บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การใช้แบบสอบวัดผลทางการศึกษามีจุดมุ่งหมายที่สำคัญ คือ การวัดความรู้ความสามารถเฉพาะด้าน (single ability) ของแต่ละวิชา ซึ่งเป็นคุณลักษณะแฝง (latent trait) ที่มีอยู่ในตัวของผู้สอบ ชุดของข้อสอบจะทำหน้าที่เป็นสิ่งเร้าให้ผู้สอบแสดงคุณลักษณะแฝงออกมา การที่แบบสอบจะสามารถวัดความสามารถเฉพาะด้านได้แสดงว่า แบบสอบนั้นจะต้องประกอบด้วยข้อสอบที่วัดความรู้ความสามารถในลักษณะเดียวกัน หรือ มีความเป็นเอกมิติ (uni-dimensionality) และการที่แบบสอบมีความเป็นเอกมิตินี้เองที่จะทำให้นักวัดผลสามารถแปลความหมายจากคะแนนได้ง่าย ชัดเจน และตรงไปตรงมา

แต่เดิมมา ในการวัดผลทางการศึกษา วิธีการสอบที่ได้รับความนิยมอย่างแพร่หลาย คือ การสอบข้อเขียน การสอบข้อเขียนมักจะใช้แบบความเรียง (essay type) หรือแบบอัตนัย ซึ่งก็พบกันว่าการทำข้อสอบแบบความเรียงนั้น ผู้ตอบต้องใช้ความรู้ความสามารถหลาย ๆ ด้านพร้อมกัน จนทำให้ยากที่จะระบุได้ชัดเจนว่า ข้อสอบนั้นวัดความสามารถเฉพาะด้านในด้านใด ประกอบกับการตรวจข้อสอบประเภทนี้ต้องใช้เวลาค่อนข้างมาก และอาจขาดความเป็นปรนัยทำให้เกิดปัญหาคุณภาพเครื่องมือ นอกจากนี้ยังมีข้อจำกัดในด้านความครอบคลุมเนื้อหาที่ต้องการวัด

เมื่อเป็นเช่นนี้ นักวัดผลจึงได้คิดรูปแบบของข้อสอบปรนัยขึ้นเพื่อแก้ปัญหาดังกล่าว ได้แก่ ข้อสอบแบบตอบสั้น แบบเติมข้อความ แบบถูกผิด แบบจับคู่ และแบบเลือกตอบ โดยเฉพาะแบบเลือกตอบเป็นแบบที่ได้รับความนิยมในเวลานี้ ในส่วนของการให้คะแนนนั้น แบบสอบปรนัยสามารถให้คะแนนแบบ 0 และ 1 หรือ ที่เรียกว่า ให้คะแนนแบบไบนารี (binary-scored item หรือ dichotomous-scored item) คือ ข้อที่ตอบผิดจะได้ 0 คะแนนและข้อที่ตอบถูกจะได้ 1 คะแนน ลักษณะของข้อสอบแบบปรนัยจะใช้เวลาในการตรวจน้อยลง สามารถใช้เครื่องจักรกลช่วยในการตรวจได้ และทำให้รายงานผลได้อย่างรวดเร็ว แม้จะมีผู้สอบจำนวนมากก็ตาม อีกทั้งยังสามารถสร้างข้อสอบให้ครอบคลุมเนื้อหาที่กำหนดได้ รวมทั้งสะดวกต่อการพัฒนาเป็นแบบสอบมาตรฐาน

แม้ว่าข้อสอบปรนัยที่ให้คะแนนแบบ 0 และ 1 จะมีประโยชน์ดังกล่าวข้างต้น แต่คุณสมบัติที่สำคัญของแบบสอบวัดผลทางการศึกษาที่พัฒนาขึ้นมาเหล่านี้ คือ การวัดความสามารถเฉพาะด้าน ดังนั้น การตรวจสอบความเป็นเอกมิติของแบบสอบ จึงเป็นขั้นตอนสำคัญในการพัฒนาแบบสอบวัดผลทางการศึกษา โดยมีปัญหาพื้นฐานอยู่ที่ว่า **"มีปัจจัยสำคัญใดบ้างที่มีผลต่อความเป็นเอกมิติของแบบสอบ"** และ **"วิธีตรวจสอบความเป็นเอกมิติแบบใดมีความเหมาะสมสำหรับการทดสอบในสถานการณ์แบบไหน"** ปัญหานี้มีความสำคัญเพราะเป็นการระบุถึงคุณภาพของเครื่องมือ ทำให้การศึกษาในเรื่องนี้ได้ขยายตัวอย่างรวดเร็ว และเกิดความคิดที่หลากหลายในการตรวจสอบความเป็นเอกมิติตามมาด้วย เริ่มต้นด้วยการที่นักวัดผลได้พยายามให้ความหมายของคำว่า "เอกมิติ" (unidimension) ขึ้น เพื่อนำไปสู่วิธีการที่นำมาใช้ในการตรวจสอบความเป็นเอกมิติ โดยมีวิวัฒนาการควบคู่กับทฤษฎีการวัดผลในช่วงระยะต่าง ๆ กัน

ในยุคของทฤษฎีวัดผลแบบดั้งเดิม (Classical Test Theory : CTT) ได้ให้ความหมายแรกของความเป็นเอกมิติว่า มีลักษณะเป็น **"ความสอดคล้องภายใน"** (internal consistency) (McArthur, 1987:36) ซึ่งเป็นการนำเอาแนวความคิดมาจากการตรวจสอบคุณภาพของเครื่องมือเกี่ยวกับความเที่ยง (reliability) ของแบบสอบ โดยถ้าแบบสอบมีความเป็นเอกมิติแล้ว จะต้องมีความสอดคล้องภายใน โดยมุ่งพิจารณาจากค่า item-intercorrelation ที่เป็นค่าแสดงถึงความสัมพันธ์ระหว่างข้อสอบแต่ละข้อในแบบสอบทั้งฉบับ ลักษณะของความเที่ยงแบบ internal consistency เป็นการดัดแปลงมาจากความเที่ยงแบบแบ่งครึ่งข้อสอบ (split halves) โดยยึดหลักที่ว่าข้อสอบแต่ละข้อมีลักษณะคู่ขนานกับแบบสอบทั้งฉบับ ค่าที่ได้เป็น intraclass correlation หรือ average inter-item correlation ซึ่งใช้กับการให้คะแนนแบบ 0 และ 1 วิธีการในชุดนี้จึงเป็นการใช้ค่าสัมประสิทธิ์ความเที่ยงจากสูตร KR-20 ของ Kuder-Richardson

$$KR\text{-}20 = \left[\frac{k}{k-1}\right]\left[1 - \frac{\Sigma\, p_i(1-p_i)}{S^2_x}\right]$$

โดย $p_i(1-p_i)$ = variance (S^2_i) ของ binary item

$$KR\text{-}20 = \left[\frac{k}{k-1}\right]\left[1 - \frac{\Sigma\, S^2_i}{S^2_x}\right]$$

