บทที่ 2

วรรณคดีที่เกี่ยวข้อง

การวิจัยในครั้งนี้ ผู้วิจัยได้ทำการศึกษาค้นคว้าวรรณคดีที่เกี่ยวข้อง ซึ่งมีประเด็นสำคัญที่เกี่ยวข้องกับการวิจัยนี้ สรุปได้เป็น 4 ประเด็นใหญ่ ๆ คือ

1. มโนทัศน์เกี่ยวกับความเป็นเอกมิติของแบบสอบ ในส่วนนี้จะได้นำเสนอแนวคิดเกี่ยวกับความเป็นเอกมิติของแบบสอบของนักวัดผลต่าง ๆ พร้อมทั้งวิธีการต่าง ๆ ในการตรวจสอบความเป็นเอกมิติของแบบสอบจากข้อมูล

2. การประยุกต์ดัชนีบ่งชี้ความเป็นเอกมิติของแบบสอบ ในส่วนนี้จะได้กล่าวถึงรายละเอียดเกี่ยวกับที่มาของวิธีการตรวจสอบเดิม 3 วิธี เพื่อให้มองเห็นแนวทางในการพัฒนาดัชนีบ่งชี้ตัวใหม่

3. การใช้โปรแกรม LISREL และโมเดล MIMIC ในการตรวจสอบความเป็นเอกมิติของแบบสอบ

4. ผลงานวิจัยที่เกี่ยวข้องกับการตรวจสอบข้อตกลงเบื้องต้นของ IRT ในเรื่องความเป็นเอกมิติของแบบสอบ

ในแต่ละประเด็นขอนำเสนอรายละเอียด ดังนี้

1. <u>มโนทัศน์เกี่ยวกับความเป็นเอกมิติของแบบสอบ</u>

ด้วยเหตุที่ในปัจจุบันได้ให้ความสนใจกับทฤษฎีคุณลักษณะแฝง (Latent Trait Theory) กันอย่างรวดเร็วมาก จึงได้มีการเน้นถึงความสำคัญของความจำเป็นในการตรวจสอบความเป็นเอกมิติของชุดของข้อสอบกันมาก แต่ก่อนที่จะทำการตรวจสอบนั้นก็พากันหันมามองที่นิยามของเอกมิติ และก็พบว่า นิยามเกี่ยวกับความเป็นเอกมิติของชุดข้อสอบประเภทสองนัย (Binary item) ยังไม่เป็นที่ยอมรับกันอย่างกว้างขวาง ปัญหาที่จะต้องพิจารณาก็คือ เน้นที่โมเดลทางคณิตศาสตร์ รวมทั้งทฤษฎีตัวประกอบร่วม ทฤษฎีคุณลักษณะแฝง และบางครั้งโยงไปถึงทฤษฎีคะแนนจริง ทฤษฎี Generalizability และทฤษฎีมิติร่วม (joint space theory) ว่ากันตามมโนทัศน์ทางคณิตศาสตร์แท้ ๆ ก็คือ สร้างความสัมพันธ์อย่างใกล้ชิดระหว่างสิ่งที่เกี่ยวข้องกับพฤติกรรมของผู้สอบกับสิ่งที่เป็นเนื้อหาข้อสอบ ซึ่งทฤษฎีทางคณิตศาสตร์จะมุ่งสร้างโมเดลนี้ขึ้นมา (Mc Donald, 1981)

Mc Donald (1981:100-117) ชี้ให้เห็นว่า แนวคิดในสมัยต้นปี 1950 เกี่ยวกับเอกมิติเอกพันธ์ (Homogeneous) และความคงที่ภายใน (internal Consistency) ของชุดข้อสอบ เป็นข้อถกเถียงของนักวิทยาศาสตร์สังคม และให้ข้อสังเกตว่า มโนทัศน์เหล่านั้นดูเหมือนต้องทำกับลักษณะของเนื้อหาที่แน่นอน หรือด้วยการวิเคราะห์เชิงสถิติกับผลการตอบของผู้เข้าสอบที่มีต่อสิ่งเร้า คือ ข้อสอบและด้วยพื้นฐานทฤษฎีทางจิตวิทยาการวัด (Psychometric) ต่าง ๆ อย่างไรก็ตามก็ยังไม่มีข้อตกลงร่วมกันว่า เอกมิติควรจะออกมาเป็นถ้อยคำภาษาที่สำคัญว่าอย่างไร

มโนทัศน์ของเอกมิติ หรือ เอกพันธ์ของแบบสอบ ที่สำคัญรวบรวมได้ 3 มโนทัศน์ คือ

1. มโนทัศน์ตามแบบของมาตรากัตแมน (Guttman scale)หรือที่เรียกว่ามาตราการวัดสมบูรณ์แบบ (perfect scale) แนวคิดนี้เกิดจากการเรียงข้อสอบตามระดับความยากง่ายของข้อสอบ และเรียงคะแนนของผู้เข้าสอบตามระดับคะแนนสูงต่ำ ซึ่งเป็นรูปแบบการตอบตามอุดมคติของ Walker (1931) และได้มีการศึกษาอย่างลึกซึ้งโดย Guttman (1950) ซึ่งเป็นที่รู้กันดีอยู่แล้วว่าชุดของข้อสอบที่มาจัดเรียงได้นั้น มีแบบฟอร์มลำดับขั้นที่ว่า ผู้ที่ผ่านข้อสอบข้อหนึ่งๆได้ ก็จะถือผู้นั้นสามารถทำข้อสอบที่มาก่อนข้อสอบข้อนั้นๆได้ทั้งหมด Ferguson (1941)ได้ใช้แนวคิดของเอกมิติมาอภิปรายถึงแนวคิดของตัวประกอบความยาก ส่วน Loevinger (1947) ได้ใช้มโนทัศน์ของมาตราการวัดสมบูรณ์แบบเป็นพื้นฐานของมโนทัศน์ความเป็นเอกพันธ์ และLumsden (1957)ได้ใช้มโนทัศน์นี้เป็นนิยามของแบบสอบเอกมิติ อย่างไรก็ตามแนวคิดนี้มีความขัดแย้งกับมโนทัศน์ของตัวประกอบร่วมเชิงเส้นในเรื่องของคะแนนสอบ (Mc Donald, 1981 : 102) นอกจากนี้ ข้อมูลที่สอดคล้อง (fit) กับรูปแบบของแบบสอบเอกมิติตามแบบของมาตรากัตแมนแล้วก็ยังบอกไม่ได้ว่า คุณลักษณะเอกมิติที่วัดนั้นเป็นอะไร (Hambleton & Swaminathan, 1985: 22)

2. มโนทัศน์ตามแบบของทฤษฎีคุณลักษณะแฝง (การวิเคราะห์โครงสร้างของลักษณะแฝง) ในโมเดลของคุณลักษณะแฝงจะสมมุติให้มีหนึ่งลักษณะแฝงหรือมากกว่า โดยกำหนดให้เป็น $x_1 \ldots\ldots x_r$ ซึ่งจะแสดงคุณสมบัติของผู้เข้าสอบแต่ละคนออกมา โดยที่ค่าของผลการตอบของผู้เข้าสอบที่มีต่อข้อสอบประเภท 0 - 1 ของทุกคนเป็นอิสระจากกันทางสถิติ และถ้ามีหนึ่งคุณลักษณะแฝง x ที่สามารถอธิบายการแจกแจงของรูปแบบการตอบสนองของข้อสอบ n ข้อได้ชุดของข้อสอบนั้นก็เป็นเอกมิติ (Mc Donald , 1985 : 103) หรือจะกล่าวในอีกคำพูดหนึ่งว่า ถ้าสมมุติว่ามีความสามารถเดียว (One ability) หรือคุณลักษณะเดียว (One trait) ที่จำเป็นต่อการอธิบายผลการตอบแบบของผู้เข้าสอบ โมเดลการตอบสนองข้อสอบก็จะถือว่ามีความสามารถ

แฝงเพียงหนึ่งเดียวซึ่งหมายถึงมีความเป็นเอกมิติ (Hambleton & Swaminathan, 1985:16)

3. มโนทัศน์ตามแบบของความคงเส้นคงวา ภายใน (internal consistency) มโนทัศน์นี้สืบเนื่องมาจากความพยายามที่จะเอา คำว่า "เอกพันธ์" (homogeneity) กับคำว่า "คงที่ภายใน" (internal consistency) มาเป็นคำเดียวกัน และนำมาใช้แทนกันในการวัดคุณสมบัติของชุดข้อสอบที่เกี่ยวข้องกับความเป็นเอกมิติ และตัวแอลฟาครอนบาค ก็ถูกนำมาชี้ความเป็นเอกมิติของแบบสอบในที่สุด ซึ่งแนวคิดในการนำเอาความเที่ยงของแบบสอบมาเกี่ยวข้องกับความเป็นเอกมิตินี้ได้ถูกพิสูจน์แล้วว่าไม่มีความสัมพันธ์กันโดย green และ คณะ (Green,1977) เขาได้สรุปว่า การวัดความคงที่ภายในและความเที่ยงไม่ควรที่จะนำมาใช้เป็นดัชนีระบุความเป็นเอกพันธ์ และสัมประสิทธิ์แอลฟาเป็นความคงที่ภายในที่ถูกนำมาใช้ในการประมาณความเที่ยงของแบบสอบ และเขายังอ้างข้อเขียนของ Crano และ Brewer (1973:229)ว่าความคงที่ภายในบรรยายได้ดีที่สุดในเรื่องของความมีดีกรีสูงของความสัมพันธ์ระหว่างกันของข้อสอบ(Green,1977: 828 - 829)

จากมโนทัศน์ของเอกมิติทั้ง 3 แนวคิดนี้ มโนทัศน์ของทฤษฎีคุณลักษณะแฝงดูเหมือนว่าจะเป็นที่ยอมรับมากที่สุด Mc Donald ชี้ให้เห็นว่า มันมีความหมายที่ไปด้วยกันได้ ระหว่างมโนทัศน์เอกมิติตามแบบของคุณลักษณะแฝง กับมโนทัศน์เอกมิติตามแบบของตัวประกอบร่วม แต่บางครั้งความสัมพันธ์ระหว่างมโนทัศน์ทั้งสองก็ไม่อาจเข้าใจได้ (Mc Donald,1981: 103) ที่จริงแล้วทฤษฎีทั่วไปของคุณลักษณะแฝง มีข้อตกลงว่า ผลของการทำข้อสอบจะแสดงถึงชุดของคุณลักษณะ K คุณลักษณะนี้ถูกกำหนดว่า เป็น K มิติแฝงในสเปซ ซึ่งมีที่ตั้งของผู้เข้าสอบแต่ละคนในสเปซแฝงได้ถูกกำหนดขึ้นมาโดยตำแหน่งที่ของผู้เข้าสอบในแต่ละคุณลักษณะแฝง ตำแหน่งที่หรือระดับความสามารถของผู้เข้าสอบที่ระดับคุณลักษณะที่ i มักจะกำหนดเป็น θ_i ดังนั้นตำแหน่งที่ของผู้เข้าสอบใน K มิติแฝงสเปซเขียนแทนด้วย เวคเตอร์ของคะแนนความสามารถดังนี้ $(\theta_1, \theta_2, \ldots, \theta_k)$ สเปซแฝงที่เรียกว่า สมบูรณ์แบบ (Complete) ได้นั้น ถ้าคุณลักษณะแฝงทั้งหมดมีอิทธิพลต่อคะแนนสอบของประชากรผู้เข้าสอบปรากฏขึ้นมา

ในทางปฏิบัติแล้ว IRT ส่วนมากได้ทำข้อตกลงว่า มีความสามารถเดียว หรือคุณลักษณะเดียวก็มีความเพียงพอในการอธิบายผลการตอบของผู้เข้าสอบ และความสัมพันธ์กันระหว่างคู่ของข้อสอบโมเดลการตอบสนองข้อสอบ โมเดลผลการตอบข้อสอบที่ถูกสมมุติให้เป็นความสามารถแฝงเดียวนี้ เรียกว่า เอกมิติ (Unidimensional) ข้อตกลงเบื้องต้นเกี่ยวกับเอกมิติสเปซแฝงนี้

เป็นเรื่องง่ายสำหรับผู้ออก ข้อสอบที่จะสร้างขึ้นมา ทั้งนี้เพราะ โดยปกติแล้วผู้ออกข้อสอบต้องการจะสร้างแบบสอบที่เป็นเอกมิติอยู่แล้ว เพื่อให้เข้าใจในการตีความชุดของคะแนนสอบ แต่แน่นอนที่ข้อตกลงเบื้องต้นเกี่ยวกับเอกมิติ ไม่สามารถพบได้อย่างจริงจัง เพราะว่ามักจะมีพุทธพิสัยอื่น ๆ บุคคลิกภาพและองค์ประกอบของการทำข้อสอบเข้ามามีผลกระทบต่อผลการทำข้อสอบเสมอ ๆ องค์ประกอบเหล่านี้รวมถึงแรงจูงใจในการทำข้อสอบ ความกังวลในการสอบ ความเร็วของการทำข้อสอบ ความซับซ้อนของการสอบ และทักษะพุทธิพิสัยอื่น ๆ เมื่อเป็นเช่นนี้สิ่งที่ข้อตกลงของเอกมิติต้องการให้พบในชุดของข้อสอบก็คือ องค์ประกอบหรือตัวประกอบเด่น (Dominant Component or factor) นั่นเอง องค์ประกอบหรือตัวประกอบเด่นก็คือ ความสามารถที่ถูกวัดโดยแบบสอบ อย่างไรก็ตามความสามารถที่ว่านี้ ไม่ควรถูกสมมุติให้ตายตัวหรือเปลี่ยนแปลงไม่ได้ ความสามารถนี้มีนิยามไว้กว้าง ๆ ว่าเป็นอะไรก็ได้ ที่แบบสอบวัด เช่น ความสามารถทางพุทธิพิสัย การวัดผลสัมฤทธิ์ ทักษะหรือศักยภาพพื้นฐาน หรือเป็นตัวแปรบุคคลิกภาพ อะไรที่เป็นความสามารถจะต้องถูกสร้างขึ้นมาด้วยวิธีการเดียวกัน คือ ภาวะสันนิษฐาน (Construct) ที่ถูกวัดด้วยแบบวัดใด ๆ นั้นจะต้องมีการแสดงให้เห็นว่าได้มีการผ่านการตรวจสอบความตรงของภาวะสันนิษฐานนั้น ๆ แล้ว

