

### สรุปและอภิปรายผลการวิจัย

การเรียนการสอนในห้องเรียน จำเป็นต้องใช้แบบสอบที่มีคุณภาพในการประเมินผลการเรียนการสอน สูตรต่าง ๆ ที่ใช้ในการตรวจสอบคุณภาพของข้อสอบต้องสร้างขึ้นมาจากข้อตกลงเบื้องต้น หนึ่งในข้อตกลงเบื้องต้นนั้นก็คือความเป็นเอกมิตินี้ของแบบสอบ แบบสอบที่นำมาวิเคราะห์ข้อสอบจะต้องมีลักษณะของความเป็นเอกมิติ เพื่อที่ผลการวิเคราะห์ที่จะได้ออกมาอย่างมีความหมายถูกต้องแม่นยำ ไม่ฝ่าฝืนข้อตกลงเบื้องต้น ดังนั้น การตรวจสอบความเป็นเอกมิติของแบบสอบจึงเป็นส่วนสำคัญก่อนการวิเคราะห์ข้อสอบ

เป้าหมายของการวิจัยครั้งนี้ เพื่อพัฒนากระบวนการตรวจสอบของดัชนีบ่งชี้ความเป็นเอกมิติของแบบสอบ โดยพัฒนาจากดัชนีบ่งชี้เดิม 3 ตัว คือ 1) พิจารณาจากการพล็อตค่าไอเกิน (EP) 2) ใช้วิธีการทดสอบด้วยไบซีเรียล (BT) และ 3) ใช้ดัชนีความเป็นเอกพันธ์ของกรีน (G) มาเป็นดัชนีบ่งชี้ใหม่ 3 ตัว คือ 1) ใช้ค่าอัตราส่วนระหว่างค่าไอเกินตัวที่ 1 กับตัวที่ 2 (ER) 2) ใช้วิธีการทดสอบด้วยไบซีเรียลประยุกต์ (ABT) และ 3) ใช้ดัชนีความเป็นเอกพันธ์ของกรีนประยุกต์ (AG)

นอกจากนี้ ยังได้ทำการตรวจสอบความไวในการบ่งชี้และศึกษาความไม่ผันแปรของดัชนีบ่งชี้เหล่านี้

การวิจัยในครั้งนี้ ได้ใช้ข้อมูลจริงจากนักเรียนชั้นประถมศึกษาปีที่ 5 ของโรงเรียนในเขตบางกะปิ 13 โรงเรียน จำนวนนักเรียน 1,000 คน การเก็บรวบรวมข้อมูลในการวิจัยนี้แบ่งออกเป็น 3 ตอน คือ

1. เก็บข้อมูลภูมิหลังของนักเรียน ซึ่งเป็นคะแนนสมรรถภาพทางด้านคณิตศาสตร์ และสมรรถภาพทางด้านภาษาอังกฤษของนักเรียน นอกจากนี้ยังมีเพศของนักเรียน เพื่อนำข้อมูลไปใช้ในการวิเคราะห์โมเดล MIMIC ต่อไป

2. เก็บข้อมูลจากแบบสอบ 2 ฉบับ คือ คณิตศาสตร์ และ ภาษาอังกฤษ ซึ่งผู้วิจัยได้สร้างขึ้นมาจากเนื้อหาแคบ ๆ และได้ผ่านการประเมินจากผู้เชี่ยวชาญ 2 กลุ่ม คือ กลุ่มอาจารย์ผู้สอบในระดับชั้นประถมศึกษาปีที่ 5 และกลุ่มนักวัดผล

ข้อมูลที่รวบรวมมาได้จากแบบสอบทั้ง 2 ฉบับได้นำมาทำให้มีระดับความเป็นเอกมิติต่าง ๆ โดยทำการคัดเลือกข้อสอบฉบับหนึ่งออกไปทีละ 2 ข้อและทำการสุ่มข้อสอบจากอีกฉบับหนึ่งเข้ามาแทนทีละ 2 ข้อเช่นกัน ทำให้แบบสอบหลักมีความเป็นเอกมิติลดลงอย่างมีระบบ ซึ่งได้รับการยืนยันโดยวิธีวิเคราะห์โมเดล MIMIC