แต่ S^2_x มีค่าเท่ากับผลรวมของความแปรปรวนของข้อสอบแต่ละข้อ (S^2_i) บวกด้วยสองเท่าของความแปรปรวนร่วมระหว่างข้อสอบ (S_{ij})

และจาก $r_{ij} = S_{ij}/S_iS_j$

ดังนั้น
$$KR\text{-}20 = \frac{r_{ij}S_iS_j}{1/k(\overline{S^2_i}) + (k-1)/k(\overline{r_{ij}S_iS_j})} \quad \text{.....(ก)}$$

$r_{ij}S_iS_j$ = inter item covariance

$\overline{S^2_i}$ = ค่าเฉลี่ยของ item variance

$\overline{r_{ij}S_iS_j}$ = ค่าเฉลี่ยของ inter item covariance

k = จำนวนข้อสอบ

(McArthur, 1987:36-37)

จากสูตร (ก) เป็นการวัดขอบเขตของ inter item-correlation โดยพิจารณาจากความเท่าเทียมกัน (equivalence) ของข้อสอบแต่ละข้อกับแบบสอบทั้งฉบับ ทำให้ค่า KR-20 ไม่ได้เป็นเพียงสัมประสิทธิ์ความเที่ยง แต่ยังมีส่วนสำคัญที่แสดงว่า ชุดข้อสอบที่ประกอบกันเป็นแบบสอบวัดคุณลักษณะเดียวกัน ซึ่งย่อมมีผลโดยตรงต่อ construct validity ของแบบสอบ ด้วยเหตุนี้ทำให้นักวัดผลหลายท่าน อาทิเช่น Allen & Yen, 1979; Gulliksen, 1950; Ghiselli,1964 และ Magnussen 1967 ใช้คำว่า ความเป็นเอกพันธ์ (homogeneous) บ่งชี้ความเป็นเอกมิติ (unidimensionality) สำหรับอ้างถึงคุณสมบัติของแบบสอบ

แต่การใช้ความเที่ยงแบบ internal consistency ซึ่งมีลักษณะบ่งชี้ถึงความเป็นเอกพันธ์ของแบบสอบเพื่อสรุปความเป็นเอกมิติของแบบสอบนั้น ยังมีข้อบกพร่องเพราะการวัดทางจิตวิทยาถือว่าความเป็นเอกพันธ์เป็นเงื่อนไขที่จำเป็น (necessary) ของความเป็นเอกมิติ แต่ไม่เพียงพอ(sufficient) ที่จะระบุความเป็นเอกมิติทั้งหมด (McArthur, 1987:37) โดยเฉพาะการใช้สูตร KR-20 เมื่อมีการเพิ่มจำนวนข้อสอบ โดยข้อสอบที่นำมาเพิ่มมีความสัมพันธ์ทางบวกกับข้อสอบเดิมทั้งหมด ค่าสัมประสิทธิ์ที่ได้จะมีค่าเข้าใกล้ 1 นอกจากนี้เมื่อมีการเปลี่ยนกลุ่มผู้สอบย่อมทำให้ค่าสัมประสิทธิ์ที่ได้เปลี่ยนแปลงตามกลุ่มผู้สอบไปด้วย ทำให้การแปลความหมายของระดับความเป็นเอกมิติไม่แน่นอนและผิดพลาดไปจากความเป็นจริงได้

นักวัดผลในยุคต่อมาเห็นว่าการพิจารณาค่า inter-item correlation ยังไม่เพียงพอที่จะแสดงถึงความเป็นเอกมิติจึงได้หันมาพิจารณาค่า item-total correlation ซึ่งเป็นค่าที่แสดงความสัมพันธ์ระหว่างข้อสอบแต่ละข้อกับแบบสอบทั้งฉบับ ผสมผสานกับแนวคิดในการวิเคราะห์องค์ประกอบ (factor analysis) โดยนำค่า item-total correlation ไปหาค่าสหสัมพันธ์กับค่าน้ำหนักองค์ประกอบ (factor loading) ขององค์ประกอบตัวแรกที่มีร้อยละความแปรปรวนสูงสุด วิธีการนี้มีชื่อเรียกว่า Biserial Test (Warm,1978:155) แต่วิธีการนี้ยังคงมีปัญหาเช่นกัน เพราะการพิจารณาค่าดัชนีที่ได้ไม่ชัดเจน จากการศึกษาของ Warm พบว่าค่าดัชนีของ Biserial Test ที่ได้จากคำนวณได้จากแบบสอบจริงมีค่าระหว่าง 0.419 ถึง 0.484 ส่วนค่าดัชนีที่คำนวณได้จากการสุ่มตัวเลขขึ้นมา มีค่าอยู่ระหว่าง 0.284 ถึง 0.348 ซึ่งดัชนีทั้งสองชุดมีค่าค่อนข้างต่ำ และที่สำคัญคือ เมื่อมีค่าการเดา (c) เกิดขึ้น ค่าดัชนีนี้ไม่มีโอกาสเป็น 1 หรือ 0 เลย เนื่องจากค่าดัชนีของ Biserial Test ที่คำนวณได้ไม่ชัดเจน จึงไม่เหมาะที่จะใช้วิธีการนี้ตรวจสอบความเป็นเอกมิติของแบบสอบ

แนวคิดต่อมาเป็นผลจากการเรียงลำดับคะแนนสอบของแต่ละบุคคลและความยากของข้อสอบในรูปของเมตริกซ์ David Walker (1931, 1936, 1940 อ้างใน McArthur, 1987 : 47) ซึ่งเป็นคนแรกที่เห็นประโยชน์จากเมตริกซ์นี้ และเริ่มค้นหาความสัมพันธ์ระหว่างรูปแบบการตอบกับรูปร่าง (shape) การแจกแจงของคะแนน Walker ได้ให้แนวคิดเกี่ยวกับรูปแบบการตอบตามอุดมคติ (ideal response) และพยายามสร้างดัชนีที่ชี้ถึงความห่างของคะแนนสอบที่ได้กับรูปแบบตามอุดมคติ แต่ความสนใจของเขาอยู่ที่ความหมายที่ได้จากการแจกแจงของคะแนนมากกว่าความหมายของผลที่ได้จากการวัด

ในปี 1944 Guttman ได้ทบทวนเรื่องนี้และคิดกระบวนการตรวจสอบถึงผลจากการสอบกับรูปแบบตามอุดมคติ ที่เรียกว่า "Guttman's Scalogram Analysis" โดยมีแนวคิดว่าเมื่อแบบสอบประกอบด้วยข้อสอบที่วัดคุณลักษณะเฉพาะด้านแล้ว หากนำข้อสอบมาเรียงตามลำดับตามค่าความยากของข้อสอบและคะแนนรวมของผู้สอบแต่ละคน ผู้สอบที่ตอบข้อสอบที่มีความยากระดับหนึ่งได้ถูกต้อง เขาจะตอบข้อสอบทุกข้อที่ง่ายกว่าข้อสอบนั้นได้ถูกต้องด้วย ในกรณีที่มีความคลาดเคลื่อนเกิดขึ้น ได้แก่ ความคลาดเคลื่อนจากการที่ผู้สอบตอบผิดในข้อที่ง่ายกว่า หรือยากเท่ากับความสามารถของเขา และความคลาดเคลื่อนจากการที่ผู้สอบตอบข้อสอบถูกในข้อที่ยากกว่าความสามารถของเขา ความคลาดเคลื่อนเหล่านี้ย่อมแสดงว่า มีคุณลักษณะแฝงมากกว่าหนึ่งมิติที่ใช้ในการทำข้อสอบชุดดังกล่าว ดัชนีที่ใช้วัดความคลาดเคลื่อนของ Guttman มีชื่อว่า "The Coefficient of Reproducibility" (REP)