วิธีการตรวจสอบความเป็นเอกมิติของข้อมูลหรือของแบบสอบนั้น ในทางปฏิบัติสามารถทำได้หลายวิธี ซึ่งมักจะมีวัตถุประสงค์เพื่อเป็นการทดสอบสมมุติฐานเกี่ยวกับข้อตกลงเบื้องต้นของทฤษฎีคุณลักษณะแฝง ก่อนที่จะนำข้อสอบนั้นไปวิเคราะห์ตามแบบของ IRT Hattie (1984) ได้รวบรวมดัชนีบ่งชี้ (indices) ทั้งหมด 8 ตัว เพื่อประเมินความเป็นเอกมิติของแบบสอบและเขาก็พบว่ามีหลายวิธีซึ่งเป็นวรรคดีเก่าให้ผลไม่พึงพอใจ ในขณะที่การใช้วิธีการพื้นฐานของการวิเคราะห์ตัวประกอบแบบไม่เชิงเส้น (non-linear factor analysis) และวิธีวิเคราะห์ค่าที่เหลือ(analysis of residuals) จะประสบความสำเร็จมากที่สุด Hambleton (1989: 174) ได้สรุปวิธีการตรวจสอบความเป็นเอกมิติของแบบสอบ ซึ่งใช้วิธีการพิจารณาทางสาระรูปสนิทดี (goodness of fit) ไว้ 6 วิธีดังนี้

1. พล็อตค่าไอเก็น (จากค่ามากที่สุดไปหาค่าน้อยที่สุด)ของเมตริกซ์สหสัมพันธ์ระหว่างข้อ (ใช้สัมประสิทธิ์สหสัมพันธ์เตตรา ให้ผลดีกว่าการใช้สัมประสิทธิ์สหสัมพันธ์ฟี) และศึกษาว่ามีตัวประกอบเด่นตัวแรกเกิดขึ้นหรือไม่ Reckase (1979) ซึ่งเป็นผู้เสนอวิธีการนี้ได้ชี้ให้พิจารณาถึงตัวประกอบเด่นตัวแรกซึ่งเป็นตัวที่มีอิทธิพลและค่าไอเก็นตัวแรกที่เกิดจากการวิเคราะห์ ตัวประกอบสำคัญ (principal component analysis) เป็นค่าที่บอกความแปรปรวนที่ตัวประกอบตัวแรกสามารถอธิบายได้ทั้งหมด ซึ่งตามเกณฑ์ของ Reckase แล้ว เกณฑ์ความแปรปรวนของตัว

ประกอบตัวแรก ควรอธิบายความแปรปรวนทั้งหมดได้ อย่างน้อย 20 เปอร์เซนต์ และเกณฑ์ค่าไอเก็นนั้นควรมีค่าเป็น 10 หรือมากกว่ากรณีที่ข้อสอบมี 50 ข้อ

2. เปรียบเทียบพล็อต(plot)ของค่าไอเก็น 2 แหล่งคือ ค่าไอเก็นที่เกิดจากเมตริกซ์สหสัมพันธ์ระหว่างข้อของข้อมูลเชิงประจักษ์ กับค่าที่เกิดจากเมตริกซ์สหสัมพันธ์ระหว่างข้อมูลเชิงสุ่มที่มีการแจกแจงเป็นปกติของข้อมูลชุดหนึ่ง ที่มาจากกลุ่มตัวอย่างและจำนวนตัวแปรเดียวกันกับข้อมูลจริง ค่าพล็อตของไอเก็นทั้ง 2 แหล่งจะถูกนำมาเปรียบเทียบกัน ถ้าข้อมูลมีข้อตกลงเบื้องต้นของความเป็นเอกมิติในข้อมูลจริงแล้ว พล็อตทั้งสองแหล่งก็ควรจะมีลักษณะเหมือนกัน ยกเว้น ค่าไอเก็นตัวแรกของการพล็อตในข้อมูลจริง ซึ่งมันควรจะมีค่าใหญ่กว่าในข้อมูลสุ่ม วิธีการนี้เสนอโดย Horn (1965) และตัวอย่างการนำวิธีการนี้ไปใช้ดูได้จากงานของ Montanelli and Humphreys (1976) และ Drasgow and Lissak (1983)

3. ตรวจสอบข้อตกลงเบื้องต้นที่เกี่ยวกับความเป็นอิสระจากตำแหน่ง (local independence) โดยตรวจสอบจากเมตริกซ์ความแปรปรวนและความแปรปรวนร่วมหรือเมตริกซ์สหสัมพันธ์ของผู้เข้าสอบที่มีช่วงความสามารถแตกต่างกัน หรือมีมาตราของคะแนนการสอบแตกต่างกัน (Mc Donald, 1981: Tucker, Humphreys & Roznowski, 1986) เมื่อสมาชิกนอกแนวทแยง (off-diagonal) ของเมตริกซ์มีค่าน้อยและใกล้ศูนย์ ก็แสดงว่าพบความเป็นเอกมิติแล้วโดยประมาณ

4. ดูความสอดคล้อง(fit)ของโมเดลวิเคราะห์ตัวประกอบเดียวแบบไม่เป็นเส้นตรง (nonlinear one-factor analysis model) กับเมตริกซ์สหสัมพันธ์ระหว่างข้อและศึกษาค่าที่เหลือจากความสอดคล้อง (McDonald, 1981; Hattie, 1985) Mc Donaldเป็นคนแรกที่เสนอให้มีการใช้การวิเคราะห์ตัวประกอบในการศึกษาความเป็นเอกมิติ โดยสามารถควบคุมความสัมพันธ์ที่ไม่ได้อยู่ในรูปของเส้นตรงระหว่างคู่ของข้อสอบและระหว่างข้อสอบกับความสามารถ (ซึ่งเป็นข้อกำหนดอยู่ในเรื่องของสเปซพหุมิติ) ตัวอย่างการใช้วิธีนี้ดูได้จาก Hambleton & Rovinelli (1986)

5. การใช้วิธีวิเคราะห์ตัวประกอบที่ขึ้นกับ IRT โดยตรง (Bock, Gibbons, & Maraki, 1985) วิธีนี้จะใช้การมองภาพพหุมิติของโค้งปกติโอจีฟชนิด 3 พารามิเตอร์ (Three parameter normal ogive) ให้สมมุติว่า เป็นการอธิบายเวคเตอร์ของการตอบสนองข้อสอบ การประมาณของพารามิเตอร์โมเดล เป็นเรื่องที่ต้องใช้เวลามากและซับซ้อน แต่ผลที่ได้รับออกมานั้นคุ้มค่า และที่น่าสนใจคือ เหมาะแก่การแก้ปัญหาหนึ่งมิติ

I1696763X

6. ข้อสอบที่ดูเหมือนว่า ฝ่าฝืนข้อตกลงเบื้องต้นจะถูกตรวจสอบให้เห็นได้ ถ้ามันทำหน้าที่ของมันแตกต่างออกไป ค่า b ของข้อสอบเหล่านี้จะถูกคำนวณแยกออกมาจากข้อสอบ จากนั้น ก็จะมีการวิเคราะห์เนื้อหา (context) ของข้อสอบที่เหลือต่อไปอีก เนื้อหาของข้อสอบที่คำนวณออกมาแล้วมีความสำคัญหนึ่งเดียวก็แสดงว่า ข้อตกลงของโมเดลถูกต้องแล้ว และถ้าพล็อตค่า b ที่คำนวณในสองเนื้อหาออกมาในเป็นเส้นตรง ข้อตกลงเกี่ยวกับเอกมิติถือว่าใช้ได้(Bejar,1980)

ความจริงแล้วความเป็นเอกมิติของแบบสอบเกี่ยวกับเนื้อหาวิชา ที่ต้องการวัดนั้นออกจะกว้างมากจนแทบจะบอกความเป็นเอกมิติไม่ได้ และยิ่งในปัจจุบันนี้การเรียนการสอนในห้องเรียนก็เน้นถึงลักษณะของการบูรณาการความรู้มาก เมื่อเป็นเช่นนี้ แนวปฏิบัติเกี่ยวกับการเรียนการสอนของครู จึงสวนทางกับแนวปฏิบัติ ของทฤษฎีทางด้านวัดผล และเมื่อเป็นเช่นนี้ นิยามของเอกมิติก็คงต้องหลากหลายจนไม่อาจเป็นเอกมิติที่แท้จริงได้ Warm (1978:101-107) ยังกล่าวข้อตกลงเบื้องต้นของเอกมิติมีความซับซ้อนมากที่สุด และยังเป็นข้อตกลงที่เข้มงวดที่สุดของ IRT และเมื่อเป็นเช่นนี้แล้ว ในฐานะของนักวัดผลควรจะทำอย่างไร จึงได้เกิดมีข้อถกเถียงกันในเรื่องนิยามของ เอกมิติ Warm จึงได้สรุปนิยามของเอกมิติไว้ 2 หลักเกณฑ์ คือ

1. เอกมิติ หมายถึง ข้อสอบที่จัดเพียง 1 สาขาวิชาของความรู้ ความสามารถเดียว แต่ข้อสอบเหล่านี้ อาจมีความสัมพันธกันได้ทั้งทางบวกและทางลบ และว่ากันโดยหลักแล้วแบบสอบที่มองดูแล้วว่ามีความเป็นเอกมิติ บางทีมันมีความเป็นเอกมิติจริง

2. เอกมิติ หมายถึง ข้อสอบที่สอบความรู้ที่ได้เรียนรู้ร่วมกันมา บางทีก็มีความเป็นเอกมิติเช่นกัน ดังนั้นข้อสอบไล่ประจำวิชาอาจตัดสินให้มีความเป็นเอกมิติก็ได้ ดังตัวอย่างการทดลองของ Bejar, Weiss และ Kingsbury (1977) สรุปได้ว่า ความเป็นเอกมิติของแบบสอบอาจมีผลเนื่องมาจากข้อสอบที่ทดสอบความรู้นั้น ได้มีการเรียนรู้ร่วมกันในห้องเรียน ดังนั้นเขาจึงมีกฎเกณฑ์ข้อหนึ่งขึ้นมาว่า ข้อสอบที่ทดสอบความรู้ต่าง ๆ ที่มีความสัมพันธ์เชิงตรรกะและเชิงลำดับขั้น อาจเป็นแบบสอบที่มีความเป็นเอกมิติก็ได้ (Bejar, Weiss & Kingsbury,1977 อ้างใน Warm, 1978:101) warm ยังกล่าวอีกว่า ไม่มีวิธีการใดที่จะทดสอบความเป็นเอกมิติของข้อสอบแบบเลือกตอบได้อย่างสมบูรณ์แบบ ด้วยเหตุผลที่ว่า การทดสอบความเป็นเอกมิติส่วนมากใช้การวิเคราะห์ตัวประกอบของเมตริกซ์สหสัมพันธ์แบบเตตราคลอริกระหว่างข้อสอบ แต่ก็โชคไม่ดีที่สหสัมพันธ์แบบเตตราคลอริกนั้นต้องมี การสมมติให้ค่าความสามารถ (O) มีการแจกแจงเป็นปกติ ซึ่งก็จะไม่มีวันเป็นไปได้ เมื่อค่าการเดา (c) ไม่ได้เท่ากับศูนย์ Warm กล่าวว่า Cristoferson (1975) ได้พยายามอย่างที่สุดที่จะพัฒนาวิธีการสอบความเป็นเอกมิติ แต่ก็ต้องใช้วิธีการทาง

คณิตศาสตร์ที่ซับซ้อน ตัวเขาเองก็มีวิธีการทดสอบความเป็นเอกมิติ 8 วิธี ซึ่งก็พบว่า 6 ใน 8 วิธี นี้ได้ใช้วิธีการวิเคราะห์ตัวประกอบ ที่มี 4 ขั้นตอนแรกเหมือนกัน คือ

1. ทำผลการตอบข้อสอบของผู้เข้าสอบให้เป็นคะแนน 0 และ 1 โดยให้ผิดเป็น 0 ถูกเป็น 1

2. คำนวณเมตริกซ์สหสัมพันธ์แบบเตตราคลอริก

3. แทนค่าในแนวทะแยงของเมตริกซ์ ด้วยค่าสหสัมพันธ์ที่มีค่ามากที่สุดในแถว (ซึ่งเป็นวิธีการที่โปรแกรมคอมพิวเตอร์ส่วนใหญ่เลือกใช้