3. ทำการเก็บรวบรวมข้อมูลจากแบบสอบมาตรฐานวิชาคณิตศาสตร์ จำนวน 10 ข้อ มีค่าความยากอยู่ระหว่าง .30 ถึง .80 มีค่าอำนาจจำแนกตั้งแต่ .39 ถึง .59 และแบบสอบมาตรฐานวิชาภาษาอังกฤษ จำนวน 20 ข้อ มีค่าความยากง่ายอยู่ระหว่าง .41 ถึง .86 และมีค่าอำนาจจำแนกตั้งแต่ .27 ถึง .77 ข้อมูลจากแบบสอบมาตรฐานนี้จะนำไปใช้ในการคำนวณดัชนีบ่งชี้ ABT

### สรุปผลการวิจัย

การวิจัยครั้งนี้ได้ข้อสรุป ดังนี้

1. พัฒนาการระบวนการตรวจสอบความเป็นเอกมิตี 3 วิธี

- 1.1 พัฒนาการระบวนการตรวจสอบด้วย EP มาเป็นการตรวจสอบด้วย ER
- 1.2 พัฒนาการระบวนการตรวจสอบด้วย BT มาเป็นการตรวจสอบด้วย ABT
- 1.3 พัฒนาการระบวนการตรวจสอบด้วย  $\mu$  มาเป็นการตรวจสอบด้วย AG

2. การตรวจสอบคุณภาพของดัชนีบ่งชี้ที่พัฒนาขึ้นใหม่ เปรียบเทียบกับดัชนีบ่งชี้เดิม 3 วิธี

2.1 เปรียบเทียบระหว่าง EP กับ ER พบว่าวิธีการพล็อตค่าไอเกิน ใช้ระบุความเป็นเอกมิตีได้เฉพาะเมื่อใช้ภาษาอังกฤษเป็นหลัก ส่วนเมื่อใช้คณิตศาสตร์เป็นหลัก ค่าไอเกินเป็นแปลงไปอย่างไม่เป็นระบบ ในขณะที่แบบสอบมีความเป็นเอกมิติลดลง ดัชนี ER มีค่าลดลงอย่างเป็นระบบ ในขณะที่แบบสอบมีความเป็นเอกมิติลดลง จำนวนข้อมีผลต่อค่าไอเกิน โดยในกรณีข้อสอบ 15 ข้อ ให้ค่าไอเกินสูงกว่าในกรณีข้อสอบ 10 ข้อ ส่วนในด้านเนื้อหา พบว่ากราฟ EP ในแบบสอบคณิตศาสตร์เป็นหลัก มีค่าไอเกินต่ำกว่ากราฟ EP ในแบบสอบที่ใช้ภาษาอังกฤษเป็นหลักทั้งในกรณีข้อสอบ 10 ข้อและ 15 ข้อ ส่วนค่า ER สรุปไม่ได้ว่า จำนวนข้อและเนื้อหา มีผลต่อ ER หรือไม่

2.2 เปรียบเทียบระหว่างดัชนี BT และ ABT พบว่าค่า BT และ ABT เปลี่ยน

แปลงค่าอย่างไม่มีทิศทางแน่นอน ในทุกเงื่อนไข ในขณะที่แบบสอบถามมีความเป็นเอกมิติลดลงอย่าง เป็นระบบ ดังนั้นค่า BT และ ABT ในการวิจัยนี้จึงไม่อาจนำมาใช้ในการระบุความเป็นเอกมิติ ของแบบสอบถามในทุกเงื่อนไข ค่า ABT ไม่ได้ให้ค่าที่ได้เปรียบกว่าค่า BT เนื่องจากค่า BT และ ABT มีการเปลี่ยนแปลงค่าที่ไม่เป็นระบบ ทั้งในกรณีข้อสอบ 10 ข้อ และข้อสอบ 15 ข้อ จึง ไม่อาจสรุปได้ว่า จำนวนข้อมีผลต่อการคำนวณค่า BT และ ABT หรือไม่ ส่วนในด้านเนื้อหาที่ ตอบไม่ได้เช่นกันว่า เนื้อหา มีผลต่อการคำนวณค่า BT และ ABT หรือไม่