แนวคิดในเรื่องนี้มีข้อโต้แย้งค่อนข้างมากเกี่ยวกับการเรียงลำดับความยากของข้อสอบ เพราะการสร้างข้อสอบให้มีความยากเรียงลำดับอย่างเป็นระบบเป็นไปได้ยาก รูปแบบการตอบข้อสอบที่ได้แต่ละครั้ง จะแตกต่างกันไปตามความสามารถของกลุ่มผู้สอบนั้น และที่สำคัญเป็นการพิจารณาบทบาทของผู้สอบมากกว่าบทบาทของข้อสอบ (McArthur, 1987 :46-52) จึงทำให้มีการพัฒนาแนวคิดนี้ไปเป็นดัชนีใช้ในการตรวจสอบ ความผิดแบบแผนการตอบของผู้สอบแต่ละคน ที่เรียกกันว่า Caution Index เพื่อใช้ประโยชน์ในการปรับปรุงการเรียนการสอน ส่วนความคิดที่จะใช้รูปแบบการตอบในการตรวจสอบความเป็นเอกมิติของ Guttman จึงเป็นอันยุติไป

ต่อมา เมื่อนักวัดผลได้เสนอทฤษฎีการตอบข้อสอบ (Item Response Theory : IRT) ในปี 1952 ทฤษฎีนี้ก็ได้รับความสนใจจากนักวิชาการอย่างกว้างขวางแทนที่ทฤษฎีแบบดั้งเดิม (CTT) อย่างรวดเร็ว เพราะ CTT เป็นทฤษฎีที่ตั้งอยู่บนข้อตกลงที่อ่อนและไม่สมเหตุสมผล (Lord, 1980 : 9; Allen & Yen,1979 : 57-60) ส่วน IRT เป็นทฤษฎีที่กล่าวถึงพฤติกรรมที่บุคคลตอบสนองต่อข้อสอบว่าถูกกำหนดขึ้นมาด้วยลักษณะบางอย่างภายในตัวของบุคคล (latent trait) ซึ่งไม่สามารถสังเกตได้โดยตรง จึงพยายามที่จะอธิบายความสัมพันธ์ระหว่าง ลักษณะภายในตัวบุคคล (trait) กับพฤติกรรมการตอบข้อสอบ และอธิบายความสัมพันธ์ดังกล่าวด้วยฟังก์ชั่นทางคณิตศาสตร์ ในการวัดผลทางการศึกษา ลักษณะของ IRT จะบอกถึงความสัมพันธ์ระหว่างความสามารถที่เป็นผลมาจากการเรียนในวิชาใดวิชาหนึ่ง กับโอกาสที่จะตอบข้อสอบแต่ละข้อในวิชานั้นได้ถูกต้อง การที่ IRT ได้รับความสนใจมากเป็นเพราะจุดเด่นที่สำคัญ 3 ประการคือ (Hambleton & Swaminathan, 1985:11)

1. ถ้ามีคลังข้อสอบ (item pool or item bank) ที่มีข้อสอบทุกข้อวัดคุณลักษณะ (trait) เดียวกัน การประมาณค่าความสามารถของผู้เข้าสอบจะเป็นอิสระจากกลุ่มตัวอย่างของข้อสอบที่ใช้ในการสอบ

2. ถ้าประชากรของผู้เข้าสอบเป็นประชากรขนาดใหญ่ การประมาณค่าพารามิเตอร์ของข้อสอบ จะเป็นอิสระจากกลุ่มตัวอย่างของผู้เข้าสอบที่ใช้ในการประมาณค่าพารามิเตอร์ของข้อสอบนั้น

3. มีค่าสถิติที่แสดงถึงความแม่นยำ (precision) ของการประมาณค่าความสามารถของผู้สอบ ค่าสถิตินี้จะเปลี่ยนแปลงไปในผู้สอบแต่ละคน

แม้ว่า IRT จะมีจุดเด่นหลายประการก็ตาม แต่ก็มีข้อตกลงเบื้องต้นที่เข้มงวด (strong assumption) ในการใช้ทฤษฎีที่สำคัญ 4 ข้อ คือ

1. ความเป็นเอกมิติของแบบสอบ (unidimensionality) คือการที่แบบสอบวัดคุณลักษณะแฝงเพียงลักษณะเดียว (single latent trait) การที่แบบสอบวัดคุณลักษณะแฝงเพียงอย่างเดียวนี้เองที่นำไปสู่ความหมายในอีกลักษณะหนึ่งของความเป็นเอกมิติ

2. ความเป็นอิสระต่อกัน (local independence) หมายถึงการที่ผู้สอบคนหนึ่งมีอิสระจากผู้สอบคนอื่น ๆ ในการทำข้อสอบ และการเป็นอิสระของผู้สอบในการทำข้อสอบข้อหนึ่ง ๆ จากการทำข้อสอบข้ออื่น ๆ

3. โค้งแสดงความสัมพันธ์ระหว่างความสามารถของผู้สอบและโอกาสที่จะตอบข้อสอบได้ถูกต้อง (item characteristic curve) ซึ่งเป็น logistic function มีลักษณะเป็น monotonic increasing function คือ เมื่อมีความสามารถเพิ่มขึ้นโอกาสที่จะตอบข้อสอบนั้นถูกต้องก็จะต้องเพิ่มมากขึ้นตามไปด้วย

4. ต้องไม่ใช่แบบสอบที่เป็น speed test ซึ่งทำให้ผู้สอบบางคนไม่ได้ทำข้อสอบในบางข้อเนื่องจากหมดเวลาการสอบ ไม่ได้เป็นเพราะผู้สอบไม่มีความสามารถที่จะทำข้อสอบนั้น

เมื่อ IRT มีลักษณะเด่นดังกล่าวและเหมาะสมในการพัฒนาแบบสอบมากกว่า CTT ทำให้การศึกษาเกี่ยวกับ IRT เริ่มขยายออกไปอย่างกว้างขวางและต่อเนื่องมาจนถึงปัจจุบัน โดยมีการศึกษาตั้งแต่ รายละเอียดเบื้องต้นของทฤษฎีเกี่ยวกับการตรวจสอบข้อตกลงเบื้องต้น วิธีการประมาณค่าพารามิเตอร์ และการนำทฤษฎีไปประยุกต์ใช้ในการเทียบมาตรา การสร้างคลังข้อสอบ การตรวจสอบความลำเอียงของข้อสอบ ตลอดจนการพัฒนาโมเดลการตอบข้อสอบจากเดิมที่ใช้กับแบบสอบที่มีคำตอบแบบ dichotomous เพื่อนำไปใช้กับคำตอบแบบ polytomous โมเดลที่เกิดขึ้นใหม่ ได้แก่ Nominal response model ของ Bock (1972) Graded response model ของ Samejima (1969) Successive interval model ของ Rost (1988) เป็นต้น