4. วิเคราะห์ตัวประกอบด้วยวิธีวิเคราะห์ตัวประกอบสำคัญ(Principal compoment or principal axis) เพื่อให้ได้ตัวประกอบ 9 ตัวแรก (เลข 9 เป็นเลขที่กำหนดขึ้นเป็นตัวอย่าง)

Warm ได้เสนอวิธีทดสอบความเป็นเอกมิติ 8 วิธีที่เขารวบรวมมาได้ ดังนี้

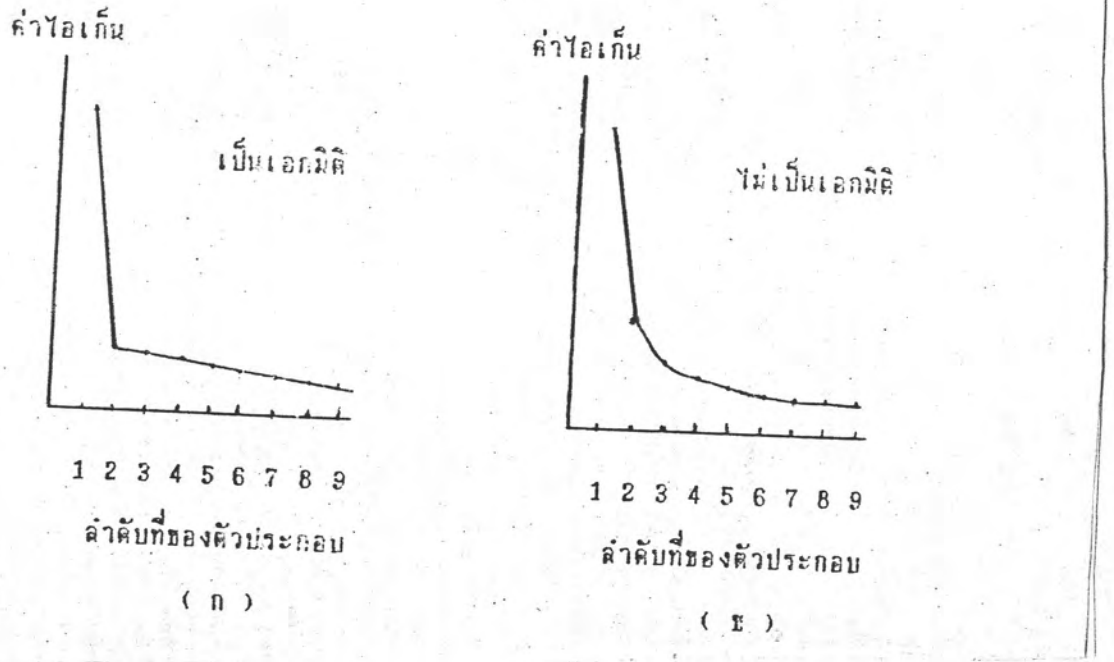
1. การทดสอบด้วยค่าไอเก็น (The Eigenvalue Test) วิธีการก็คือ พล๊อตค่าไอเก็นของตัวประกอบต่าง ๆ กับค่าตำแหน่งของตัวประกอบ ดังรูป 2.1 ก. และ ข. จะสังเกตว่า ข้อสอบที่ถูกตัดสินว่ามีความเป็นเอกมิติ เมื่อค่าไอเก็นของตัวประกอบที่ 1 มีค่าใหญ่โดยเทียบกับตัวประกอบที่ 2 และค่าไอเก็นของตัวประกอบที่เหลือมีค่าพอๆกัน กราฟที่มีลักษณะคล้ายๆกับกราฟ 2.1 ก. เมื่อข้อสอบเป็นเอกมิติและถ้ากราฟมีลักษณะ ดังรูป 2.1 ข. ข้อสอบก็จะไม่เป็นเอกมิติ (Lord & Novick,1968 อ้างใน Warm,1978:104)

2. การทดสอบเส้นฐานเชิงสุ่ม (The Random Baseline Test) การทดสอบนี้เป็นส่วนเพิ่มเติมของการทดสอบด้วยค่าไอเก็น ซึ่งวิธีนี้จำเป็นต้องทำการทดสอบด้วยค่าไอเก็นก่อน และเมื่อจะทดสอบด้วยเส้นฐานเชิงสุ่ม ก็จะต้องสร้างข้อมูลเชิงสุ่มที่เป็นเมตริกซ์ 0 และ 1 ที่มีลำดับเช่นเดียวกับเมตริกซ์สหสัมพันธ์ระหว่างข้อ จากขั้นตอนในการทำวิเคราะห์ตัวประกอบ แล้วทำการพล๊อตค่าไอเก็นจากข้อมูลเชิงสุ่มในกราฟเดียวกันกับการพล๊อตค่าไอเก็นจากข้อมูลเชิงประจักษ์ ความเป็นเอกมิติจะแสดงออกมาเมื่อค่าไอเก็นของตัวประกอบตัวแรกจากข้อมูลทั้ง 2 ชุดแตกต่างกัน (Mc Bride & Weiss, 1974: P 30 อ้างใน Warm, 1978: P 104)ดังรูป 2.2 ก. และ 2.2 ข.

3. การทดสอบด้วยไบซีเรียล (The Biserial Test) วิธีการนี้ให้คำนวณค่าสหสัมพันธ์ไบซีเรียลระหว่างข้อสอบกับแบบสอบทั้งฉบับ แล้วนำค่าสหสัมพันธ์นี้มาสัมพันธ์กับค่าน้ำหนักของข้อสอบบนตัวประกอบตัวแรก ถ้าค่าความสัมพันธ์นี้ให้ค่าสหสัมพันธ์สูง (.80 หรือ สูงกว่า) ก็

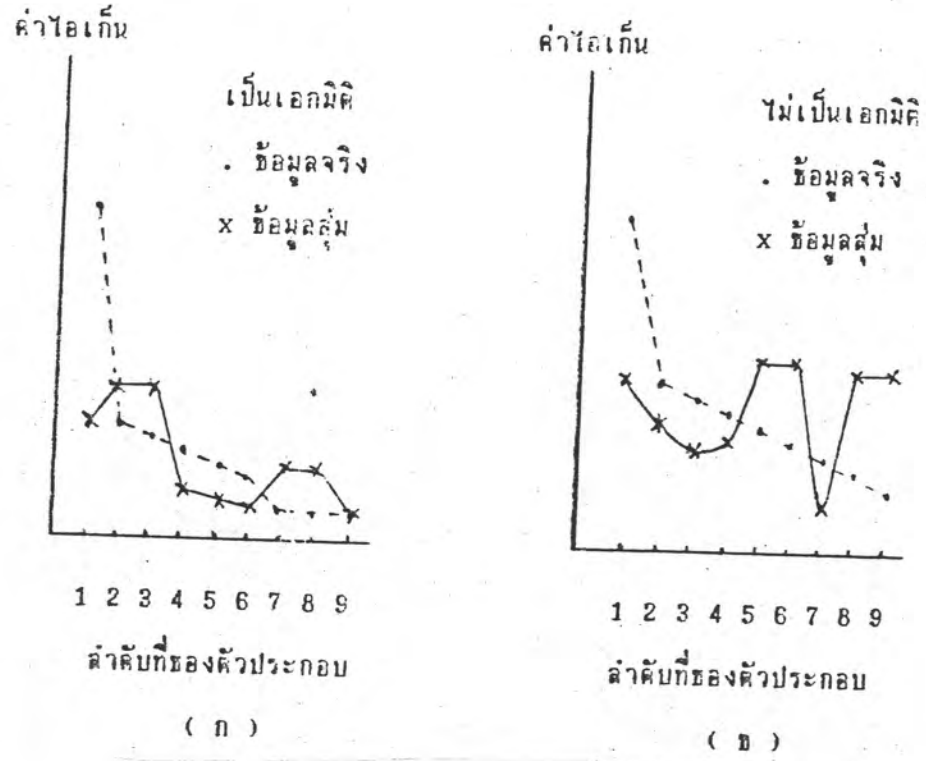
สนับสนุนข้อตกลงเบื้องต้นของความเป็นเอกมิติ (Mc Bride & Weiss, 1974 อ้างใน Warm, 1978 : 104)

4. การทดสอบค่าน้ำหนักขององค์ประกอบ (The Factor Loading Test) วิธีการนี้ความเป็นเอกมิติจะสังเกตที่น้ำหนักบนตัวประกอบ ถ้าน้ำหนักของตัวประกอบตัวแรกของข้อสอบทั้งหมดมีนัยสำคัญและมีเครื่องหมายเหมือนกัน (คือ + หรือ -) (Mc Bride & Weiss, 1974 อ้างใน Warm, 1978 : 104)



รูป 2.1 ก. และ ข. แสดงการทดสอบความเป็นเอกมิติด้วยค่าไอเก็น

(ดัดแปลงจาก Warm 1978 : p 105)



รูป 2.2 ก. และ ข. แสดงการทดสอบความเป็นเอกมิติด้วยเส้นฐานเชิงสุ่ม

(ดัดแปลงจาก Warm 1978 : 105)

5. ทดสอบความคล้องจอง (The Congruence Test) วิธีนี้ ถ้าผู้เข้าสอบสามารถแบ่งออกได้เป็น 2 กลุ่มย่อยอย่างมีความหมาย เช่น ผู้ชายและผู้หญิงหรือ ผิวดำและผิวขาว เป็นต้น แล้วทำการวิเคราะห์ตัวประกอบของแต่ละกลุ่ม ค่าสัมประสิทธิ์ความคล้องจอง (Coefficient of congruence) (C_{AB}) ของน้ำหนักของข้อสอบบนตัวประกอบที่ 1 ระหว่างกลุ่มทั้งสองจะเข้าใกล้ศูนย์ ความเป็นเอกมิติจะขึ้นอยู่กับตัวแปรบนกลุ่มที่กำหนดขึ้น สูตรสัมประสิทธิ์ของความสอดคล้องเป็นดังนี้ (Tommell,1970 อ้างใน Warm,1978:106)

$$C_{AB} = \sqrt{\frac{\sum_{i}^{n}(L_{ia} - L_{ib})^2}{n}}$$

เมื่อ L_{ia} = น้ำหนักของข้อสอบ i บนตัวประกอบที่ 1 ของกลุ่ม a

L_{ib} = น้ำหนักของข้อสอบ i บนตัวประกอบที่ 1 ของกลุ่ม b

n = จำนวนข้อสอบในแบบสอบ

C_{AB} = สัมประสิทธิ์ของความคล้องจองระหว่างกลุ่ม A และกลุ่ม B

6. ทดสอบด้วยค่าคอมมูนาลิตี (The Communality Test) วิธีนี้วอร์มดัดแปลงมาจากสูตรดัชนีความเอกพันธ์ของ Green (1977, 836) โดยใช้ค่า G แทนตัว μ ของกรีน ดังนี้

$$G = \frac{\sum\sum \frac{r_{ij}}{\sqrt{h_i^2 h_j^2}}}{n} \qquad i>j$$

เมื่อ r_{ij} = ค่าสหสัมพันธ์เตตราคลอริกระหว่างข้อสอบ

h_i, h_j = ค่าคอมมูนาลิตีของข้อสอบ i และ j ตามลำดับ

π = จำนวนค่าสหสัมพันธ์

ค่า G เข้าใกล้ 1 ก็แสดงว่าข้อสอบเข้าใกล้ความเป็นเอกมิติ

7. การทดสอบบางส่วนและทั้งหมด (The Part / Whole Test) วิธีนี้จะแยกข้อสอบออกเป็น 2 ส่วนตามเนื้อหา แล้วประมาณค่า a และค่า b (ค่าอำนาจจำแนกและค่าความยาก)ของข้อสอบแต่ละส่วนและของข้อสอบทั้งฉบับ ถ้าพารามิเตอร์ที่ประมาณได้ภายใต้ 2 เงื่อนไข (บางส่วนและทั้งหมด)มีความสัมพันธ์กันสูงก็สนับสนุนว่าข้อสอบเหล่านั้นมีความเป็นเอกมิติ (Bejar, 1977:107)

8. การทดสอบความถี่ของเวคเตอร์ (The Vector Frequency Test)วิธีนี้สมมุติให้ค่า มีการแจกแจงเป็นปกติ และกำหนดพารามิเตอร์ของข้อสอบมาให้ ก็สามารถคำนวณค่าความถี่คาดหวังของแบบแผนการตอบสนองที่เป็นไปได้ทั้งหมด การเปรียบเทียบค่าความถี่ที่สังเกตได้กับค่าความถี่ที่เป็นไปได้ทั้งหมดของแบบแผนการตอบ ถ้าแบบสอบมีความเป็นเอกมิติอยู่ก็จะทำให้ได้ค่าไคสแควร์ที่ไม่มีนัยสำคัญ (Bock & Lieberman,1970 อ้างใน Warm,1978:107)

การตรวจสอบความเป็นเอกมิติที่ใช้วิธีการสะกัดตัวประกอบด้วย Principal component analysis มีมโนทัศน์พื้นฐานของนิยามเกี่ยวกับค่าแรงค์ ของเมตริกซ์อยู่ หรืออาจกล่าวได้ว่า แรงค์ของเมตริกซ์สหสัมพันธ์ระหว่างข้อ (Rank of interitem correlation matrix)