2.3 เปรียบเทียบระหว่างดัชนี  $\mu$  และ AG พบว่า ค่า  $\mu$  และค่า AG มีค่าเกิน 1.000 ในบางเงื่อนไข และมีการเปลี่ยนแปลงค่าที่ไม่มีทิศทางแน่นอน ในขณะที่แบบสอบถาม มีความเป็นเอกมิติลดลงอย่างเป็นระบบ ดังนั้น ค่า  $\mu$  และ AG ในการวิจัยนี้จึงไม่อาจนำมาใช้ในการ ระบุความเป็นเอกมิติของแบบสอบถามได้ และค่า AG ก็ให้ผลการวิเคราะห์ที่ไม่ดีไปกว่าค่า เนื่องจากค่า  $\mu$  และ AG เปลี่ยนแปลงค่าไม่เป็นระบบ จึงบอกไม่ได้ว่าจำนวนข้อและเนื้อหา มี ผลต่อการคำนวณค่า  $\mu$  และ AG หรือไม่

3. การเปรียบเทียบความไวในการเปลี่ยนแปลงค่าของดัชนีบ่งชี้ที่พัฒนาขึ้นมาใหม่กับ ดัชนีบ่งชี้เดิม

3.1 ดัชนี ER ให้ความไวในการเปลี่ยนแปลงค่าที่ไม่เกิน 100 % และมีความไวใน การเปลี่ยนแปลงค่าอย่างเป็นระบบเข้าสู่ 100 ในขณะที่แบบสอบถามมีความเป็นเอกมิติลดลง

3.2 ดัชนี BT และ ABT ไม่สามารถแสดงความไวในการเปลี่ยนแปลงค่าอย่างเป็น ระบบได้ ดังนั้นจึงบอกไม่ได้ว่าดัชนี ABT จะมีความไวในการบ่งชี้ความเป็นเอกมิตินอกจาก BT หรือไม่

3.3 ดัชนี  $\mu$  และ AG ไม่สามารถแสดงความไวในการเปลี่ยนแปลงค่าอย่างเป็น ระบบ ในขณะที่แบบสอบถามมีความเป็นเอกมิติลดลง ดังนั้นจึงบอกไม่ได้ว่าดัชนี AG จะมีความไวใน การบ่งชี้ความเป็นเอกมิตินอกจาก  $\mu$  หรือไม่

3.4 ดัชนี ER เป็นดัชนีบ่งชี้ความเป็นเอกมิติได้ดีที่สุด

### อภิปรายผลการวิจัย

1. โมเดล MIMIC ใช้พิจารณาความสอดคล้องระหว่างข้อมูลกับโมเดล ว่ามีคุณลักษณะ

แผงเพียง 1 คุณสมบัติ ดังนั้นจึงนำมาใช้ในการเปรียบเทียบความเป็นเอกมิติของแบบสอบ กับ วิธีตรวจสอบแบบอื่น ๆ การวิเคราะห์โมเดลเป็นการตรวจสอบความสอดคล้องระหว่างข้อมูลกับ โมเดลตามทฤษฎี ซึ่งได้อาศัยค่า  $\chi^2$  และค่าความสอดคล้องกับข้อมูล (GFI) ถ้าข้อมูลมีความ สอดคล้องกับโมเดลมาก ค่า  $\chi^2$  จะมีค่าต่ำ และค่า GFI จะมีค่าสูง นั่นคือในขณะที่ข้อมูล เชิงประจักษ์ได้ให้แบบสอบที่มีความเป็นเอกมิติสูง และลดต่ำลงอย่างเป็นระบบ ค่าความสอดคล้องระหว่างข้อมูลกับโมเดลก็จะให้ค่าตัวเลขที่เปลี่ยนแปลงอย่างเป็นระบบคล้ายตามกัน เช่นเดียวกับดัชนี ER ในขณะที่ดัชนีบ่งชี้ BT, ABT, และ AG ไม่สามารถให้ค่าที่มีความต่อเนื่อง ได้ เมื่อแบบสอบมีความเป็นเอกมิติเปลี่ยนแปลงไป