สิ่งที่น่าสนใจเกี่ยวกับการศึกษาข้างต้น คือ มีการศึกษารายละเอียดนับตั้งแต่การตรวจสอบข้อตกลงเบื้องต้นไปจนถึงการประยุกต์ใช้ทฤษฎี ทำให้ทฤษฎีขยายขอบเขตออกไปอย่างไม่หยุดยั้ง แม้ว่าการศึกษาในรายละเอียดเบื้องต้นของทฤษฎีจะยังไม่พบการกำหนดรูปแบบการปฏิบัติที่แน่นอนชัดเจนก็ตามที ซึ่งแตกต่างไปจากทฤษฎีอื่น ๆ ที่พบกัน และจะเห็นว่าการศึกษาเกี่ยวกับการตรวจสอบความเป็นเอกมิติของแบบสอบซึ่งเป็นข้อตกลงพื้นฐานที่สำคัญของทฤษฎีนี้ก็เริ่มมีการศึกษาในประเทศไทยโดย วรนุช แหยมแสง (2537) ได้ข้อค้นพบเบื้องต้นที่น่าสนใจ คือ

ดัชนี Eigen Ratio (ER) ที่ได้จากอัตราส่วนของค่า eigen ขององค์ประกอบที่ 1 และ eigen ขององค์ประกอบที่ 2 ของการวิเคราะห์องค์ประกอบเชิงสำรวจ (exploratory factor analysis) จาก tetrachoric correlation matrix และดัชนีความสอดคล้องของโมเดลตามโมเดล MIMIC (GFI) ของ Muthen มีความเหมาะสมที่จะใช้เป็นตัวบ่งชี้ความเป็นเอกมิติของแบบสอบ แต่ในการศึกษายังไม่ได้พิจารณารายละเอียดเกี่ยวกับจุดอ่อนของการวิเคราะห์องค์ประกอบที่เกิดจาก correlation matrix ชนิดต่าง ๆ ตลอดจนการเปลี่ยนแปลงค่าดัชนีภายใต้ปัจจัยที่เกี่ยวข้องกับข้อสอบและผู้สอบ จึงยังไม่มีข้อค้นพบที่ชัดเจนในทางปฏิบัติ หากมีการศึกษารายละเอียดเบื้องต้นชัดเจนและแน่นอนแล้ว น่าจะทำให้การนำ IRT ไปใช้ในการพัฒนาแบบสอบวัดผลทางการศึกษาได้เหมาะสมยิ่งขึ้น

จุดมุ่งหมายของการพัฒนาแบบสอบวัดผลทางการศึกษามีลักษณะคล้ายกันทั้ง CTT และ IRT คือการวัดความสามารถเฉพาะด้าน โดยความสามารถที่ต้องการวัดมีลักษณะเป็นคุณลักษณะแฝง (latent trait) มักมีคำถามอยู่เสมอว่า แบบสอบที่สร้างขึ้นวัดความสามารถในการทำข้อสอบเฉพาะด้านหรือไม่ ซึ่ง Lord และ Novick (1968 อ้างใน Hulin, Drasgow & Parsons, 1985:155) ได้กล่าวว่า IRT ตั้งอยู่บนข้อตกลงที่เข้มงวด (strong assumption) และไม่มีแบบสอบชุดใดที่จะมีคุณสมบัติครบตามข้อตกลงนี้อย่างสมบูรณ์ ลักษณะที่เป็นไปได้มากที่สุด คือ การพิจารณา **มิติที่เด่นเพียงมิติเดียว** (dominant dimension) ทำให้มีผู้พยายามเลี่ยงไปใช้ข้อตกลงที่ผ่อนคลายลง (Stout,1987 : 589-618) คือ ให้แบบสอบนั้นวัดคุณลักษณะที่ต้องการวัดเป็นส่วนใหญ่ และอาจมีคุณลักษณะอื่นที่เกี่ยวข้องเพียงบางข้อซึ่งหลีกเลี่ยงไม่ได้ ความเป็นเอกมิติในลักษณะนี้เรียกว่า **เอกมิติสำคัญ** (essential unidimension)

ในเรื่องความเป็นเอกมิติของแบบสอบนี้ Lord และ Novick ได้เน้นที่การสร้างข้อคำถามแต่ละข้อให้มีเนื้อหาสอดคล้องกัน และคัดเลือกข้อสอบอย่างสมเหตุสมผลเป็นอันดับแรก ซึ่งทำให้แบบสอบที่ได้มีลักษณะเป็นเอกพันธ์ (homogeneity) และเป็นส่วนหนึ่งที่แสดงถึงความเป็นเอกมิติของแบบสอบ แต่ Hulin, Drasgow & Parsons (1983 :40-41) มีความเห็นว่าข้อตกลงเรื่องความเป็นเอกมิติของแบบสอบเป็นเรื่องที่สำคัญของ IRT การพิจารณาความเป็นเอกพันธ์ของแบบสอบเพียงอย่างเดียวยังไม่เพียงพอ และควรมีการตรวจสอบให้ชัดเจนก่อนที่จะนำ IRT ไปใช้

Lumsden (1961 อ้างใน Hambleton & Swaminathan, 1985:155) ได้ทบทวนการสร้างแบบสอบที่มีความเป็นเอกมิติและสรุปว่า การวิเคราะห์องค์ประกอบ (factor

analysis) น่าจะเป็นวิธีการที่ดีที่สุดในการทดสอบความเป็นเอกมิติ และในเวลา 15 ปีต่อมา เขายังคงยืนยันความคิดเกี่ยวกับการใช้การวิเคราะห์องค์ประกอบตรวจสอบความเป็นเอกมิติ วิธีการของ Lumsden เป็นการใช้การวิเคราะห์องค์ประกอบวิเคราะห์ข้อสอบที่สร้างขึ้น ข้อสอบใดที่ไม่อยู่ใน dominant factor จะถูกตัดออกไป จากนั้นนำข้อที่เหลือมาวิเคราะห์องค์ประกอบต่อไปเรื่อย ๆ จนได้ผลเป็นที่พอใจ และใช้สัดส่วนของ first factor variance ต่อ second factor variance เป็นดัชนีบ่งชี้ความเป็นเอกมิติ (index of unidimensionality) แต่วิธีการของ Lumsden จะมีปัญหาเรื่องความครบถ้วนของเนื้อหา ถึงกระนั้น Hambleton และ Traub (1973 อ้างใน Hambleton & Swaminathan, 1985:156) ยังคงเสนอให้ใช้การวิเคราะห์องค์ประกอบในการตรวจสอบความเป็นเอกมิติเช่นเดียวกัน นอกจากนี้ยังมีนักวัดผลอีกหลายท่านที่เสนอให้ใช้การวิเคราะห์องค์ประกอบ โดยใช้ principal component analysis และพิจารณาค่าความแปรปรวนสูงสุดจากองค์ประกอบหลักตัวแรก (first major component) (Hambleton, 1988; Hambleton et al., 1991; Knol & Bergerenschede, 1988; Lord, 1980 อ้างใน Roznowski & Tucker, 1991) เมื่อความแปรปรวนในองค์ประกอบแรกมีสัดส่วนมากกว่าองค์ประกอบอื่น ๆ มากก็น่าเชื่อได้ว่าแบบสอบนั้นมีแนวโน้มเป็นเอกมิติ