เป็น ดัชนีบ่งชี้ความเป็นเอกมิติของแบบสอบได้ Muliak (1972 : 42-43) ได้ให้นิยามแรงค์ของเมตริกซ์ A ขนาด m x n ว่าเป็นจำนวนเวคเตอร์คอลัมภ์อิสระเชิงเส้นสูงสุดใน A และแนวคิดเกี่ยวกับแรงค์ของเมตริกซ์ ก็สอดคล้องกับแนวคิดของมิติในสเปซที่บรรจุอยู่ในคอลัมภ์เวคเตอร์ของเมตริกซ์ แรงค์ของเมตริกซ์เป็นมโนทัศน์ที่สำคัญในการวิเคราะห์ตัวประกอบ Thurstone (1977 อ้างใน Muliak, 1972 : 43) ได้ใช้แรงค์ในการกำหนดจำนวนตัวประกอบที่จำเป็นในการอธิบายความแปรปรวนร่วมในระหว่างชุดของตัวแปร การหาค่าแรงค์ของเมตริกซ์สามารถใช้ดีเทอร์มีเน้นท์ (determenant) ในการคำนวณได้ โดยใช้หลักการที่ว่า ถ้าเราทำเมตริกซ์ย่อยชนิดจตุรัส (square submatrics)จากสมาชิกในแถวและคอลัมภ์ของเมตริกซ์หนึ่งๆ ซึ่งก็จะให้ค่าดีเทอร์มีเน้นท์ของเมตริกซ์ย่อยเหล่านี้ได้ และดีเทอร์มีเน้นท์ของเมตริกซ์ย่อยเหล่านี้ ซึ่งรู้จักกันในชื่อที่เรียกว่าไมเนอร์ของเมตริกซ์ จากนั้นแรงค์ของเมตริกซ์ก็คือ อันดับของไมเนอร์เมตริกซ์ที่ใหญ่ที่สุดที่มีค่าไม่เป็นศูนย์ นอกจากนี้ ยังมีวิธีการหาค่าแรงค์ของเมตริกซ์ด้วยวิธีอื่น ๆ อีกเช่นวิธีของการจัดกระทำในแนวแถว (Row operation)

เนื่องจาก การหาค่าแรงค์ของเมตริกซ์ขนาดใหญ่มีความซับซ้อนมาก ดังนั้นในบางครั้งการหาค่าแรงค์จึงต้องใช้วิธีการประมาณ และความรู้เกี่ยวกับมโนทัศน์ของค่าแรงค์ต่ำสุด (Concept of minimum rank) จึงเป็นสิ่งจำเป็นที่ต้องรู้ การประมาณค่าคอมมูนาลิตี (Communalities) ในเมตริกซ์สหสัมพันธ์ระหว่างข้อมีผลต่อการประมาณค่าแรงค์ของเมตริกซ์ โดยที่แกรมเมียนเมตริกซ์ (gramian matrix) จะให้ค่าไอเก็นที่ไม่เป็นลบเท่านั้น และแรงค์ของมันก็คือจำนวนของค่าไอเก็นที่เป็นบวกที่เมตริกซ์นั้นมีอยู่ แล้วค่าไอเก็นที่เหลือจะเป็นศูนย์ Thurstone ตระหนักว่า ถ้าคอมมูนาลิตีของตัวแปรถูกประมาณค่าสูงกว่าที่ควรจะเป็น (overestimated) ตัวประกอบก็จะอธิบายสหสัมพันธ์นอกแนวทแยงได้มากกว่า และไม่ได้ลดค่าแรงค์ของเมตริกซ์ ซึ่งขัดแย้งกับโมเดล ในทางตรงข้าม ถ้าค่าคอมมูนาลิตีถูกประมาณต่ำกว่าที่ควรจะเป็น ก็จะได้เมตริกซ์ที่มีค่าแรงค์สูง ซึ่งเป็นผลเนื่องมาจากการได้เมตริกซ์ที่ไม่เป็นแกรมเมียน (non-gramian) ที่ซึ่งให้เมตริกซ์ที่มีค่าไอเก็นเป็นลบ Thurstone ย้ำว่า แรงค์ของจัสตุรัสเมตริกซ์สหสัมพันธ์ที่สมมาตรกัน จะมีค่าเท่ากับจำนวนของค่าไอเก็นของเมตริกซ์ที่มีค่าเป็นบวกและลบ แต่ตัวประกอบทั้งหลายที่สอดคล้องกับค่าไอเก็นที่เป็นลบนั้น ในทางเรขาคณิต (geometric) เรียกว่า เป็นมิติของจำนวนจินตภาพ (imaginary numbers) หรือถ้าพูดในเทอมของความแปรปรวนก็กล่าวว่าคะแนนตัวประกอบของตัวประกอบเหล่านี้ให้ค่าความแปรปรวนที่มีค่าเป็นลบซึ่งไม่มีความหมายอะไร Thurstone จึงเชื่อว่า มีจุดซึ่งอยู่ระหว่างค่าคอมมูนาลิตีที่ประมาณสูงกว่าความเป็นจริง และต่ำกว่าความเป็นจริง (overestimated and lowerestimated)ที่ซึ่งแรงค์ของเมตริกซ์สหสัมพันธ์

หวลกลับ (reduced correlation matrix) จะมีค่าต่ำสุด

2. การประยุกต์ดัชนีบ่งชี้ความเป็นเอกมิติของแบบสอบ 3 วิธี

ดังที่ได้กล่าวมาแล้ว จะเห็นว่าวิธีการในการบ่งชี้ความเป็นเอกมิติของแบบสอบนี้ สามารถทำการตรวจสอบได้หลายวิธี บางวิธีก็มีแนวคิดหลักการคล้าย ๆ กัน เช่น วิธีการบ่งชี้ด้วยผลลัพธ์หลังการสะกัดตัวประกอบสำคัญ ซึ่งมีทั้งค่าไอเก็น ค่าน้ำหนักของตัวแปรบนตัวประกอบ ค่าความแปรปรวนของตัวประกอบที่ 1 เปอร์เซ็นต์ของตัวประกอบร่วมตัวแรก ฯลฯ และบางวิธีก็มีการกำหนดค่าเป็นดัชนีบ่งชี้ความเป็นเอกมิติที่แน่นอน เช่น วิธีการทดสอบด้วยไปซีเรียล บางวิธียังอยู่ในสูตรที่น่าจะได้รับการพัฒนาต่อไป เพื่อใช้การบ่งชี้ความเป็นเอกมิติ ให้ได้ผลที่มีคุณภาพมากขึ้น ในที่นี้ได้มีการประยุกต์วิธีการบ่งชี้ 3 วิธีเดิม เป็น 3 วิธีใหม่ ดังนี้

2.1 การประยุกต์วิธีการบ่งชี้ความเป็นเอกมิติของแบบสอบที่ใช้การพล็อตค่าไอเก็น

การบ่งชี้ความเป็นเอกมิติของแบบสอบที่ใช้การพล็อตค่าไอเก็น เป็นวิธีการที่ได้รับความนิยมมาก เพราะการสะกัดตัวประกอบสำคัญ(Principal factor anallysis)สามารถทำได้ด้วยคำสั่ง FACTOR ในโปรแกรมสำเร็จ SPSS พร้อมทั้งสามารถสั่งให้มีการแสดงผลการพล็อตกราฟระหว่างค่าไอเก็นกับลำดับที่ของตัวประกอบร่วม ทำให้สามารถพิจารณาจากกราฟและบอกได้เลยว่าข้อสอบชุดนั้นมีความเป็นเอกมิติหรือไม่ การพิจารณาก็เพียงแต่ดูตำแหน่งของค่าไอเก็น ตัวที่ 1 ตัวที่ 2 ตัวที่ 3 และตัวอื่น ๆ ประกอบบ้าง โดยถือว่าตำแหน่งของค่าไอเก็นตัวที่ 1 อยู่สูงกว่าตำแหน่งของค่าไอเก็นตัวที่ 2 มาก ๆ และตำแหน่งของค่าไอเก็น ตัวที่ 3 และตัวอื่น ๆ อยู่ในระดับเดียวกันและใกล้เคียงกับตำแหน่งของค่าไอเก็นตัวที่ 2 (ดูจากรูป 2.3 ก. และ 2.3 ข.) การพิจารณาเช่นนี้ บางครั้งก็อาจจะเกิดจุดที่ต้องทำการตัดสินใจยาก ทำให้การบอกความเป็นเอกมิติของแบบสอบโดยวิธีนี้ค่อนข้างเป็นอัตนัย Reckase ได้พยายามเสนอให้ใช้ค่าอัตราส่วนระหว่างค่าไอเก็นตัวที่ 1 กับค่าไอเก็นตัวที่ 2 (E_1/E_2) แต่ก็ไม่ได้บอกเกณฑ์ที่ชัดเจนว่า ควรจะใช้ค่าตั้งแต่เท่าไรขึ้นไป

ที่มาของการมองเห็นความสำคัญของค่าไอเก็นในการอธิบายความเป็นเอกมิติ อยู่ที่ว่านักวัดผลพยายามหาค่าแรงค์เพื่อบอกจำนวนของมิติของเมตริกซ์สหสัมพันธ์ระหว่างข้อโดยอาศัยมโนทัศน์ของค่าแรงต่ำสุด (Concept of minimum rank) ที่ว่าการประมาณค่าคอมมูนาลิตีมีผลต่อการประมาณค่าแรงค์ของเมตริกซ์สหสัมพันธ์ระหว่างข้อและแกรมเมียนเมตริกซ์ จะให้ค่าไอเก็น

ที่เป็นบวกเท่านั้น และค่าแรงค์ของแมตริกซ์นี้ก็จะมีค่าเท่ากับจำนวนค่าไอเก็นที่เป็นบวก(Mulaik, 1972:136) Thurstone กล่าวว่า การประมาณค่า h^2 ที่ต่ำกว่าความเป็นจริง (underestimate) ก็จะทำให้ matrix มีค่าแรงค์สูง และยังส่งผลให้ค่าไอเก็นที่เป็นลบอีก ดังนั้น การทำการวิเคราะห์ตัวประกอบด้วยวิธีการวิเคราะห์ตัวประกอบสำคัญซึ่งไม่มีการประมาณค่าคอมมูนาลิตี จะทำให้ได้ตัวแทนของคุณสมบัติของบุคคลสัก 2-3 ตัว(a few central characteristic of the person) ซึ่งเรียกว่าคุณลักษณะแฝง (Latent Trait) (Green,1978 : 342)การใช้วิธีการวิเคราะห์ตัวประกอบโดยการสะกัดแบบตัวประกอบสำคัญ จึงทำให้ได้จำนวนไอเก็นที่เป็นตัวบอกจำนวนแรงค์ชั้นต่ำ และตัวประกอบที่สะกัดมาได้ ก็จะเป็นตัวแทนของคุณลักษณะแฝง ส่วนค่าไอเก็นซึ่งได้มาจากการวิเคราะห์ตัวประกอบโดยวิธีนี้ จะมีค่ามากที่สุดบนตัวประกอบแรกและลดต่ำลงบนตัวประกอบ ต่อ ๆ ไป ทำให้การพิจารณาเกี่ยวกับความเป็นเอกมิติ สามารถอาศัยค่าไอเก็นตัวแรก ซึ่งเกิดจากตัวประกอบตัวแรกเป็นคุณลักษณะเด่นที่ต้องการวัด และควรที่จะมีค่ามากกว่าค่าไอเก็นตัวที่ 2 และตัวต่อ ๆ ไปมาก ๆ จึงจะถือว่าคุณลักษณะเด่นที่ต้องการวัดนั้นมีตัวเดียวจริง ๆ ส่วนที่เหลืออยู่นั้น มีค่าน้อนจนถือว่าเป็นความคลาดเคลื่อน

ด้วยเหตุนี้ นักวัดผลจึงนิยมที่จะสังเกตค่าไอเก็นตัวที่ 1 ตัวที่ 2 และ ตัวที่ 3 มาเป็นเครื่องพิจารณาความเป็นเอกมิติของแบบทดสอบด้วยวิธีการต่าง ๆ เช่น จากการพล็อตค่าไอเก็นเรียงตามลำดับที่ของตัวประกอบหรือเทียบตัวประกอบตัวที่ 1 กับตัวที่ 2 บ้างตามที่ได้กล่าวมาแล้ว

แต่จุดอ่อนของวิธีการพิจารณาเหล่านี้ อยู่ที่การสังเกตจากกราฟ ซึ่งใช้รูปร่างกราฟเป็นเครื่องตัดสินซึ่งค่อนข้างเป็นการตัดสินที่มีความเป็นอัตนัยมาก ในที่นี้ได้ใช้แนวคิดของ Rackase เป็นตัวระบุความเป็นเอกมิติของแบบสอบ

โดยใช้สูตร $ER = E_1 / E_2$

เมื่อ ER คือ อัตราส่วนของค่าไอเก็น

E_1 คือ ค่าไอเก็น ตัวที่ 1

E_2 คือ ค่าไอเก็น ตัวที่ 2

ER มีค่าเป็นบวกและอยู่ระหว่างค่า 0 ถึง 1.000 โดยที่ค่า ER ยิ่งสูง ก็ยิ่งแสดงความเป็นเอกมิติสูง