2. การวิจัยในครั้งนี้พบว่าดัชนี BT, ABT,  $\mu$  และ AG ไม่อาจนำมาใช้ในการ ตรวจสอบความเป็นเอกมิติของแบบสอบทุกกรณี ทั้งนี้อาจเป็นเพราะสาเหตุดังต่อไปนี้

2.1 ดัชนี BT ซึ่งเกิดจากค่าความสัมพันธ์ระหว่างค่าอำนาจจำแนกรายชื่อ (RYT) กับค่าน้ำหนักของข้อบ่งชี้ตัวประกอบร่วมตัวแรก (COMM) มีค่าต่ำและมีการเปลี่ยนแปลงค่าที่ไม่เป็น ระบบ ในขณะที่แบบสอบมีความเป็นเอกมิติลดต่ำลงอย่างเป็นระบบ เมื่อพิจารณาตาราง (ง.1) ในภาคผนวก (ง) ซึ่งแสดงความสัมพันธ์ระหว่างค่าอำนาจจำแนกรายชื่อ (RYT) ค่าความตรง รายชื่อ (MSI) และค่าน้ำหนักของข้อบ่งชี้ตัวประกอบร่วมตัวแรก (COMM) จะเห็นว่าค่า RYT แต่ ละข้อมีค่าตั้งแต่ .379 ถึง .596 ซึ่งมีความแปรปรวนน้อย และค่า COMM มีค่าตั้งแต่ .491 ถึง .850 ซึ่งมีความแปรปรวนน้อยเช่นกัน ผลของความสัมพันธ์จึงมีค่าน้อย (Ferguson, 1985 : 111) และการเปลี่ยนแปลงของค่า RYT และ COMM ไม่สอดคล้องกับการเปลี่ยนแปลงของเอกมิติ จึงทำให้ค่า BT เปลี่ยนแปลงค่าอย่างไม่เป็นระบบ ในขณะที่แบบสอบมีเอกมิติลดลงอย่างเป็นระบบ

ดัชนี ABT ซึ่งเกิดจากค่าความสัมพันธ์ระหว่างค่าความตรงรายชื่อ (MSI) กับ ค่าน้ำหนักของข้อบ่งชี้ตัวประกอบร่วมตัวแรก (COMM) มีค่าต่ำ และมีการเปลี่ยนแปลงค่าที่ไม่เป็น ระบบเช่นเดียวกัน สาเหตุอาจเป็นเพราะ (1) แบบสอบไม่มีความสัมพันธ์กับแบบสอบมาตรฐาน (2) ค่า MSI มีค่าต่ำและมีความแปรปรวนน้อย (.06 ถึง .13) และค่า COMM มีค่าน้อยเช่น เดียวกัน จึงทำให้ความสัมพันธ์ระหว่าง MSI และ COMM มีค่าน้อย และเช่นเดียวกันการเปลี่ยนแปลงของค่า MSI และของ COMM ไม่สอดคล้องกับการเปลี่ยนแปลงระดับของเอกมิติ

จึงสรุปได้ว่า ดัชนี BT และ ABT ไม่เหมาะที่จะใช้ตรวจสอบความเป็นเอกมิติ ในการวิจัยนี้

2.2 ที่มาของสูตรความเป็นเอกพันธ์ของกรีน เกิดจากนิยามที่ว่า ถ้าแบบสอบถามมีเพียงมิติเดียวจะให้ตัวประกอบร่วม (Common factor) เพียงตัวเดียว และเมื่อหาค่าเมตริกซ์สหสัมพันธ์หวนกลับ (r-reproduce) คือ  $AA'$  ก็ยอมให้เมตริกซ์ที่มีค่าแนวทแยงมุมหลักเป็นค่า  $h_1^2$  และเมตริกซ์นี้ ก็จะทำให้ค่าแรงค์เป็น 1.000 ด้วย เมื่อเป็นเช่นนั้น ไมเนอร์เมตริกซ์ของเมตริกซ์  $AA'$  ทุกไมเนอร์ จะมีค่าเป็นศูนย์ นั่นคือ