อย่างไรก็ตาม การใช้วิธีการวิเคราะห์องค์ประกอบยังพบว่ามีจุดอ่อนเกี่ยวกับ correlation matrix ที่ใช้เป็น matrix เริ่มต้นของการวิเคราะห์ โดย correlation matrix ที่ใช้กับแบบสอบที่ให้คะแนนแบบ 0 และ 1 อยู่มี 2 ชนิด (Hambleton & Swaminathan ,1985: 21) คือ

1. Correlation matrix ที่คำนวณจาก phi correlation เป็นวิธีการที่พัฒนามาจาก Pearson Product Moment Correlation วิธีการนี้จะให้จำนวนองค์ประกอบ (factor) ที่มากเกินความเป็นจริง องค์ตัวประกอบที่เกินมาเรียกว่า "Difficulty Factor" เป็นผลมาจากช่วงของค่าความยาก (b) กว้างมาก การแก้ไขทำได้โดยการจำกัดค่า b ไม่ให้สูงจนเกินไป (extreme) ซึ่งการสร้างข้อสอบโดยควบคุมความยากไปด้วยในขณะเดียวเป็นสิ่งที่ทำได้ยาก (McDonald & Ahlawat ,1974 อ้างใน Hambleton & Swaminathan, 1985:156)

2. Correlation matrix ที่คำนวณด้วย tetrachoric correlation โดย Lord & Novick (1968) เห็นด้วยกับวิธีการนี้ แม้ว่าวิธีนี้จะให้องค์ประกอบร่วมเพียงตัวเดียว แต่ก็ยังพบว่า ถ้ามีค่าการเดาเกิดขึ้น ทำให้ผู้สอบที่มีความสามารถต่ำตอบข้อสอบที่ยากเกินความ

สามารถได้เช่นเดียวกับผู้สอบที่มีความสามารถสูงเป็นผลให้ข้อสอบนั้นมีอำนาจจำแนกต่ำไปด้วย และทำให้ correlation matrix ที่ได้อาจจะไม่เป็น positively definite ซึ่งเป็นปัญหาสำคัญของการวิเคราะห์องค์ประกอบ

จากข้อจำกัดข้างต้นเมื่อข้อสอบมีค่าการเดาเกิดขึ้น ทำให้ correlation matrix ที่ได้ ไม่เป็น positively definite หรือที่เรียกว่า non-Gramian matrix ซึ่งเป็นเมตริกซ์ที่ไม่เหมาะสมในการวิเคราะห์องค์ประกอบ (Lord & Novick,1961:349;Roznoski & Tucker,1991:109-127) ทำให้ผลการวิเคราะห์องค์ประกอบมีความคลาดเคลื่อน ซึ่งปัญหาเรื่องนี้ได้มีการแก้ไขในโปรแกรมสำเร็จรูป ที่วิเคราะห์องค์ประกอบด้วย tetrachoric correlation matrix ได้แก่ โปรแกรม TESTFACT และ LISREL โดยจะมีคำสั่งย่อยที่ช่วยปรับข้อมูลให้เรียบ

ส่วนในเรื่องของการแก้อิทธิพลของการเดา ถ้าสามารถปรับแก้ค่าการเดาในการคำนวณค่า tetrachoric correlation จะช่วยในแก้ปัญหา non-Gramian matrix และทำให้ผลการวิเคราะห์องค์ประกอบถูกต้องมากขึ้น Carroll (1945 อ้างใน Hulin,Drasgow & Parsons,1983 : 249-255) เป็นผู้หนึ่งที่ได้เสนอวิธีการแก้ไขค่าการเดาที่มีผลต่อการคำนวณค่า tetrachoric correlation coefficient

นอกจากนี้ยังมีผู้เสนอให้ใช้ variance-covariance matrix แทน correlation matrix ดังกล่าวข้างต้น (Roznowski,1991:109-207) และพบว่าผลที่ได้มีความคงที่มากกว่า phi correlation matrix และ tetrachoric correlation matrix ในการวิเคราะห์องค์ประกอบเชิงสำรวจ (exploratory factor analysis) ตลอดจนมีการเสนอให้ใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor analysis) แทนการวิเคราะห์องค์ประกอบเชิงสำรวจ (Jöreskog & Sörbom,1989:230-232) ในการตรวจสอบความเป็นเอกมิติ

ปัญหาของการใช้วิธีการวิเคราะห์องค์ประกอบที่ควรจะศึกษา คือ "การเลือกใช้ correlation matrix ว่าควรใช้วิธีใด และใช้ควบคู่กับการวิเคราะห์องค์ประกอบเชิงสำรวจ (exploratory factor analysis) หรือ การวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor analysis) จึงจะเหมาะสมสำหรับตรวจสอบความเป็นเอกมิติของแบบสอบ"

การศึกษาเกี่ยวการตรวจสอบความเป็นเอกมิติได้ขยายตัวต่อไปอีก เมื่อนักวิชาการพบความสัมพันธ์ระหว่างข้อตกลงของ IRT เรื่องความเป็นเอกมิติ (unidimensionality) และความเป็นอิสระต่อกันในการตอบข้อสอบ (local independence) แม้ว่าในปี 1978 Warm (1978:107) เสนอความเห็นไว้ว่า ถ้าแบบสอบมีความเป็นเอกมิติแล้ว ก็เป็นเงื่อนไขที่เพียงพอว่าแบบสอบนั้นมีความเป็นอิสระต่อกันในการตอบข้อสอบ แต่ถ้าแบบสอบมีความเป็นอิสระต่อกันในการตอบข้อสอบแล้ว ยังไม่เพียงพอที่จะบอกได้ว่าแบบสอบนั้นมีความเป็นเอกมิติหรือไม่ ซึ่งไม่เห็นด้วยกับการนำวิธีการตรวจสอบความเป็นอิสระต่อกันมาใช้ตรวจสอบความเป็นเอกมิติ แต่ต่อมาในปี 1985 Hambleton & Swaminathan (1985:22-25) ได้กล่าวถึงความเป็นอิสระต่อกันในการตอบข้อสอบว่า สามารถใช้แสดงถึงความเป็นเอกมิติของแบบสอบได้ จึงมีนักวัดผลหลายท่านสนใจศึกษาถึงความเป็นอิสระต่อกันในการตอบข้อสอบอีกครั้งหนึ่งดังจะเห็นได้จากการศึกษา Roznowski, Tucker & Humphreys (1991:109-127) ที่ใช้แนวคิดของการใช้ความเป็นอิสระต่อกันในการตอบข้อสอบเพื่อบ่งชี้ความเป็นเอกมิติ ซึ่งมีรายละเอียดของแนวทางการการตรวจสอบดังนี้