2.2 การประยุกต์ดัชนีบ่งชี้ความเป็นเอกมิติของแบบสอบที่ใช้ไบซีเรียลทดสอบ

การใช้ไบซีเรียลทดสอบ ก็เป็นวิธีการบ่งชี้ความเป็นเอกมิติของแบบสอบที่มีผู้นิยมใช้อีกวิธีหนึ่ง วิธีการใช้ไบซีเรียลทดสอบนี้ ที่มาของการตรวจเอกมิติโดยวิธีนี้อยู่ที่ว่า ผู้พัฒนาแต่แรกนั้น มีความเชื่อว่าถ้าข้อสอบแต่ละข้อมีความคงที่ภายใน ก็จะส่งผลให้ข้อสอบเหล่านี้ มีคุณสมบัติสอดคล้องกับมิติที่ต้องการวัดได้ นั่นคือการหาความสัมพันธ์ระหว่างข้อสอบกับคะแนนรวมของข้อสอบฉบับนั้น จะให้สาระสนเทศที่สำคัญว่าคนที่ทำข้อสอบชุดนี้ได้คะแนนสูงก็ควรที่จะทำข้อสอบข้อหนึ่ง ๆ นั้นได้ถูกเป็นส่วนมากและคนที่ทำข้อสอบชุดนี้ได้คะแนนน้อยก็ควรทำข้อสอบข้อนั้น ๆ ผิดด้วย ถ้าเป็นเช่นนี้ค่าไบซีเรียลสูงหรือต่ำกับคะแนนรวม ก็บอกถึงความสามารถของชุดของข้อสอบเหล่านี้ว่ามีค่าอำนาจจำแนกคนเก่งคนอ่อนได้มากน้อยแค่ไหน และเมื่อนำคุณสมบัติประจำข้อเหล่านี้ไปสัมพันธ์กับน้ำหนักของคุณลักษณะเด่นที่แบบสอบนั้นวัด จึงดูเหมือนว่ามันมีเหตุผลดีที่จะบอกว่านอกจากข้อสอบนั้นมีคุณสมบัติของความคงที่ภายในสูง มีความสามารถจำแนกคนเก่งและคนอ่อนได้ดีแล้ว ยังยืนยันได้อีก มีความสัมพันธ์กับคุณลักษณะเด่นที่ต้องการวัดอีกด้วยและได้ตั้งเกณฑ์ของความสัมพันธ์ในขั้นสุดท้ายว่าควรจะมีค่าไม่ต่ำกว่า .8

แต่จุดอ่อนที่ผู้พัฒนาวิธีการตรวจสอบแบบนี้ อาจจะมองข้ามไปก็คือว่า ถ้าข้อใดมีค่าไบซีเรียลต่ำ สอดคล้องกันกับค่าน้ำหนักของตัวประกอบตัวแรกซึ่งตรงกับข้อนั้นมีค่าต่ำด้วย ก็ยังคงส่งผลให้ค่าสหสัมพันธ์ในขั้นสุดท้ายนี้ มีค่าสูงไปด้วย (มากกว่า .8) ซึ่งก็ไม่น่าที่จะตีความได้ว่าแบบสอบชุดนี้มีความเป็นเอกมิติ เพราะค่าน้ำหนักของข้อสอบบนตัวประกอบตัวแรก หรือที่เรียกว่าคุณลักษณะเด่น มีค่าต่ำแล้วแบบสอบชุดนี้ก็ไม่เหมาะที่จะบ่งบอกถึง การมีความเป็นเอกมิติของแบบสอบอยู่ อีกประการหนึ่ง การที่ข้อสอบมีสหสัมพันธ์ไบซีเรียลกับคะแนนรวมของแบบสอบทั้งฉบับมีค่าสูง ก็ไม่ได้รับประกันว่าข้อสอบข้อต่าง ๆ ที่มีค่าไบซีเรียลสูงเหล่านั้นจะมีความตรงเกณฑ์เป้าหมายของคุณลักษณะที่ต้องการจะจัด

แนวคิดในการแก้ไขจุดอ่อนของวิธีการนี้น่าจะนำเอาแนวคิดเกี่ยวกับการวิเคราะห์ของ Gulliksen มาใช้ซึ่งจำเป็นต้องอาศัยตัวแปรภายนอกมาช่วยในการบ่งบอกถึงความตรงของข้อสอบว่าได้มุ่งวัดในสิ่งที่ต้องการวัดหรือไม่ ซึ่งในบางครั้งสิ่งที่เราต้องการจะมุ่งวัดอาจมีมากกว่า 1 คุณลักษณะ และเมื่อเป็นเช่นนี้ค่าความสัมพันธ์ระหว่างค่าความตรงตามแนวของ Gulliksen นี้กับน้ำหนักของข้อสอบบนตัวประกอบที่เกิดจากแบบสอบที่เราต้องการจะวิเคราะห์ เป็นตัวบอกถึงความเป็นเอกมิติของแบบสอบ ดังนั้นถ้าน้ำหนักของข้อสอบบนตัวประกอบตัวแรกมีค่าสูง เมื่อคิดเปอร์เซ็นต์ของความแปรปรวนแล้ว มากกว่า 50 เปอร์เซ็นต์ เมื่อเทียบกับความแปรปรวนของตัวประกอบอื่น ๆ เราก็สามารถสรุปได้ว่า สิ่งที่ข้อสอบส่วนใหญ่มุ่งวัดอยู่บนตัวประกอบที่ 1 แต่ก็ไม่

อาจบอกได้ว่าตัวประกอบที่ 1 ที่ว่ามีความตรงตามเกณฑ์ที่เราต้องการมุ่งวัดหรือไม่ จึงจำเป็นต้องตรวจเช็คด้วยโดยการหาตัวแปรภายนอกที่มีความสัมพันธ์กับสิ่งที่เรามุ่งวัดมาเป็นเกณฑ์ในการตรวจสอบค่าความตรงประจำข้อ ซึ่งตัวแปรหรือแบบสอบเกณฑ์ภายนอกนี้ ไม่จำเป็นจะต้องมีความเป็นเอกมิติ แต่หาว่ามีบางสิ่งบางอย่างที่เราต้องการวัดสัมพันธ์อยู่ จึงอาจให้ค่าความตรงประจำข้อสูงได้ แต่ในการหาค่าน้ำหนักของข้อบนตัวประกอบแล้ว ไม่มีความเกี่ยวข้องกับความตรง เพียงแต่จะตรวจสอบว่าแบบสอบชุดนี้มีความเป็นเอกมิติหรือไม่เท่านั้น ส่วนการตรวจสอบความตรงประจำข้อนั้นก็เพื่อยืนยันว่าเอกมิติที่มุ่งวัดนั้นมีความสัมพันธ์ความตรงตามเกณฑ์ที่ต้องการของแต่ละข้อ ดังนั้นค่าสหสัมพันธ์ในขั้นสุดท้ายนี้จะมีค่าสูง เมื่อข้อสอบมีค่าความตรงสูง-ต่ำ สอดคล้องกับค่าน้ำหนักของข้อนั้น ๆ กับตัวประกอบตัวแรก เพราะข้อสอบที่มีความบ่งชี้ความเป็นเอกมิติของแบบสอบที่ประยุกต์จากวิธีการทดสอบไบซีเรียลเดิมเป็นดังนี้

1. หาค่าสหสัมพันธ์แบบไบซีเรียลกับข้อสอบที่ต้องการวิเคราะห์แต่ละข้อ ค่าสหสัมพันธ์ที่ได้นี้ใช้สัญลักษณ์ว่า r_{ixc} คือ ค่าสหสัมพันธ์ไบซีเรียลระหว่างข้อที่ i กับคะแนนรวมจากแบบสอบเกณฑ์ ซึ่งจะเป็นค่าที่มีความหมายมากกว่า

2. หาความหมายในแง่ของความตรงประจำข้อ (Item Validity) ด้วยการนำ ค่า r_{ixc} คูณกับค่าส่วนเบี่ยงเบนมาตราฐานของข้อ i (s_i) นั่นก็คือ $r_{ixc}\ s_i$ เป็นค่าดัชนีความตรงประจำข้อตามแนวคิดของ Gulikson (Gulliksen,1967:384-388)

3. และเมื่อนำค่า $r_{pbxc}\ s_i$ ของ Gulliksen มาหาค่าสหสัมพันธ์กับค่าน้ำหนักของข้อบนตัวประกอบร่วมตัวแรก ก็จะได้ค่าดัชนีบ่งชี้ความเป็นเอกมิติของแบบสอบอีกตัวหนึ่งที่ประยุกต์มาจาก BT (Biserial Test) เรียกว่า ABT (The Application of Gulliksen Item Validity Index and the Biserial Test Index) ซึ่งดัชนีที่ประยุกต์มาใหม่ 2 ตัวนี้ จากวิธีการเดิม BT จะบอกความสัมพันธ์กับตัวประกอบร่วมตัวแรกและยังบ่งบอกถึงความตรง ในการวัดคุณลักษณะที่ต้องการวัดของแบบสอบอีกด้วย

2.3 การประยุกต์ดัชนีบ่งชี้ความเป็นเอกมิติของแบบสอบที่ใช้ค่าดัชนีความเป็นเอกพันธ์ของ Green

วิธีการตรวจสอบความเป็นเอกมิติของแบบสอบ โดยใช้ดัชนีความเอกพันธ์ของ Green นั้น มีที่มาจากการเริ่มต้นด้วยมโนทัศน์ที่เกี่ยวกับการได้มา ซึ่งเมตริกซ์ที่มีค่าแรงค์เป็นหนึ่ง โดยมีหลักของเหตุผลที่ว่า ถ้าชุดของข้อสอบมีตัวประกอบร่วมเดียวแล้วน้ำหนักของข้อสอบข้อหนึ่งๆ บนตัวประกอบนั้น จะมีค่าเท่ากับรากที่ 2 ของค่าคอมมูนาลิตีของข้อนั้นๆ นั่นคือ $a^2_i = h^2_i$,

$a^2_{ij} = h^2_{ij}$ และเช่นเดียวกันค่าสหสัมพันธ์ระหว่างข้อสอบคู่หนึ่งๆ ก็จะมีค่าเท่ากับผลคูณของค่าน้ำหนักของข้อทั้งสองนั้น ($r_{ij} = a_i a_j$) ดังนั้น $|r_{ij}| = \sqrt{h_i^2 h_j^2}$ ผลของการถอดรูททำให้ค่า r_{ij} มีค่าเป็นบวกเสมอ ดังนั้น ที่มาของสูตรของ Green จึงเป็น

$|r_{ij}| < \sqrt{h_i^2 h_j^2}$ และ

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

ความแม่นยำของการประมาณค่า อยู่ที่การประมาณค่าคอมมูนาลิตี้ (h^2_{ij}) และเขาก็ได้เสนอให้ใช้ค่ากำลังสองพหุคูณระหว่างข้อสอบอื่นๆ กับข้อนั้นๆ (SMC_j) เพราะเขาว่าการประมาณค่า h^2_{ij} ด้วยวิธีนี้ก็เพื่อทำนายข้อสอบข้อหนึ่งๆ ในแบบสอบจากข้อสอบข้ออื่นๆ และสามารถคำนวณได้ไม่ยากจากโปรแกรมคอมพิวเตอร์ โดย $R^2_j = 1 - s^2_j$ เมื่อ R^2_j คือ SMC_j และ $s^2 j$ ก็คือ ส่วนกลับของสมาชิกในแนวทะแยงตัวที่ j ของอินเวอร์สเมตริกซ์ของเมตริกซ์สหสัมพันธ์ระหว่างข้อสอบ แต่เนื่องจากค่า R^2_j เป็นตัวประมาณค่า h^2_{ij} ที่ขีดจำกัดต่ำ (lower bound estimator) ดังนั้น ตัวดัชนีความเป็นเอกพันธ์ที่คำนวณได้จากสูตรของ Green จึงอาจจะให้ค่าที่เกินกว่า 1.00 ได้ ทั้งที่ความจริงแล้ว

$|r_{ij}| \leq \sqrt{h_i^2 h_j^2}$ และ

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

ดังที่กล่าวมาแล้ว ซึ่งค่าดัชนีความเป็นเอกพันธ์ ไม่ควรจะมีค่าเกินกว่า 1.00

จุดบกพร่องนี้ Green ได้เสนอค่าแก้ r_{ij} ที่ไม่ขึ้นอยู่กับค่าคอมมูนาลิตี้ โดยหารค่า r_{ij} ด้วย $h^2_i h^2_j$ เหมือนอย่างที่เราทำแอทเทนนูเอชั่น (Attenuation) กับค่าสหสัมพันธ์อื่น ๆ แล้วนำผลเฉลี่ยของ r_{ij} มาเป็นดัชนีความเป็นเอกพันธ์ที่จะให้ค่าอยู่ระหว่าง 0.00 ถึง 1.00 อย่างไรก็ดี Warm (Warm, 1978:106) ได้พัฒนาสูตรของ Green ใหม่ โดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างข้อที่ได้มาจากสูตรสหสัมพันธ์เตตราคลอริกและเมื่อการประมาณค่า h^2_{ij} ด้วย R^2_i นั้น ทำให้ได้ค่าดัชนีความเป็นเอกพันธ์เกิน 1.00 ดังที่กล่าวมาแล้ว Warm จึงได้ใช้ค่าแก้ตามวิธีการที่ Green แนะนำไว้ข้างต้น ดังนั้น สูตรของ Warm จึงเปลี่ยนจากดัชนี μ เป็นดัชนี G ดังนี้

$$G = \frac{\sum\sum_{i \neq j} \frac{r_{i,j}}{\sqrt{h^2_i \, h^2_j}}}{n} \quad , \quad i > j$$