$$\begin{pmatrix} h_1^2 & r_{1,j} \\ r_{1,j} & h_1^2 \end{pmatrix} = 0$$

ดังนั้น  $|r_{1,j}| = \sqrt{h_1^2 h_1^2}$

$$\text{ซึ่ง } |r_{1,j}| / \sqrt{h_1^2 h_1^2} = 1$$

สูตร ดัชนีความเป็นเอกพันธ์ของกรีน จึงเป็น

$$\mu = \frac{\sum_{i \neq j} |r_{i,j}|}{\sum_{i \neq j} \sqrt{h_i^2 h_j^2}}$$

โดย มีค่าอยู่ระหว่าง 0 ถึง 1 และ  $\mu$  จะเข้าใกล้ 1 เมื่อแบบสอบถามมีความเป็นเอกพันธ์มากขึ้น แต่ในการวิจัยนี้ให้ค่า  $\mu$  ที่เกิน 1.000 ในบางเงื่อนไข ทั้งนี้อาจเป็นเพราะ

การประมาณค่า  $h_1^2$  จากค่า  $SMC_1$  (Square Multiple Correlation) ซึ่งให้ค่า  $h_1^2$  ที่ต่ำกว่าความเป็นจริง จึงทำให้ค่า  $r_{1,j} > \sqrt{h_1^2 h_1^2}$  (Green, 1977:836)

ผลการวิจัยเช่นนี้สอดคล้องกับงานวิจัยของ Warm (Warm, 1978 : 107) ซึ่งเขาได้พัฒนาสูตรของ Green ใหม่ โดยประมาณค่า  $h_1^2$  จาก  $SMC_1$  และได้ใช้ค่าปรับแก้ ดังนี้

$$G = \frac{\sum \sum \frac{|r_{i,j}|}{\sqrt{h_i^2 h_j^2}}}{n}$$

Warm ได้ศึกษาวิธีการตรวจสอบความเป็นเอกมิติวิธีการต่างๆ รวมทั้งดัชนี G และ ได้ข้อสรุปว่าค่าของ G ไม่ได้เข้าสู่ 1.000 เมื่อข้อมูลเป็นเอกมิติ และไม่ได้เข้าสู่ 0.000 เมื่อข้อมูลไม่เป็นเอกมิติ ยิ่งไปกว่านั้นยังรายงานว่ ดัชนี G เป็นดัชนีบ่งชี้ไม่ได้ไปกว่าการใช้ค่าเปอร์เซ็นต์ของความแปรปรวนร่วมของตัวประกอบตัวแรก (first factor percent of common

variance)

ส่วนดัชนี AG ยังคงให้ค่าที่เกิน 1.000 ทั้งนี้เป็นเพราะว่าการประมาณค่า  $h_1^2$  ยังคงต่ำกว่าความเป็นจริง เพราะได้ใช้ค่า  $h_1^2$  จากการสกัดตัวประกอบแบบอิมเมจซึ่งเป็นการวิเคราะห์ตัวประกอบแบบตัวประกอบร่วม (Common Factor Analysis) การประมาณค่า  $h_1^2$  ยังคงให้ค่าที่ต่ำกว่าความเป็นจริง (Mulaik, 1972 : 148)

2.3 ค่าความยากของข้อสอบมีผลต่อการคำนวณค่าดัชนี BT, ABT,  $\mu$  และ AG กล่าวคือ ดัชนีทั้ง 4 ตัว ต้องใช้การวิเคราะห์ตัวประกอบมาช่วยในการตรวจสอบ และค่าความยากก็มีผลต่อการวิเคราะห์ตัวประกอบของเมตริกซ์สหสัมพันธ์ระหว่างข้อ (Hambleton & Swaminathan, 1985:21) จึงอาจมีผลต่อดัชนีทั้ง 4 ตัว