ความเป็นอิสระต่อกันในการตอบข้อสอบ (Local Independence) หมายถึง การตอบข้อสอบของผู้สอบในแต่ละข้อ จะมีความเป็นอิสระต่อกันทางสถิติ คือ เนื้อหาของข้อสอบข้อหนึ่งต้องไม่มีผลต่อการตอบข้อสอบข้ออื่น ลำดับของข้อสอบจะต้องไม่มีผลต่อการทำข้อสอบ ดังนั้นความน่าจะเป็นของการตอบข้อสอบทั้งชุดเท่ากับผลคูณของความน่าจะเป็นของการตอบข้อสอบแต่ละข้อ สิ่งที่แสดงถึงความสัมพันธ์ระหว่างข้อตกลงเรื่องความเป็นเอกมิติของแบบสอบและความเป็นอิสระต่อกันในการตอบข้อสอบ เริ่มจากชุดของข้อสอบที่วัดความสามารถเฉพาะด้าน และใช้ผู้สอบที่ระดับความสามารถ (θ) เดียวกัน แล้วพบว่าการตอบข้อสอบเป็นอิสระต่อกันทางสถิติ แบบสอบชุดนั้นจะมีความเป็นเอกมิติ ถ้าที่ระดับความสามารถเดียวกัน การตอบข้อสอบไม่เป็นอิสระต่อกันทางสถิติ ทำให้เห็นว่าผู้สอบบางคนมีคะแนนมากกว่าผู้สอบคนอื่น ๆ ที่ระดับความสามารถเดียวกัน แสดงว่าผู้สอบต้องใช้ความสามารถมากกว่า 1 ด้านในการทำข้อสอบ แบบสอบนั้นจึงขาดคุณสมบัติของความเป็นเอกมิติ

เนื่องจากความเท่าเทียมกันของความเป็นเอกมิติและความเป็นอิสระต่อกันในการตอบข้อสอบ จึงสามารถทดสอบความเป็นอิสระต่อกันในการตอบข้อสอบด้วยการวิเคราะห์ตัวประกอบและตรวจสอบความเป็นเอกมิติด้วยความเป็นอิสระต่อกันในการตอบข้อสอบ (Lord, 1953a อ้างใน Hambleton & Swaminathan,1985:22-25) การทดสอบความเป็นอิสระต่อกันในการตอบข้อสอบ ทำได้โดยการคำนวณค่า chi-square test ของข้อสอบทีละคู่ที่ระดับความสามารถเดียวกัน แล้วเปลี่ยนค่าความสามารถที่อยู่ในสเกลเดียวกันไปเรื่อย ๆ การทดสอบจะมีจำนวนครั้ง

เท่ากับจำนวนคู่ของข้อสอบคูณด้วยช่วงของความสามารถ เช่นถ้ามีข้อสอบ 10 ข้อ จะจัดคู่ได้ 45 คู่ และถ้าแบ่งความสามารถเป็น 10 ช่วง การทดสอบจะเกิดขึ้นทั้งหมด 450 ครั้ง ซึ่งมีขั้นตอนค่อนข้างมากและแปลความหมายไม่ชัดเจน หากมีข้อสอบคู่หนึ่งไม่เป็นอิสระต่อกันที่ระดับความสามารถหนึ่งแล้ว ส่วนข้อสอบคู่ที่เหลือเป็นอิสระต่อกันในการตอบข้อสอบ ด้วยเหตุนี้จึงทำให้นักวัดผลยังไม่นิยมใช้การตรวจสอบความเป็นอิสระต่อกันในการตอบข้อสอบเพื่ออ้างอิงความเป็นเอกมิติ

เมื่อพิจารณาวิธีการพัฒนากระบวนการตรวจสอบความเป็นเอกมิติทั้งหลายที่กล่าวมา จะเห็นได้ว่าวิธีการที่ได้รับความสนใจในปัจจุบัน คือ การวิเคราะห์องค์ประกอบเชิงสำรวจ และการวิเคราะห์องค์ประกอบเชิงยืนยัน ซึ่งเป็นการนำผลที่ได้จากการวิเคราะห์มาพัฒนาเป็นดัชนีบ่งชี้ความเป็นเอกมิติ โดยในการวิเคราะห์องค์ประกอบเชิงสำรวจ ผลที่นำมาใช้เป็นดัชนี ได้แก่ ดัชนีอัตราส่วนไอเกน (eigen ratio : ER) ได้จากอัตราส่วนระหว่างค่า eigen ขององค์ประกอบที่ 1 และ ค่า eigen ขององค์ประกอบที่ 2

ส่วนการวิเคราะห์องค์ประกอบเชิงยืนยัน ซึ่งพิจารณาความเป็นเอกมิติ จากการทดสอบความสอดคล้องระหว่างข้อมูลกับโมเดล ที่กำหนดให้มีตัวแปรแฝงเพียงตัวเดียวที่อยู่เบื้องหลังตัวแปรที่สังเกตได้ทั้งชุด การวิเคราะห์แบบนี้ใช้ในการวิเคราะห์โมเดลโครงสร้าง (structural model analysis) อาศัยหลักการพื้นฐานของการวิเคราะห์โมเดลโครงสร้างความแปรปรวนร่วม (covariance structure model) การวิเคราะห์ structural equation model มีโมเดลในการวิเคราะห์ 2 โมเดล คือ (Gerbing,1979; Joreskog & Sorbom,1978 อ้างใน Anderson & Gerbing,1982 : 453)

1. Measurement model เป็นโมเดลที่แสดงความสัมพันธ์ระหว่างตัวแปรที่สังเกตได้ (observed variables) หรือ ตัวบ่งชี้ (indicators) กับตัวแปรแฝง (latent variable) หรือ ภาวะสันนิษฐาน (construct)

2. Structural equation model เป็นโมเดลที่แสดงถึงความสัมพันธ์ระหว่างตัวแปรแฝง หรือ ภาวะสันนิษฐาน

ใน measurement model โดยที่ตัวแปรแฝงเป็นสิ่งที่ไม่สามารถสังเกตได้จึงต้องอาศัยตัวบ่งชี้ที่เป็นตัวแทนของตัวแปรแฝงนั้น ตัวแปรแฝงแต่ละตัวจะมีตัวบ่งชี้เพียงตัวเดียว หรืออาจมีตัวบ่งชี้เป็นชุดก็ได้ เรียกว่า multiple indicators คุณสมบัติที่สำคัญของชุดตัวบ่งชี้ คือ การที่ทุกตัวบ่งชี้เป็นตัวแทนของตัวแปรแฝงตัวเดียวกัน นั่นคือ ความเป็นเอกมิติของชุดตัวบ่งชี้จึงควรมีการตรวจสอบความเป็นเอกมิติของชุดตัวบ่งชี้เสียก่อน ที่จะตรวจสอบโมเดลโครงสร้าง