อย่างไรก็ดี Warm ได้มีการนำวิธีการในการตรวจสอบเอกมิติวิธีต่างๆ รวมทั้งดัชนี G นี้ มาเปรียบเทียบกันแล้ว ได้ให้ข้อสรุปว่า (Warm,1978:107) ค่าของ G นี้ไม่ได้ เข้าสู่ 1.00เมื่อข้อมูลเป็นเอกมิติ และไม่ได้สู่ 0.00 เมื่อข้อมูลไม่เป็นเอกมิติ และยิ่งไปกว่านั้น Warm ยังรายงานอีกว่าดัชนี G เป็นดัชนีบ่งชี้ความเป็นเอกมิติที่ไม่ดีไปกว่าการใช้ค่าเปอร์เซนต์ของความแปรปรวนร่วมของตัวประกอบตัวแรก (first factor percent of common variance)

จะเห็นได้ว่า สูตรของ Green เป็นที่น่าสนใจ ได้ถูกนำมาพัฒนาแล้วก็ยังได้ผลไม่เป็นที่พอใจ อีกประการหนึ่งการเปรียบเทียบวิธีการบ่งชี้ความเป็นเอกมิติวิธีต่าง ๆ ของ Warm ได้ใช้ข้อมูลจริงจากแบบสอบ 6 ชุด และข้อมูลเชิงสุ่ม 6 ชุดซึ่งเป็นข้อมูลของ Mc Bride & Weiss ที่ได้ทำการศึกษาและพบว่า ข้อมูลจากแบบทดสอบจริงได้รับการตรวจสอบจากวิธีการ 3 วิธี คือ วิธีการที่ใช้ข้อมูลเชิงสุ่มเป็นฐาน (Random Baseline),วิธีที่ใช้สหสัมพันธ์ไปซีเรียล(Biserial Test) และวิธีการที่ใช้น้ำหนักตัวประกอบ (Factor Loading Test) บ่งชี้ว่าเป็นข้อสอบที่มีความเป็นเอกมิติ แต่สำหรับค่า G ของ Warm ซึ่งใช้ข้อมูลชุดเดียวกันพบว่าให้ค่า G ตั้งแต่ .419 ถึง .484 และมีความสัมพันธ์กับเปอร์เซนต์ของตัวประกอบร่วมตัวแรก (first common factor variance) โดยมีค่า ρ = 1.00 สำหรับข้อมูลจากแบบสอบจริง และให้ค่า G .284 ถึง .348 ซึ่งเป็นค่าที่มีความสัมพันธ์กับเปอร์เซนต์ของตัวประกอบร่วมตัวแรกด้วยค่า ρ = .60 Warm จึงสรุปว่า ค่า G ไม่ได้เข้าสู่ 1.00 เมื่อข้อสอบเป็นเอกมิติและไม่เข้าสู่ 0.00 เมื่อข้อสอบไม่เป็นเอกมิติยิ่งไปกว่านั้นยังสรุปอีกว่า G ไม่ได้เป็นดัชนีบ่งชี้ที่ดีไปกว่าค่าเปอร์เซนต์ของตัวประกอบร่วมตัวแรก จากข้อรายงานดังกล่าวนี้ ผู้วิจัยยังไม่แน่ใจในความไม่แน่นอนของค่า G อาจเป็นเพราะวิธีการตรวจสอบของ Warm ยังไม่ได้แสดงให้ผู้วิจัยได้เห็นเด่นชัดถึงความเป็นดัชนีที่ไม่ดี ดังนั้น ผู้วิจัยจึงต้องการที่จะปรับปรุงและพัฒนาสูตร μ ของ Green ใหม่ โดยขอยึดหลักการเดิมของ Green ที่ว่า $|r_{i,j}| \leq \sqrt{h^2_i \, h^2_j}$ และการประมาณค่าดัชนี ให้ได้ใกล้เคียงกับความเป็นจริงนั้นอยู่ที่การประมาณค่า h^2_i

ในการประมาณค่า h^2_i ได้ใกล้เคียงเพียงใดนั้นตรวจสอบได้หลังจากการสะกัดตัวประกอบแล้วและ ให้แฟกเตอร์เมตริก (F) ที่คูณกับค่าทรานสโปสของมันแล้ว จะให้ค่าสหสัมพันธ์ย้อนกลับ (R-reproduce) ที่แตกต่างจากค่าสหสัมพันธ์เริ่มต้น (R-origin) น้อยที่สุด ซึ่งวิธีการ

นี้เป็นหลักการที่โปรแกรมคอมพิวเตอร์นำมาใช้ในการสะกัดตัวประกอบ และภายหลังการสะกัดตัวประกอบด้วยวิธีใดก็ตาม ก็จะมีการรายงานค่าคอมมูนาลิตี้ในส่วนที่เป็นสถิติขั้นสุด (Final Statistic) ของผลการประมวล (out put) เสมอ และเมื่อ Green เห็นว่าค่าคอมมูนาลิตี้ควรประมาณครั้งแรกจากค่า R^2_i ดังนั้น ผู้วิจัยจึงของเสนอแนะให้ใช้ค่าคอมมูนาลิตี้ ภายหลังการ

สะกัดตัวประกอบด้วยวิธิอิมเมต ซึ่งเป็นวิธีการที่เหมาะสมคือ วิธีนี้จะเป็นวิธีการประมาณค่าคอมมูนาลิตี้ได้ใกล้เคียงกับความเป็นจริงมากที่สุด และสมเหตุสมผลที่ว่า ผู้วิจัยต้องการค่าคอมมูนาลิตี้ในความหมายที่เป็นความสัมพันธ์ระหว่างข้อสอบข้อหนึ่ง ๆ กับข้อสอบที่เหลือทั้งหมดในเมื่อค่าความสัมพันธ์ระหว่างข้อสอบ 2 ข้อเท่ากับรากที่ 2 ของผลคูณของค่าความสัมพันธ์ระหว่างข้อสอบข้อหนึ่ง (ใน 2 ข้อนั้น) สัมพันธ์ข้ออื่นๆ ที่เหลือกับความสัมพันธ์ระหว่างข้อสอบอีกข้อหนึ่งที่เหลือกับข้ออื่นๆ ที่เหลือก็หมายความว่า ค่าสหสัมพันธ์ระหว่าง 2 ข้อใด ๆ คู่หนึ่ง มีค่าเท่ากับความสัมพันธ์ของข้อสอบข้อเหล่านั้นกับข้อสอบอื่น ๆ ที่เหลือ

ซึ่งก็แสดงว่า เมื่อผลรวมของคู่ของความสัมพันธ์ระหว่างข้อแต่ละคู่รวมกันแล้ว ไปมีค่าเท่ากับผลรวมของความสัมพันธ์ระหว่างข้อหนึ่ง ๆ กับข้อที่เหลือ ก็เท่ากับว่าข้อสอบเหล่านั้นมีความสัมพันธ์กันในเรื่องเดียวกัน ไม่ว่าจะคิดจากผลรวมของความสัมพันธ์แต่ละคู่ หรือคิดจากผลรวมของความสัมพันธ์ระหว่าง แต่ละข้อกับข้ออื่น ๆ ที่เหลือ ก็เป็นเรื่องเดียวกันหมดจึงให้ผลรวมของความสัมพันธ์เท่ากัน ดังนั้นผู้วิจัยจึงเสนอดัชนีความเป็นเอกพันธ์ของแบบสอบพร้อมตัวแก้ $r_{i,j}$ ให้อิสระจาก h^2_i h^2_j ตามข้อเสนอแนะของ Green ดังนี้

$$AG = \frac{\sum_{i \neq j} \sum \sqrt{\frac{|r_{i,j}|}{(a^2_{ik})(a^2_{jk})}}}{n} \quad ; \quad \begin{matrix} i = 1, \ldots, P-1 \\ j = 2, \ldots, P \end{matrix}$$

เมื่อ AG คือ ดัชนีบ่งชี้ความเป็นเอกมิติของแบบสอบที่ประยุกต์มาจากของ Green

$r_{i,j}$ คือ ค่าสัมประสิทธิ์สหสัมพันธ์ แบบเตตราคลอริก ระหว่างข้อสอบ

n คือ จำนวนค่าสหสัมพันธ์มีค่าเท่ากับ P (P - 1)/2

p คือ จำนวนข้อ

a_i, a_j คือ ค่าน้ำหนัก ของข้อ i และ j บนทุกตัวประกอบ

3. <u>การใช้โปรแกรม LISREL และโมเดล MIMIC ในการตรวจสอบความเป็นเอก-มิติของแบบสอบ</u>

LISREL เป็นโปรแกรมสำเร็จที่ใช้ประโยชน์ได้อเนกประสงค์ เกี่ยวกับการวิเคราะห์โมเดลเชิงสาเหตุ (Causal model) โดยมีวัตถุประสงค์เพื่อศึกษาความสัมพันธ์เชิงโครงสร้าง (Structural relations) ระหว่างตัวประกอบหรือตัวแปรแฝงของทั้งตัวแปรต้นและตัวแปรตามที่เขียนขึ้นมาเป็นโมเดล LISREL ไม่ใช่โมเดล แต่เป็นโปรแกรมคอมพิวเตอร์ที่มีวิธีการวิเคราะห์โมเดลรูปต่าง ๆ LISREL เป็นชื่อที่ JORESKOG ตั้งขึ้น (Pedharzur, 1982: P 637) มีชื่อเต็มว่า Linear Structural Relationship ใช้วิเคราะห์ความสัมพันธ์ของตัวแปรในรูปโมเดลสมการโครงสร้างด้วยวิธีการทางคอมพิวเตอร์ ซึ่งมีวิธีการประมาณค่าพารามิเตอร์ต่าง ๆ โดยยึดพื้นฐานของทฤษฎีทางสถิติ ที่เรียกว่า แมกซิมั่มไลลิฮูด (Maximum - likelihood) LISREL สามารถวิเคราะห์โมเดลต่าง ๆ ดังต่อไปนี้

1. วิเคราะห์โมเดลที่มีทิศทางของสาเหตุไปในทิศทางเดียว(Recursive model) เช่น การวิเคราะห์โมเดลเส้นทาง (Path - analysis model)
2. วิเคราะห์โมเดลที่มีค่าที่เหลือ (Residual) ที่สัมพันธ์กันเอง
3. วิเคราะห์โมเดลที่มีทิศทางของสาเหตุสองทิศทาง (Non Recursive model)
4. วิเคราะห์โมเดลที่มีดัชนีของตัวแปรที่วัดไม่ได้ หรือตัวแปรแฝงหลายตัว(Pedharzur,1982:638)

ด้วยเหตุที่ LISREL มีประโยชน์ใช้ได้ในเอนกประสงค์ โปรแกรมของมันจึงค่อนข้างซับซ้อน และมีคำสั่งมากมาย ในที่นี้จะขอกล่าวถึงการทำงานของ LISREL โดยย่อและการใช้ LISREL ในการวิเคราะห์ข้อสอบ

โมเดลที่ใช้โปรแกรม LISREL ในการวิเคราะห์ จะประกอบด้วย 2 ส่วน คือ ส่วนที่แสดงความสัมพันธ์ระหว่างตัวแปรต้นแฝง (exogenous) และตัวแปรตามแฝง(Endogenous) กับส่วนที่แสดงการวัดของตัวแปรแฝงเหล่านั้น ดังนั้น LISREL จึงประกอบด้วยส่วนสำคัญย่อย 2 ส่วน (major subdivisions) คือ โมเดลสมการโครงสร้าง (Structural equation model) และโมเดลการวัด (measurement model)

1. โมเดลสมการโครงสร้าง แสดงความสัมพันธ์ระหว่างตัวแปรต้นแฝง และตัวแปรตามแฝง ดังนี้

$$\eta = B\eta + \Gamma\xi + \zeta \qquad (1)$$

เมื่อ η คือ เวคเตอร์ m x 1 ของตัวแปรตามแฝง

ξ คือ เวคเตอร์ m x 1 ของตัวแปรต้นแฝง

B คือ เมตริกซ์ m x m ของสัมประสิทธิของผลกระทบระหว่างตัวแปรตามแฝง กับตัวแปรตามแฝง

Γ คือ เมตริกซ์ m x n ของเมตริกซ์ของสัมประสิทธิของผลกระทบระหว่างตัวแปรต้นแฝง ที่มีต่อตัวแปรตามแฝง

ζ คือ เวคเตอร์ m x 1 ของค่าที่เหลือ หรือ ความคลาดเคลื่อนของสมการ

ข้อตกลงของสมการ ก็คือ ค่าเฉลี่ยของตัวแปรทุกตัวในสมการ มีค่าเป็นศูนย์ นั่นคือ ตัวแปรทุกตัวแสดงอยู่ในรูปของคะแนนความแตกต่าง (diviate score) และยังมีข้อตกลงว่า ξ และ ζ ไม่สัมพันธ์กัน และ B ต้องหาค่าดีเทอร์มีแนนได้ (non singular matrix)

2. โมเดลการวัด แสดงความสัมพันธ์ระหว่างตัวแปรแฝงซึ่งวัดไม่ได้กับตัวแปรสังเกตที่วัดได้ ดังนี้

ในส่วนของตัวแปรตามแฝง สมการจะเป็นดังนี้

$$Y = \Lambda_y\eta + \epsilon \qquad (2)$$

และในส่วนของตัวแปรต้นแฝง สมการจะเป็นดังนี้

$$X = \Lambda_x\xi + \delta \qquad (3)$$

โดยที่ y และ X เป็นเวคเตอร์ของการวัดตัวแปรตาม และตัวแปรต้น มีขนาดเป็น P x 1 และ q x 1 ตามลำดับ

