสารสนเทศของแบบสอบที่มีข้อสอบทดสอบเลขที่ไม่เป็นอิสระต่อกันต่ำกว่าค่าฟังก์ชันสารสนเทศของแบบสอบที่มีข้อสอบมีความเป็นอิสระต่อกัน ส่วนช่วงความสามารถ 0 ถึง 3 ค่าฟังก์ชันสารสนเทศของแบบสอบที่มีข้อสอบทดสอบเลขที่ไม่เป็นอิสระต่อกันสูงกว่าค่าฟังก์ชันสารสนเทศของแบบสอบที่มีข้อสอบมีความเป็นอิสระต่อกัน เนื่องจากฟังก์ชันสารสนเทศของแบบสอบเป็นผลรวมของเชิงพีชคณิตของค่าฟังก์ชันสารสนเทศของข้อสอบ

แต่ละข้อรวมเข้าด้วยกันทั้งฉบับ ณ ตำแหน่ง θ เดียวกัน ซึ่งค่าฟังก์ชันสารสนเทศของข้อสอบจะสูงหรือต่ำขึ้นอยู่กับความสามารถ θ ของผู้สอบว่าใกล้เคียงหรือแตกต่างจากค่าความยากของข้อสอบมากน้อยเพียงใด ทั้งนี้ เมื่อค่าอำนาจของข้อสอบเพิ่มขึ้นและค่าโอกาสในการเดามีค่าเข้าใกล้ศูนย์ จะทำให้ค่าฟังก์ชันสารสนเทศของข้อสอบมีค่าสูงขึ้น (Sireci, Thissen, & Wainer, 1991) ดังนั้น ในกลุ่มผู้สอบที่มีช่วงความสามารถ 0 ถึง 3 ซึ่งมีความสามารถใกล้เคียงกับค่าความยากของข้อสอบ ในการทำข้อสอบที่มีค่าอำนาจจำแนกสูง และโอกาสในการเดาข้อสอบต่ำ จึงมีแนวโน้มที่การประมาณค่าฟังก์ชันสารสนเทศของข้อสอบสูงกว่าความเป็นจริง ในทำนองเดียวกับที่ในกลุ่มผู้สอบที่มีช่วงความสามารถ -3 ถึง -1 ซึ่งมีความสามารถไกลจากค่าความยากของข้อสอบ ในการทำข้อสอบที่มีค่าอำนาจจำแนกสูง และโอกาสในการเดาข้อสอบต่ำ จึงมีแนวโน้มที่การประมาณค่าฟังก์ชันสารสนเทศของข้อสอบต่ำกว่าความเป็นจริง เมื่อค่าฟังก์ชันสารสนเทศของข้อสอบมีค่าสูงหรือต่ำกว่าความเป็นจริง จึงส่งผลต่อฟังก์ชันสารสนเทศของแบบสอบด้วยเช่นกัน

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 ข้อเสนอแนะในการนำไปประยุกต์ใช้ในสถานการณ์จริง

1.1.1 การวิจัยในครั้งนี้ แม้ว่าภายใต้บริบทของทฤษฎีการตอบสนองนั้นข้อสอบจะยังมีความแกร่งในการฝ่าฟันข้อตกลงเรื่องความเป็นอิสระต่อกันของข้อสอบ แต่ผลการเพิ่มจำนวนข้อสอบทดสอบทดสอบที่ ไม่เป็นอิสระในข้อสอบที่เป็นอิสระ พบว่า เมื่ออัตราส่วนจำนวนข้อที่สัมพันธ์กันต่อจำนวนข้อที่เป็นอิสระต่อกันเพิ่มมากขึ้นค่าดัชนีวัดความสอดคล้องความเป็นเอกมิติลดลง แต่ในทางตรงกันข้ามกลับให้ผลการประมาณค่าฟังก์ชันสารสนเทศของข้อสอบที่สุทธเกินความเป็นจริง ดังนั้น จากผลการวิจัย จึงควรเพิ่มจำนวนข้อสอบทดสอบทดสอบที่ ไม่เป็นอิสระต่อจำนวนข้อสอบที่เป็นอิสระ ด้วยอัตราส่วนไม่เกิน 15:85 เพื่อให้การประมาณค่าพารามิเตอร์ในสถานการณ์มีความถูกต้องแม่นยำมากที่สุด