ทั้งหมด ในลักษณะเดียวกันแบบสอบแต่ละชุดก็ทำหน้าที่เป็นชุดตัวบ่งชี้ตัวแปรแฝงเกี่ยวกับความสามารถด้านใดด้านหนึ่งของผู้สอบ โดยทั่วไปนิยมการทดสอบความสอดคล้องของโมเดลกับข้อมูล หากพบว่ามีความสอดคล้องเกิดขึ้นย่อมหมายความว่า ชุดของตัวบ่งชี้ของตัวแปรแฝงแต่ละตัวใน measurement model เป็นตัวแทน (represent) ตัวแปรแฝงเดียวกัน การตรวจสอบความเป็นเอกมิติของแบบสอบด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน จึงทดสอบความสอดคล้องของข้อมูลกับโมเดลในส่วนของ measurement model เท่านั้น

วิธีการที่ใช้ทดสอบความสอดคล้องของโมเดลกับข้อมูลวิธีหนึ่ง คือ การใช้ χ^2-test เนื่องจากค่า χ^2 จะมีค่าเพิ่มขึ้นตามขนาดของกลุ่มตัวอย่าง เป็นผลให้การทดสอบมีแนวโน้มจะมีนัยสำคัญ จึงไม่เหมาะที่จะนำมาเป็นดัชนีบ่งชี้ความเป็นเอกมิติ ทำให้มีการพัฒนาดัชนีอื่น ๆ ขึ้น ได้แก่ Goodness of fit index (GFI) แต่ดัชนี GFI มีค่าที่มีความสัมพันธ์กับค่า χ^2 ซึ่ง Maiti & Mukherjee (1990 อ้างใน Jöreskog & Sörbom,1993:122) พบว่า ดัชนี GFI มีความสัมพันธ์กับค่า χ^2 แบบ monotonic relationship จึงมีการพัฒนา adjusted goodness of fit index (AGFI) โดยเป็นการปรับดัชนี GFI ด้วยค่า degrees of freedom (df)

ในปัจจุบันพบว่ามีดัชนีนอกเหนือจาก GFI และ AGFI อีกหลายตัวที่ได้รับการพัฒนาขึ้นเพื่อตรวจสอบความสอดคล้องของโมเดลเช่นเดียวกัน (Jöreskog & Sörbom, 1993:125-131) ได้แก่

1. Root mean square residual error (RMR)
2. Normed fit index (NFI) และ non-normed fit index (NNFI) ของ Bentler & Bonett (1980)
3. Relative fit index (RFI) และ incremental fit index (IFI) ของ Bollen (1986,1989a)
4. Critical N (CN) ของ Hoelter (1983)

เมื่อเป็นเช่นนี้ คำถามต่อมาของการตรวจสอบความเป็นเอกมิติ จึงอยู่ที่ว่า **"ดัชนีตัวไหนที่เหมาะสมในการบ่งชี้ความเป็นเอกมิติมากกว่ากัน"**

นอกจากนี้ในการสอบแต่ละครั้งยังมีปัจจัยอื่น ๆ ที่เกี่ยวข้องกับผลคะแนนสอบที่ได้ ได้แก่

1. กระบวนการที่ใช้ในการเรียนการสอน เช่น วิธีการสอนของครูผู้สอน ระยะเวลาในการสอน อุปกรณ์การสอน ความสามารถของครูผู้สอน เป็นต้น

2. การใช้จำนวนข้อสอบที่แตกต่างกันในการสอบแต่ละครั้ง เมื่อมีจำนวนข้อสอบเพิ่มขึ้นจะมีผลโดยตรงต่อค่า correlation coefficient ทำให้ค่าที่คำนวณได้เพิ่มขึ้นตามไปด้วย

3. การใช้แบบสอบชุดเดียวกันกับจำนวนผู้สอบที่แตกต่างกันย่อมส่งผลต่อค่า correlation coefficient เช่นเดียวกับจำนวนข้อสอบ

4. ค่าความยากของข้อสอบ ถ้ามีช่วงกว้างมากจะทำให้เกิด difficulty factor ในการวิเคราะห์ตัวประกอบ

5. การเลือกใช้โมเดลในการวิเคราะห์ ใน IRT ประกอบด้วย 3 โมเดล คือ 1, 2 และ 3 พารามิเตอร์ ย่อมมีผลที่แตกต่างกันออกไป โดยเฉพาะในโมเดล 3-พารามิเตอร์ ซึ่งมักใช้กับแบบสอบวัดผลทางการศึกษา เมื่อมีค่าการเดา (c) เกิดขึ้นจะมีผลต่อการคำนวณค่า tetrachoric correlation และส่งผลต่อไปยังการวิเคราะห์ตัวประกอบ (Carroll,1945 อ้างใน Hulin,Drasgow & Parsons,1983 : 249-250) ดังได้กล่าวมาแล้วข้างต้น

Reckase (1990 อ้างใน Nandakumar & Stout,1993 : 41-68) จึงให้ความเห็นว่า ความเป็นเอกมิติจะต้องเป็นคุณสมบัติของทั้งแบบสอบและผู้สอบที่ทำแบบสอบนั้นด้วย ดังนั้นประเด็นปัญหาที่สำคัญอีกประเด็นหนึ่งของการตรวจสอบความเป็นเอกมิติ จึงควรศึกษาถึง **"ดัชนีที่มีคุณภาพเหมาะสมที่จะใช้ในการตรวจสอบความเป็นเอกมิติภายใต้ปัจจัยต่าง ๆ ที่เกี่ยวข้องกับแบบสอบและตัวผู้สอบ"** ส่วนกระบวนการในการเรียนการสอนนั้น แม้ว่าจะเกี่ยวข้องคะแนนสอบที่ได้ แต่ไม่ได้นำมาพิจารณาร่วมด้วย เนื่องจากเป็นกระบวนการที่เกิดขึ้นก่อนการสอบ มีช่วงเวลาในดำเนินการ และมีองค์ประกอบที่ซับซ้อนมาผสมผสานกัน ไม่ว่าจะเป็นวิธีการสอน พฤติกรรมการสอนของครู หรือพฤติกรรมการเรียนของผู้เรียน อุปกรณ์ที่ใช้ในการเรียนการสอน ตลอดจนการใช้สื่อการสอนประเภทต่าง ๆ การควบคุมให้สถานการณ์เหล่านี้มีลักษณะเดียวกันหมดจึงเป็นสิ่งที่ทำได้ค่อนข้างยาก จึงควรมีการศึกษาโดยเฉพาะแยกออกไป