Λ_y และ Λ_x คือ เมตริกของน้ำหนัก (Loading) ของตัวแปรที่วัดได้ y และ X บนตัวแปรตามแฝงและตัวแปรต้นแฝง มีขนาดเป็น p x m และ q x n ตามลำดับ

ส่วน Θ_y และ Θ_x เป็นเวคเตอร์ที่แสดงความคลาดเคลื่อนจากการวัดค่า y และ X มีขนาดของเวคเตอร์เป็น p x 1 และ q x 1 ตามลำดับ

Pedhurzer (1982:42) ได้แสดงให้เห็นว่าพื้นฐานของข้อตกลงของ LISREL ก็คือ เมตริกของความแปรปรวนรวมของตัวแปรที่วัดได้ (Σ) นั่นเอง ซึ่งเป็นฟังชั่นของเมตริกซ์ 8 เมตริกซ์ดังนี้

1. Λ_y เป็นเมตริกซ์ของสัมประสิทธิ์หรือน้ำหนัก ซึ่งเกี่ยวข้องกับดัชนี (indicator) ของตัวแปรตาม กับตัวแปรตามแฝง (η)
2. Λ_x เป็นเมตริกซ์ของสัมประสิทธิ์หรือน้ำหนัก ซึ่งเกี่ยวข้องกับดัชนี (indicator) ของตัวแปรต้น กับตัวแปรต้นแฝง (ξ)
3. β (เบต้า) เป็นเมตริกของสัมประสิทธิ์ของผลกระทบจากตัวแปรตามแฝง ที่มีต่อตัวแปรตามแฝง
4. γ (แกรมมา) เป็นเมตริกของสัมประสิทธิ์ของผลกระทบจากตัวแปรต้นแฝงที่มีต่อตัวแปรตามแฝง
5. ϕ (ฟี) เป็นเมตริกซ์ของความแปรปรวนร่วมของตัวแปรต้นแฝง (ξ)
6. Ψ (ซาย) เป็นเมตริกซ์ของความแปรปรวนร่วมของค่าที่เหลือ (δ)
7. Θ_y (ทีต้า) เป็นเมตริกซ์ของความแปรปรวนร่วมของความคลาดเคลื่อน จากการวัดค่า y
8. Θ_x (ทีต้า) เป็นเมตริกซ์ของความแปรปรวนร่วมของความคลาดเคลื่อน จากการวัดค่า x

สมาชิกที่อยู่ในเมตริกซ์ข้างต้นนี้ อาจเป็นได้ 3 ชนิด คือ

1. พารามิเตอร์กำหนด (fixed parameter) เป็นเมตริกซ์ของพารามิเตอร์ที่ถูกกำหนดค่าขึ้น

2. พารามิเตอร์กึ่งกำหนด (Constrained parameters) เป็นเมตริกซ์ที่ไม่รู้ค่าพารามิเตอร์ แต่มีค่าเท่ากับ 1 หรือ มากกว่าพารามิเตอร์อื่น ๆ

3. พารามิเตอร์ที่ต้องการหาค่า (free parameter) เป็นเมตริกซ์ที่ไม่รู้ค่าพารามิเตอร์และไม่อาจกำหนดให้มีค่าเท่ากับพารามิเตอร์อื่น ๆ ได้ (Joreskog & Sorbom, 1978 อ้างใน Pedhurzer, 1982 : 642)

Joreskog ใช้เมตริกของความแปรปรวนร่วมของตัวแปร สังเกตที่วัดได้มาประมาณสมาชิกในเมตริกทั้ง 8 ชนิดข้างต้นนี้

ในการตรวจสอบเอกมิติของข้อมูล ได้มีนักวิจัยหลายท่านได้นำเอาส่วนที่เป็นโมเดลการวัดมาใช้ในการตรวจสอบ โดยให้ข้อกระทงที่วัดได้เป็นตัวแปร y และคุณลักษณะที่ต้องการวัดเป็นตัวแปรแฝง (η) ซึ่งมีเพียงตัวเดียว ค่า Λ_y จึงเป็นสัมประสิทธิ์หรือน้ำหนักของข้อกระทงที่วัดได้บนตัวแปรคุณลักษณะแฝง (η) และความคลาดเคลื่อนจากการวัด ก็คือ นั่นคือการวิเคราะห์องค์ประกอบแบบยืนยันนั่นเอง (Confirmatory factor analysis) ซึ่งมีสมการหรือโมเดลการวัดเป็น

$$Y = \Lambda_y \eta + \epsilon$$

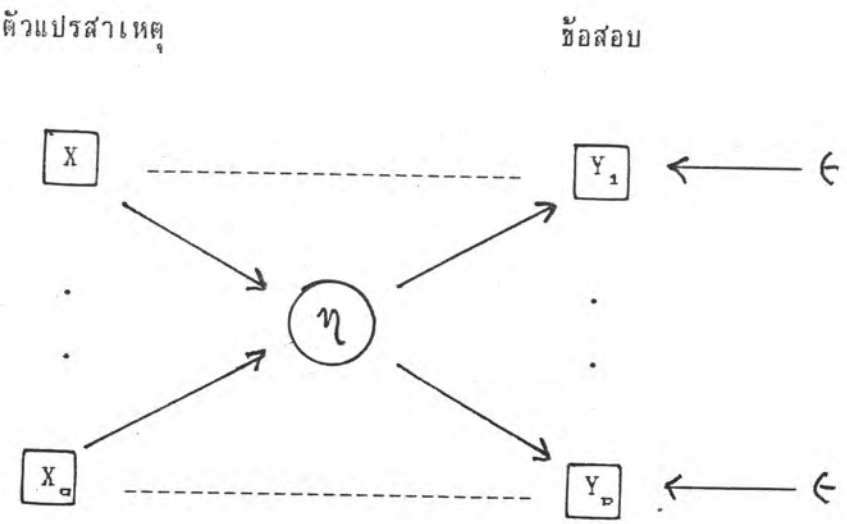
Muthen (Muthen อ้างใน Wainer & Brun, 1985: 213) ได้นำโมเดลสมการโครงสร้างมาใช้ในการศึกษาความตรง โดยการโยงเอา IRT ไปเชื่อมต่อกับตัวแปรภายนอก ซึ่งตัวแปรภายนอกนี้อาจเป็นตัวแปรประเภทตัวแปรกลุ่มหรือตัวแปรต่อเนื่องก็ได้ สูตรโมเดลของ Muthen จึงเป็นโมเดลที่แสดงความสัมพันธ์ระหว่างตัวแปรภายนอกกับข้อสอบที่ตอบได้ (Response items) เขากล่าวว่าถ้าข้อมูลมีคุณภาพและมีเพียงพอโมเดลนี้สามารถให้สารสนเทศและให้อำนาจในการวิเคราะห์ภาวะสันนิษฐาน (Construct) ของการวัดได้มากกว่าการใช้ IRT มาตราฐานอย่างเดียว ลักษณะทั่วไปของโมเดลใหม่นี้ก็คือสามารถอธิบายประเด็นสำคัญ ๆ พร้อม ๆ กันทีเดียว 4 ประเด็นเกี่ยวกับการวิเคราะห์ข้อสอบ คือ (Muthen อ้างใน Wainer & Brun, 1985 : 213)

1. ประมาณค่าพารามิเตอร์ของข้อสอบตามแบบของ IRT
2. ประเมินความแข็งแกร่งของคุณลักษณะแฝงเดิมของนักเรียนที่ถูกสมมุติขึ้นมา เพื่อการวิเคราะห์
3. ขจัดข้อสอบที่ลำเอียง
4. ใช้ทดสอบและการผ่อนปรน ข้อตกลงของ IRT เกี่ยวกับเอกมิติและเงื่อนไขความเป็นอิสระ

โมเดลที่ Muthén ใช้ในการวิเคราะห์เป็นโมเดลหนึ่งในโปรแกรม LISREL ที่มีชื่อว่า MIMIC การวิจัยนี้ได้อาศัยโมเดล MIMIC ในการวิเคราะห์หาระดับต่าง ๆ ของความเป็นเอกมิติของแบบสอบฉบับต่างๆ และยึดเอาเทคนิควิธีนี้เป็นเกณฑ์ในการตรวจสอบความเป็นเอกมิติ ดังนั้นในที่นี้จึงขอทบทวนโมเดลไมมิคในหัวข้อต่อไป

โมเดลไมมิค (MIMIC model)

MIMIC เป็นคำที่ย่อมาจาก Multiple Indicators and Multiple Causes ซึ่งหมายถึง โมเดลที่ประกอบด้วยดัชนีบ่งชี้หลายตัวที่บอกคุณลักษณะของตัวแปรแฝง และมีตัวแปรเหตุที่วัดได้หลาย ๆ ตัว ดังนี้



รูปที่ 2.3 โมเดลทั่วไปของไมมิค ดัดแปลงจาก A MIMIC structural profit model ของ Muthén

รูปแบบอย่างง่ายของโมเดลไมมิค ประกอบด้วยตัวแปรแฝงที่วัดไม่ได้เพียวตัวเดียว ซึ่งเป็นผลมาจากตัวแปรเหตุ x ที่วัดได้หลาย ๆ ตัว และยังแสดงออกมาเป็นตัวแปร y ที่วัดได้หลาย ๆ ตัวอีกด้วย มีสมการโมเดล ดังนี้

$$Y = \lambda\eta + \epsilon \qquad (4)$$

$$= \gamma' x + \varsigma \qquad (5)$$

เมื่อ $Y = (y_1 , y_2 \ldots\ldots\ldots y_p)$ เป็นดัชนีบ่งชี้ของตัวแปรแฝง

$X = (x_1 , x_2 \ldots\ldots\ldots x_q)$ เป็นตัวแปรสาเหตุของตัวแปรแฝง

เมื่อเปรียบเทียบกับโมเดล LISREL เต็มรูปนั้น สมการ (4) ก็คือ โมเดลการวัดของตัวแปร η และสมการ (5) ก็เป็นสมการโครงสร้างของตัวแปร η ส่วน ϵ และ ξ มีข้อตกลงเบื้องต้นว่า ไม่สัมพันธ์กันอย่างอิสระจากกัน (mutually uncorrelated) ในสมการ (1) พูดได้ว่าตัว y ทั้งหลายเป็นตัวแทนของการวัดตัว η (congeneric measures of η) และสมการ (5) พูดได้ว่าตัว η เป็นความสัมพันธ์เชิงเส้น(linear) กับค่า x ทั้งหลาย ซึ่งอยู่ในรูปของการแจกแจงแบบสุ่ม โมเดลนี้ยังสามารถให้ภาพของโมเดลถดถอยพหุที่มีกฎบังคับเฉพาะ 2 ข้อคือ

1. เมตริกซ์ถดถอยต้องมีแรงค์เป็น 1
2. เมตริกซ์ของความแปรปรวนร่วมของที่เหลือต้องสอดคล้อง(satisfy) กับโมเดลการวัดที่เป็นตัวแทน (Congeneric measurement model)

ข้อพิสูจน์นี้ สามารถแสดงให้เห็นได้ด้วยการแทนค่าสมการ (4) ด้วยสมการ (5) ดังนี้

$$Y = \Lambda\gamma' x + \Lambda\xi + \epsilon$$

$$= \Pi x + Z$$

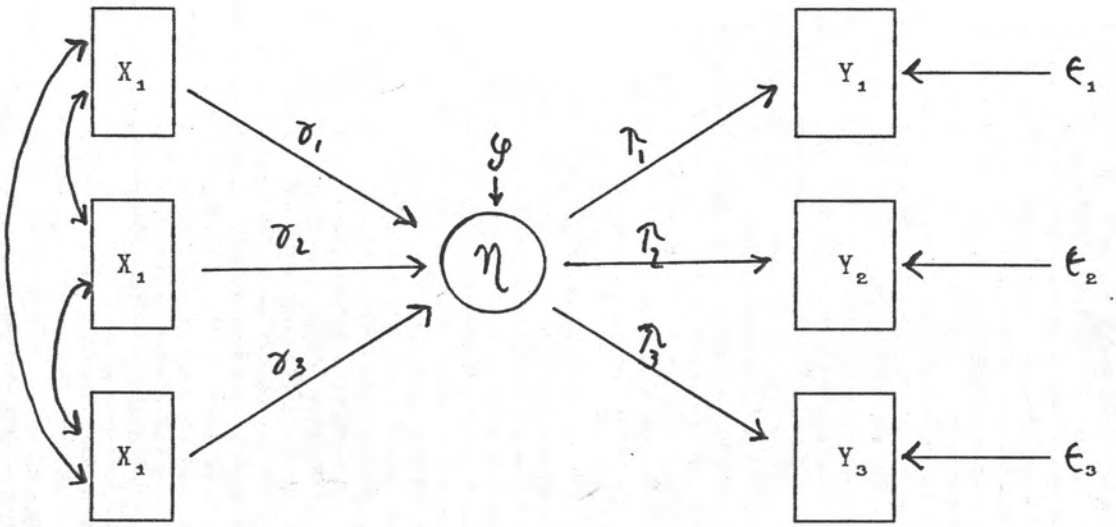
จะเห็นได้ว่า $\Pi = \Lambda\gamma'$ และ $Cov(Z) = \Lambda\beta'\Psi + \epsilon$ เมื่อ $\Psi = Var (\xi)$ และ Θ_ϵ เป็นแนวทะแยงของเมตริกซ์ ความแปรปรวนร่วมของ