1.1.2 แม้ว่ากรณีที่ข้อสอบมีจำนวนมาก จะทำให้การประมาณค่าฟังก์ชันสารสนเทศมีค่าเพิ่มขึ้น แต่การตรวจสอบความเป็นอิสระระหว่างข้อสอบแต่ละข้อมีจำนวนคู่ในการตรวจสอบเพิ่มขึ้น เวลาที่ใช้ในการประมาณค่าจะนานขึ้นด้วย ดังนั้นการนำไปใช้ในสถานการณ์จริงต้องคำนึงถึงขีดความสามารถของคอมพิวเตอร์ที่ใช้ด้วย และจากผลการวิจัยพบว่า ในสถานการณ์จริง จำนวนข้อสอบควรไม่น้อยกว่า 40 ข้อ ขึ้นไป

1.1.3 การนำผลการศึกษาไปใช้ในสถานการณ์การสอบระดับชาตินั้น ประเด็นที่ควรพิจารณาอย่างยิ่งคือ

1) ความเป็นเอกมิติของข้อสอบ ซึ่งในการศึกษาครั้งนี้เป็นการศึกษาภายใต้ความเป็นเอกมิติ แต่ในสถานการณ์ของการสอบระดับชาตินั้น แบบสอบแต่ละฉบับอาจวัดหลายมิติ ดังนั้น การนำผลการศึกษาไปประยุกต์ใช้จึงต้องพิจารณาแบ่งแบบสอบออกเป็นมิติย่อย เพื่อประมาณค่าพารามิเตอร์ของข้อสอบ

2) ขนาดกลุ่มตัวอย่าง การศึกษาครั้งนี้สามารถนำไปใช้ได้กับสถานการณ์ของการสอบที่มีผู้สอบ 1,000 คน ขึ้นไป รวมถึงสถานการณ์ของการสอบระดับชาติที่อาจมีผู้สอบมากกว่าหนึ่งแสนคนขึ้นไป อย่างไรก็ตาม แม้ในสถานการณ์ของการสอบระดับชาตินั้น กลุ่มผู้สอบจะเปลี่ยนกลุ่มไม่ซ้ำกัน และการสร้าง

แบบสอบเพื่อใช้ในการสอบจะไม่มี การนำแบบสอบที่ใช้สอบระดับชาติไปแล้วมาใช้ใหม่ แต่ยังคงมีความจำเป็น อย่างยิ่งที่ควรพิจารณาถึงการประมาณค่าพารามิเตอร์ของข้อสอบที่มีข้อสอบที่สัมพันธ์กัน เพื่อให้ได้ ค่าพารามิเตอร์ที่ใกล้เคียงกับค่าพารามิเตอร์ที่แท้จริงมากที่สุด เพื่อให้สามารถนำไปใช้ในการวางแผนการ คัดเลือกข้อสอบเพื่อจัดทำคลังข้อสอบ หรือวางแผนการพัฒนาการศึกษาของประเทศมีความถูกต้องต่อไป

1.2 ข้อเสนอแนะในเชิงเทคนิคของการจำลองข้อมูล

1.2.1 ในขั้นตอนการตรวจสอบแบบแผนการตอบข้อสอบแบบสองค่า (0, 1) ควรมีการ พิจารณาว่า ข้อมูลเป็นเอกมิติตามข้อตกลงของทฤษฎีการตอบสนองข้อสอบหรือไม่ แม้ว่าโมเดลทฤษฎีการ ตอบสนองข้อสอบมีความแข็งแกร่งปานกลาง (moderately robust) ในการฝ่าฝืนจากความเป็นเอกมิติ (Cooke & Michie, 1997) แต่เพื่อความถูกต้องแม่นยำของการศึกษา ถ้าไม่เป็นไปตามเงื่อนไขดังกล่าวควรตัดข้อมูลทิ้ง และทำการสร้างข้อมูลใหม่

1.2.2 การวิจัยครั้งนี้ผู้วิจัยสร้างข้อมูล และตรวจสอบความถูกต้องของข้อมูลที่จำลองด้วย โปรแกรม R ทุกขั้นตอนของการศึกษา ตั้งแต่ขั้นตอนของการสร้างข้อมูล (generate data) ไปจนถึงขั้นตอน การประมาณค่าพารามิเตอร์ และการตรวจสอบความอิสระและความเป็นเอกมิติ เนื่องจากสามารถวิเคราะห์ได้ ในคราวเดียวกัน อย่างไรก็ตามการใช้โปรแกรมดังกล่าวจำเป็นที่จะต้องอาศัยผู้ที่มีความรู้ความเชี่ยวชาญเฉพาะ ด้านในการพัฒนาโปรแกรม แม้จะมี package ที่สามารถประยุกต์ใช้ได้ แต่จะต้องระมัดระวังในขั้นตอนของการ เขียนคำสั่ง และคำนวณค่าระหว่างเวกเตอร์หรือเมตริกซ์ เพราะจะทำให้ผลที่ได้เกิดข้อผิดพลาดได้ง่าย