3. การที่กราฟ EP ไม่สามารถใช้ตรวจสอบความเป็นเอกมิติของแบบสอบ เมื่อใช้คณิตศาสตร์เป็นหลักเพราะกราฟ EP มีค่า ไอเกินที่เปลี่ยนแปลงไปอย่างไม่เป็นระบบทั้งกรณีข้อสอบ 10 ข้อ และ 15 ข้อ เมื่อใช้แบบสอบคณิตศาสตร์เป็นหลักที่เป็นเช่นนี้ อาจเป็นเพราะแบบสอบคณิตศาสตร์มีเนื้อหาปานกลางถึงค่อนข้างยาก (P .190 ถึง .686) เมื่อมีการสุ่มข้อสอบภาษาอังกฤษซึ่งมีความง่ายกว่า (P .471 ถึง .907) เข้ามาแทนที่ข้อสอบคณิตศาสตร์ที่ถูกคัดออกไป ทำให้แบบสอบคณิตศาสตร์มีการเกาะกลุ่มของข้อต่าง ๆ ขึ้นมาใหม่ ซึ่งมีอิทธิพลต่อความสัมพันธ์ระหว่างข้อเพิ่มขึ้นแทนที่จะมีความสัมพันธ์ลดลง จึงทำให้ค่าไอเกินมีค่าเพิ่มขึ้นในขณะที่แบบสอบมีความเป็นเอกมิตลดลง

4. การที่ดัชนี ER สามารถนำมาใช้ในการตรวจสอบเอกมิติของแบบสอบได้เพราะดัชนี ER มีการเปลี่ยนแปลงค่าที่สอดคล้องกับดัชนี GFI และนอกจากนี้ดัชนี ER ยังมีความไวในการเปลี่ยนแปลงค่ามากกว่าดัชนี GFI และเป็นดัชนีที่มีความไวอย่างเป็นระบบ เมื่อเทียบกับดัชนีตัวอื่น ๆ ซึ่งมีค่าเปลี่ยนแปลงอย่างไม่เป็นระบบ คือ มีความไวขึ้น ๆ ลง ๆ ดัชนี ER ได้เปรียบกว่ากราฟ EP เพราะดัชนี ER สามารถให้ค่าความมากน้อยในการบ่งชี้ได้

#### ข้อเสนอแนะ

1. ในทางปฏิบัติโดยทั่วไป การตรวจสอบความเป็นเอกมิติของแบบสอบควรจะใช้ดัชนี บ่งชี้ ER เพราะเป็นดัชนีที่มีความสอดคล้องกับดัชนี GFI ในโมเดล MIMIC

2. ควรได้มีการวิจัยแบบเดียวกันนี้ แต่มีการคัดเลือกกลุ่มตัวอย่างที่มีความสามารถแตกต่างกันมาก โดยคัดเลือกนักเรียนที่เก่งคณิตศาสตร์อย่างเฉียวและอ่อนในวิชาภาษาอังกฤษ กับนักเรียนที่เก่งภาษาอังกฤษอย่างเฉียวและอ่อนวิชาคณิตศาสตร์ เพราะการคัดเลือกตัวอย่างเช่นนี้น่าจะทำให้การคำนวณดัชนีบ่งชี้ต่าง ๆ มีการเปลี่ยนแปลงอย่างเป็นระบบขึ้น

3. เนื่องจากแบบสอบที่นำมาใช้ในการวิจัยทั้งสองชุดไม่ได้ควบคุมค่าความยากให้อยู่ในช่วงจำกัด จึงเสนอแนะว่า การวิจัยครั้งต่อไปควรมีการจำกัดค่าความยากให้อยู่ในช่วงแคบๆ เพื่อไม่ให้ค่าความยากมีผลต่อการวิเคราะห์ตัวประกอบ

4. การวิจัยนี้ใช้ข้อสอบเพียง 10 ข้อ และ 15 ข้อ เนื่องจากความจำกัดของโปรแกรม จึงเสนอแนะให้มีการวิจัยที่ใช้จำนวนข้อมากกว่านี้

5. น่าจะได้มีการวิจัยที่ทำการคิดข้อสอบเข้าออกทีละ 1 ข้อ ไปจนครบจำนวนข้อของแบบสอบ เพื่อสังเกตการเปลี่ยนแปลงค่าของดัชนีต่าง ๆ

---