เพื่อประมาณโมเดลไมมิคด้วย LISREL ทำได้โดยตรง เมื่อสมการ (4) และ(5) ก็คือ สมการ (1) และ (2) ตามลำดับ ส่วนสมการ (3) ในที่นี้คือ $X \equiv \xi$ นั่นคือ $\Lambda_x = I$ และ $\Theta_\delta = 0$ ต่อจากนั้นก็ใช้ตัวเลือก (option) F I ซึ่งเท่ากับว่าให้ $\Phi = Cov (\xi)$ ซึ่งจะประมาณได้ โดย S_{xx} อันเป็นเมตริกซ์ ความแปรปรวนร่วมของ X ของกลุ่มตัวอย่าง

ตัวอย่างปัญหาและโปรแกรม ในการวิเคราะห์โมเดลไมมิค กรณีของ X = 3 และ Y = 3 ได้มาจาก Hodge & treiman (1968) (อ้างในหนังสือ DOS - LISREL and DOS - PRELIS) เรื่องสถานภาพทางสังคมและการมีส่วนร่วมทางสังคม โจทย์ปัญหามีว่า ในการศึกษาความสัมพันธ์ระหว่างสภานภาพทางสังคม และการมีส่วนร่วมทางสังคม ได้ใช้กลุ่มตัวอย่างผู้หญิงจำนวน 550 คน มีการวัดตัวแปรสถานภาพทางสังคม 6 ตัว ชื่อสถานภาพทางสังคม 6 ตัวนี้ ปรากฏอยู่ในตารางเมตริกซ์ข้างล่างนี้ และโมเดลได้แสดงอยู่ในรูปที่ 2

ตารางที่ 2.1 แสดงค่าสหสัมพันธ์ของตัวแปรใน โมเดลไมมิค พิมพ์อยู่ในไฟล์ชื่อ EX54.COR

	X_1	X_2	X_3	Y_1	Y_2	Y_3
Income	1.000					
Occupation	.304	1.000				
Education	.305	.344	1.000			
Church attendance	.100	.156	.158	1.000		
Memberships	.284	.192	.324	.360	1.000	
Friends seen	.176	.136	.226	.210	.365	1.000



รูปที่ 2.4 แสดงแผนผังเส้นทางของโมเดลไมมิค

ตัวแปรทั้งหมดถูกแปลงให้เป็นคะแนนมาตราฐาน ตัวแปร y ทั้งหลายเป็นตัวแปรบ่งชี้คุณลักษณะแฝง ที่เป็นอิสระจากกัน (การมีส่วนร่วมทางสังคม) ซึ่งเกิดจากตัวแปรสาเหตุ คือ x ทั้งหลาย ไฟล์คำสั่ง LISREL ที่ใช้ในการวิเคราะห์ คือ

```
Social Status and Participation
DA NI = 6    NO = 530    MA = KM
LA
Income Occupation education Churchat Membersh Friends
KM FI = Ex54.COR
3E
4  5  6  1  2  3
MO NY = 3  NE = 1    NX = 3    FI    LY = FR
LE
Socialph
FI Ly (1)
VA 1 LY(1)
OU SE TV
```

สำหรับ Muthen เขาได้ทำการวิจัยโดยใช้ โดเดล MIMIC เต็มรูปแบบโดยการเชื่อมโปรแกรม LISREL โมเดล MIMIC กับโปรแกรม LOGIST วิเคราะห์ค่าพารามิเตอร์ต่างๆ ตามแนวทฤษฎี IRT ในที่นี้จะยกส่วนโมเดล MIMIC ที่เขาได้ทำการวิจัยในมหาวิทยาลัยแคลิฟอร์เนียร์ เขาได้ใช้ตัวแปรภายนอก (X) ซึ่งเป็นภูมิหลัง 18ตัว และใช้ข้อสอบ (Y) เพียง 8 ข้อ และเขาก็สรุปว่าการได้วิเคราะห์ด้วยโมเดล MIMIC ก่อนที่จะทำ IRT นั้น สามารถให้ค่าพารามิเตอร์ที่ถูกต้องตรงความเป็นจริงมากกว่า ใช้ IRT อย่างเดียว (Muthen อ้างใน Wainer & Braun, 1985 : 213 - 238)

4. <u>งานวิจัยเกี่ยวข้องกับการตรวจสอบเอกมิติของแบบสอบ</u>

งานวิจัยเกี่ยวข้อง กับการตรวจสอบเอกมิติของแบบสอบ ส่วนมากจะเน้นถึงผลของการ

ฝ่าฝืนข้อตกลงของการวิเคราะห์ข้อสอบแบบ IRT ข้อตกลงเบื้องต้นที่สำคัญของ IRT แม้จะมี 2 ข้อคือ แบบสอบจะต้องมีคุณสมบัติที่ทุกข้อวัดในสิ่งเดียวกันเพียงเรื่องเดียวและผลการตอบข้อสอบในแบบสอบนั้นจะต้องให้ผลที่เป็นอิสระจากกันหรือที่เรียกว่า อิสระจากตำแหน่งของข้อสอบ (item local independence)แต่ความจริงแล้วการตรวจสอบความเป็นเอกมิติของแบบสอบอย่างเดียวก็สามารถคลุมไปถึงข้อตกลงของความเป็นอิสระสถานที่ด้วย เนื่องจากว่าข้อตกลงสองข้อนี้เป็นสิ่งที่เท่าเทียมกัน (equivalent) การตรวจสอบความเป็นเอกมิติของแบบสอบจึงมีผู้สนใจศึกษาและตรวจสอบมากกว่าข้ออื่น ๆ และยิ่งในปัจจุบันนี้ มีคนสนใจนำทฤษฎีการวิเคราะห์ข้อสอบแบบ IRT มาใช้กันมาก การตรวจสอบความเป็นเอกมิติก็ยิ่งทำกันอย่างแพร่หลายด้วยวิธีการต่าง ๆ ส่วนมากจะเริ่มต้นที่การวิเคราะห์ตัวประกอบ และศึกษาจากผลที่ได้จากการวิเคราะห์ตัวประกอบนั่นเอง

ในปี 1986 Harison(1986) ได้ศึกษาถึงความคงทนของการประมาณค่าพารามิเตอร์ของIRT ที่มีต่อการฝ่าฝืนข้อตกลงเบื้องต้นเกี่ยวกับเอกมิติ ซึ่งผลการวิจัยได้พิสูจน์ให้เห็นว่า วิธีการประมาณด้วย LOGIST มีความแกร่งต่อการฝ่าฝืนข้อตกลงเกี่ยวกับเอกมิติมาก และความยาวของแบบสอบเพิ่มขึ้น LOGIST จะประมาณค่าคุณลักษณะทั่วไป (general trait)ได้มีประสิทธิภาพมากขึ้น นอกจากนี้จำนวนตัวประกอบร่วมมากขึ้น การประมาณค่าพารามิเตอร์จะมีประสิทธิภาพเพิ่มขึ้นด้วย ยกเว้น กรณีแบบสอบยาว 30 ข้อ และลักษณะการกระจายของข้อสอบเป็นยูนิฟอร์ม (uniform) ค่าเฉลี่ยของค่าความคลาดเคลื่อนในการประมาณพารามิเตอร์ a (RMSD ระหว่าง a และ a) ในกรณี 8 ตัวประกอบจะมีค่าน้อยกว่ากรณี 4 ตัวประกอบ ทั้งในกรณีของการกระจายถ่วงน้ำหนัก(loading)ของข้อสอบบนตัวประกอบร่วมเป็นแบบยูนิฟอร์ม และเป็นแบบเบ้ Harrison ยังได้เสนอแนะว่า ตัวประกอบร่วมอาจจะมากกว่า 1ได้ ทั้งที่ได้มีการพิจารณาที่รูปแบบข้อสอบแล้ว เช่น กรณี 150 ด้าน (facets) ของเชาว์ปัญญาที่เสนอโดยกิลฟอร์ด (Guilford, 1967) น่าจะใช้การวิเคราะห์ตัวประกอบชั้นที่ 2(second - order factors) ซึ่งน่าสนใจกว่า

ในปี 1987 Ratna Nandakuma ได้ศึกษาวิธีการของ stout ที่เรียกว่า Stout's procedure และมีความเห็นว่าสามารถนำมาใช้เพื่อการประเมินเอกมิติของแบบสอบได้ในหลายๆ สถานการณ์ในทางปฏิบัติ และเขายังได้เสนอแนะว่า การวิจัยครั้งต่อไป ควรคำนึงถึงการประยุกต์ของวิธีการสเต้า เพื่อวัตถุประสงค์อื่นในทางการวัดผลการศึกษาอื่น เช่น การขจัดความลำเอียงของแบบสอบและการเทียบแบบสอบ (test equating)

Abdelgadir (1987) ได้ทำการประเมินแบบเชิงประจักษ์ โดยการวิเคราะห์ตัวประกอบที่เรียกว่า การวิเคราะห์องค์ประกอบของข้อสอบเชิงสาระสนเทศเต็มรูปแบบ (Full - Information Item Factor Analysis) เรียกย่อ ๆ ว่า FIFA ซึ่งวิธีการนี้ได้นำเอาทั้งโมเดลของ IRT model และ Thurstone's mulltiple Factor model มาวิเคราะห์ร่วมกัน FIFA ถูกนำมาใช้ในการประเมินใน 4 สถานการณ์ คือ (1) การมีจำนวนตัวประกอบสำคัญ (major factor) ต่างๆ (2) การแจกแจงของตัวประกอบเหล่านี้ (3) การประมาณพารามิเตอร์การเดา (4) การมีองค์ประกอบย่อย (Minor factor) นอกจากนี้ ยังมีการศึกษาผลการลำเอียงในการผสมผสานเงื่อนไขเหล่านี้ ผลการศึกษาสรุปได้ว่า ถ้าข้อมูลมีข้อตกลงตามการวิเคราะห์แบบ FIFA การวิเคราะห์สรุปได้ว่า ถ้าข้อมูลมีข้อตกลงตามการวิเคราะห์แบบ FIFA การวิเคราะห์ตัวประกอบด้วยเมตริกซ์สหสัมพันธ์แบบเตตราคลอริก จะให้ข้อค้นพบที่ดีกว่าแบบที่วิเคราะห์จาก FIFA ตัวประกอบเมเจอร์ที่ไม่ปกติ (Non-normally) จะมีตัวประกอบไมเนอร์ และจำนวนตัวประกอบสำคัญก็ไม่ให้ผลกระทบที่ดีต่อวิธีการ แม้ว่าการประมาณการเดาจะไปแก้ความคลาดในการประมาณพารามิเตอร์ตัวอื่น ๆ ก็ตาม มันก็ไม่ได้ตัวประกอบที่สำคัญออกมา โดยเฉพาะอย่างยิ่งถ้าค่าสหสัมพันธ์เพียร์สันของค่าที่เหลือและการเปลี่ยนแปลงของความแปรปรวน เมื่อถูกนำมาใช้แทนสหสัมพันธ์เตตรา

ในปี 1988 Richard M Pomplum (Richard , 1988) ศึกษาผลของการฝ่าฝืนข้อตกลงของ IRT เกี่ยวกับความอิสระจากตำแหน่งของข้อสอบ เอาจากผลการตอบข้อสอบระดับวิทยาลัยในวิชาการสื่อสารและวิชาคณิตศาสตร์ มาใช้สมมุติเป็นโมเดลทฤษฎีการตอบสนองรายข้อกรณีพหุมิติ แล้วจัดให้ชุดของข้อมูลเหล่านี้มีระดับของการไม่อิสระต่าง ๆ กันด้วยวิธีการของโมเดลของ Ackernam (Ackerman's model) และพบผลกระทบของการประมาณความสามารถ โดยโมเดลหนึ่งพารามิเตอร์และสามพารามิเตอร์ เขาจึงเสนอแนะให้ระวังในการแปรความหมายของการประมาณความสามารถที่เป็นเอกมิติในขณะที่แบบสอบยังมีระดับของความไม่อิสระขั้นต่าง ๆ อยู่

Richard C Gershon (Richard ,1989) ได้ศึกษาผลของความกังวลต่อการทำข้อสอบ และการเรียงลำดับความยาก ซึ่งปรากฏว่า การลำดับความยากของข้อสอบ ความกังวลในการสอบ รวมถึงความสามารถของนักเรียน มีผลกระทบต่อการทำข้อสอบและการฝ่าฝืนข้อตกลงของ IRT และยังแสดงให้เห็นว่าทฤษฎีละเลยการอธิบายถึงความแตกต่างระหว่างบุคคลในเรื่องพฤติกรรมการสอบ และยังเกี่ยวข้องกับเครื่องมือวัด จึงได้เสนอแนะให้การทดสอบใหม่ซึ่งประกอบด้วยบุคลิกภาพและอุปนิสัยในการสอบอยู่ในผลการตอบข้อสอบด้วย

จะเห็นได้ว่า การตรวจสอบความเป็นเอกมิติของแบบสอบนั้น เป็นสิ่งจำเป็นอย่างยิ่ง และการฝ่าฝืนข้อตกลงนี้ ก็มีผลต่อการประมาณค่าความสามารถของบุคคล การวิจัยในครั้งนี้ จึงตระหนักถึงคุณภาพของดัชนีบ่งชี้ความเป็นเอกมิติของแบบสอบ ซึ่งได้เสนอขั้นตอนในการดำเนินการวิจัยไว้ใน บทที่ 3