1.2.3 ในการใช้โปรแกรม R ในการศึกษา สิ่งที่ควรคำนึงถึงคือ แม้จะมี package สำเร็จรูปให้ เลือกใช้ในการวิเคราะห์ข้อมูล ซึ่งหลาย package ใช้สำหรับวิเคราะห์ข้อมูลด้วยเทคนิคเดียวแต่มีเงื่อนไขในการ ใช้และข้อจำกัดแตกต่างกัน ดังนั้น ผู้จะใช้โปรแกรม R ในการวิเคราะห์ข้อมูล จึงควรคำนึงถึงทฤษฎีที่จะนำไปใช้ ประกอบด้วย

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 เนื่องจากการวิจัยครั้งนี้เป็นการศึกษาการเปลี่ยนแปลงดัชนีวัดความสอดคล้องความเป็นเอก มิติของแบบสอบที่มีการเพิ่มจำนวนข้อสอบทดสอบที่ทดสอบที่ไม่เป็นอิสระต่อกันตามลำดับชุด เพื่อให้การศึกษา ครอบคลุมมากยิ่งขึ้น ควรศึกษาการเปลี่ยนแปลงดัชนีวัดความสอดคล้องความเป็นเอกมิติเมื่อตำแหน่งของ ข้อสอบทดสอบแตกต่างกัน

2.2 ควรศึกษาควรตรงเชิงเนื้อหาของแบบสอบเมื่อแบบสอบมีการเพิ่มจำนวนข้อสอบทดสอบที่ไม่ เป็นอิสระต่อกัน

2.3 การศึกษาครั้งนี้ เป็นการศึกษาภายใต้ความเป็นเอกมิติโดยจำลองข้อมูลแบบเลือกตอบ 4 ตัวเลือก ซึ่งเป็นข้อตกลงเบื้องต้นที่สำคัญของทฤษฎีการตอบสนองข้อสอบที่มีการตรวจให้คะแนนแบบ 2 ค่า ที่มีโอกาสในการตอบถูกสูงสุดเท่ากับ 0.25 ควรมีการศึกษาเปรียบเทียบผลกับการศึกษาภายใต้แบบสอบ ประเภทอื่นที่มีการตรวจให้คะแนน 2 ค่า แต่โอกาสในการตอบถูกต่างกัน

2.4 ควรมีการศึกษาเปรียบเทียบผลการศึกษายภายใต้โมเดลทฤษฎีการตอบสนองข้อสอบแบบ เทสต์เลท กับผลการศึกษายภายใต้โมเดลทฤษฎีการตอบสนองข้อสอบ

2.5 เนื่องจากงานวิจัยนี้ เป็นการศึกษาศึกษาจากสถานการณ์การจำลองข้อมูล ซึ่งเป็นการศึกษาในเพื่อให้ได้ข้อสรุปในเชิงทฤษฎี ในการศึกษาครั้งต่อไป ควรนำผลการศึกษาเชิงทฤษฎีไปประยุกต์ใช้กับข้อมูลจริงให้สามารถเลือกใช้ได้อย่างสอดคล้องกับสภาพจริงมากยิ่งขึ้น

เอกสารอ้างอิง

- จงกล บัวแก้ว, ไพรัตน์ วงษ์นาม และ สมพงษ์ ปันพูน. (2561). การประมาณค่าพารามิเตอร์ของแบบสอบเทสต์เลทตามโมเดลทฤษฎีการตอบสนองข้อสอบและโมเดลทฤษฎีการตอบสนองเทสต์เลท. *วารสารราชภัฏสุราษฎร์ธานี*, 5(1), 139-155.
- วรนุช แหยมแสง. (2537). *การพัฒนากระบวนการตรวจสอบความเป็นเอกมิติของแบบสอบ*. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต ภาควิชาวิจัยการศึกษา จุฬาลงกรณ์มหาวิทยาลัย.
- Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey, Calif.: Brooks/Cole Pub. Co.
- Baghaei, P. & Ravand, H. (2016). Modeling local item dependence in cloze and reading comprehension test items using testlet response theory. *Psicológica*, 37, 85-104.
- Brandt, S. (2012). Robustness of multidimensional analyses against local item dependence. *Psychological Test and Assessment Modeling*, 54(1), 36-53.
- Cao, Y., Lu, R., & Tao, W. (2014). Effect of item response theory (irt) model selection on testlet-based test equating. *ETS Research Report Series*, 2014(2), 1-13. doi.org/10.1002/ets2.12017
- Eckes, T. & Baghaeib, P. (2015). Using testlet response theory to examine local dependence in c-tests. *Applied Measurement in Education*, 28, 85–98. DOI: 10.1080/08957347.2014.1002919
- Edwards, M. C. & Cai, Li. (2011). *A new procedure for detecting departures from local independence in item response models*. CA: Department of Psychology, The Ohio State University.
- Hambleton, R. K. & Swaminathan, H. (1985). *Item response theory: Principles and application*. Boston: Kluwer-Nijhoff.
- Hambleton, R. K., & Traub, R. E. (1974). The effects of item order on test performance and stress. *Journal of Experimental Education*, 43(1), 40–46. https://doi.org/10.1080/00220973.1974.10806302
- Jiao, H., Wang, S., & He, W. (2013). Estimation methods for one-parameter testlet models. *Journal of Educational Measurement*, 50, 186–203. DOI: 10.1111/jedm.12010

- Li, F. (2017). An information-correction method for testlet-based test analysis: From the perspectives of item response theory and generalizability theory. *ETS Research Report Series*, 2017, 1-25. doi.org/10.1002/ets2.12151
- Liu, Y. & Maydeu-Olivares, A. (2012). Local dependence diagnostics in irt modeling of binary data. *Educational and Psychological Measurement*, 73(2), 254–274. DOI: 10.1177/0013164412453841
- Lord, F. M. (1980). *Application of item response theory to practical testing problems*. Hillsdak NJ: Erlbaum.
- McDonald, R. P. (1983). Exploratory and confirmatory nonlinear confirmatory nonlinear common factor analysis. In H. Wainer and S. Messick (Eds), *Principals of modern psychological measurement: A Festschrift for Frederic M.Lord*. Hillsdale NJ: Erlbaum. (p.197-213).
- Nandakumar, R. & Stout, W. (1993) Refinements of stout's procedure for assessing latent trait unidimensionalit. *Journal of Educational Statistics*, 18, 41-68.
- Rajlic, G. (2019). *Violations of unidimensionality and local independence in measures intended as unidimensional: assessing levels of violations and the accuracy in unidimensional irt model estimates* (Doctoral dissertation). the University of British Columbia, Vancouver, Canada. DOI: 10.14288/1.0380235
- Ravand, H. (2015). Assessing testlet effect, impact, differential testlet, and item functioning using cross-classified multilevel measurement modeling. *SAGE Open*, (April-June 2015), 1–9. DOI: 10.1177/2158244015585607
- Roznowski, M., Tucker, L. R., & Humphreys, L. G. (1991). Three approaches to determining the dimensioinality of binary items. *Applied Psychological Measurement*, 15, 109-127.
- Rubright, J. D. (2018). Impact of both local item dependencies and cut-point locations on examinee classifications. *Educational Measurement: Issues and Practice*, 37(3), 40-45. DOI: 10.1111/emip.12183
- Setiawati, F. A., Izzaty, R. E., & Hidayat, V. (2018). Items parameters of the space-relations subtest using item response theory. *Data in Brief*, 19(2018), 1785–1793. doi.org/10.1016/j.dib. 2018.06.061
- Sireci, S. G., Thissen, D., & Wainer, H. (1991). On the reliability of testlet-based tests. *Journal of Educational Measurement*, 28(3), 237-247. https://www.jstor.org/stable/1434845
- Tao, W. & Cao, Y. (2016). An extension of IRT-based equating to the dichotomous testlet response theory model. *Applied Measurement in Education*, 29(2), 108–121.

- Van den Wollenberg, A. L. (1982). Two new test statistics for the Rasch model. *Psychometrika*, 47(2), 123–140. <https://doi.org/10.1007/BF02296270>
- Wainer, H., Bradlow, E. T., & Wang, X. (2007). *Testlet response theory and its applications*. New York, NY: Cambridge University Press. doi.org/10.1017/ CBO9780511618765
- Warm, T. A. (1978). *A primer of item response theory*. Oklahoma: Coast Guard Institute.
- Yen, W. M. (1984). Effect of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, 8, 125-145.
- Zhan, P., Liao, M., & Bian, Y. (2018). Joint testlet cognitive diagnosis modeling for paired local item dependence in response times and response accuracy. *the Journal Frontiers in Psychology*, 9(April 2018), 1-14. DOI: 10.3389/fpsyg. 2018.00607