ในเรื่องของดัชนีที่ใช้ในการบ่งชี้ความเป็นเอกมิติ ลักษณะของดัชนีที่ได้ควรมีความคงที่ภายใต้การเปลี่ยนแปลงของปัจจัยต่าง ๆ เพราะการสร้างข้อสอบวัดผลทางการศึกษาในปัจจุบัน มีแนวโน้มที่จะสร้างเป็นคลังข้อสอบ (item bank) ขนาดใหญ่ และเลือกใช้ข้อสอบเพียงจำนวนหนึ่งซึ่งมีจำนวนแตกต่างกันไปในการสอบแต่ละครั้ง การตรวจสอบความเป็นเอกมิติของข้อสอบทั้งหมดในคลังข้อสอบในคราวเดียวกันจึงเป็นไปได้ยาก เนื่องมาจากจำนวนข้อสอบ ประกอบกับขีดจำกัดของเครื่องคอมพิวเตอร์และโปรแกรมสำเร็จรูปที่ใช้ในการวิเคราะห์ นอกจากนี้ดัชนีที่ได้ควรจะต้องมีความไวต่อการเปลี่ยนแปลงความเป็นเอกมิติพอสมควร เมื่อมีการเจือปนของข้อสอบในมิติอื่นอีกด้วย

งานเกี่ยวกับการศึกษาวิธีการตรวจสอบความเป็นเอกมิติมีความสำคัญเพิ่มขึ้นเมื่อ Anderson & Gerbing (1983)ได้ศึกษาพบว่า การใช้ชุดของตัวแปรที่สังเกตได้ (observed variables) บ่งชี้ตัวแปรแฝง (latent variable) ในโมเดลการวัด (measurement model) ซึ่งเป็นโมเดลที่แสดงความสัมพันธ์ระหว่างตัวแปรที่สังเกตได้กับตัวแปรแฝงของ covariance structure equation model นั้น และเรียกชุดของตัวแปรที่สังเกตได้นี้ว่า multiple indicators หากการวัดตัวแปรแฝงตัวหนึ่งต้องใช้ตัวแปรที่สังเกตได้หลายตัวในการวัดแล้ว ชุดของตัวแปรที่สังเกตได้ หรือตัวบ่งชี้เหล่านั้นก็ควรบ่งชี้ตัวแปรแฝงเพียงตัวเดียว และมีลักษณะเป็นโครงสร้างเดี่ยว (single construct) คือ มีความเป็นเอกมิตินั่นเอง Anderson & Gerbing ยังพบว่า นอกจากความเป็นเอกมิติของชุดตัวบ่งชี้ จะส่งผลต่อความแม่นยำในการวัดตัวแปรแฝงแล้ว ยังส่งผลไปยังโมเดลโครงสร้าง (structural model) ซึ่งเป็นโมเดลที่แสดงถึงความสัมพันธ์ระหว่างตัวแปรแฝงกับตัวแปรแฝงในการศึกษาโมเดลเชิงสาเหตุ (causal model) ด้วย covariance structure equation model ทำให้ผลการทดสอบความสอดคล้องระหว่างข้อมูลและโมเดลคลาดเคลื่อนไปจากความเป็นจริง

แต่อย่างไรก็ตาม โปรแกรมสำเร็จรูปที่ใช้วิเคราะห์เชิงยืนยันเพิ่งจะเริ่มแพร่หลายเข้ามาในประเทศไทย โปรแกรมที่รู้จักกัน คือ LISREL ของ Jöreskog และ Sörbom ซึ่งยังใช้กันอยู่ในวงจำกัด ทำให้นักวัดผลส่วนใหญ่ยังคงใช้การวิเคราะห์องค์ประกอบเชิงสำรวจ ตรวจสอบความเป็นเอกมิติ และมีโปรแกรมที่นิยมใช้กัน 2 โปรแกรม คือ SPSS และ SAS โดยเฉพาะ SPSS ซึ่งมีการใช้อย่างกว้างขวางในเวลานี้

ดังนั้น ในการศึกษาความเป็นเอกมิติของแบบสอบครั้งนี้จึงมุ่งพิจารณาถึง "ปัจจัยต่าง ๆ เกี่ยวกับการสอบประกอบด้วยการเปลี่ยนแปลงของจำนวนข้อสอบ การเปลี่ยนแปลงของจำนวนผู้สอบ และค่าความยากของข้อสอบ ควบคู่ไปกับชนิดของเมตริกซ์และวิธีวิเคราะห์องค์ประกอบที่มีผลต่อการเปลี่ยนแปลงของดัชนีที่ใช้บ่งชี้ความเป็นเอกมิติของแบบสอบ สำหรับแบบสอบที่ให้คะแนนแบบไบนารี โดยพิจารณาดัชนีที่ได้จากการวิเคราะห์องค์ประกอบที่มีคุณสมบัติในแง่ความคงที่ หรือมีการเปลี่ยนแปลงน้อยที่สุดภายใต้การเปลี่ยนแปลงของของปัจจัยเกี่ยวกับ จำนวนข้อสอบ จำนวนผู้สอบ และค่าความยากของข้อสอบ รวมทั้งเป็นดัชนีที่มีความไวต่อการเปลี่ยนแปลงความเป็นเอกมิติของแบบสอบ"

วัตถุประสงค์ของการวิจัย

1. เพื่อวิเคราะห์ดัชนีบ่งชี้ความเป็นเอกมิติของแบบสอบ ที่ได้จากการวิเคราะห์องค์ประกอบเชิงสำรวจ และการวิเคราะห์องค์ประกอบเชิงยืนยัน ด้วย correlation matrix 4 ชนิด ประกอบด้วย tetrachoric correlation matrix 3 ชนิด คือ เมตริกซ์เดิม เมตริกซ์ที่มีการปรับเรียบข้อมูล เมตริกซ์ที่มีการปรับแก้ค่าการเดา และ variance-covariance matrix ภายใต้การเปลี่ยนแปลงของจำนวนข้อสอบ จำนวนผู้สอบ และค่าความยากของข้อสอบ

2. เพื่อศึกษาถึงความไวและความคงที่ของดัชนีแต่ละประเภท ในการตรวจสอบความเป็นเอกมิติของแบบสอบ

สมมุติฐานการวิจัย

การตรวจสอบความเป็นเอกมิติของแบบสอบ ด้วยการใช้ดัชนีที่ได้จากการวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis : EFA) และการวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis : CFA) โดยวิธีการวิเคราะห์ทั้ง 2 แบบมีแนวคิดที่สำคัญดังนี้

EFA เป็นการวิเคราะห์องค์ประกอบ เมื่อผู้วิจัยต้องการศึกษาเชิงสำรวจกลุ่มตัวแปรด้วยวิธีทางคณิตศาสตร์ โดยไม่จำเป็นต้องใช้ทฤษฎีใดโดยเฉพาะ ผู้วิจัยจะต้องเก็บข้อมูลจากกลุ่มตัวแปรนั้น แล้วจึงนำมาจัดหมวดหมู่ด้วยการวิเคราะห์องค์ประกอบ องค์ประกอบที่ได้จะถูกตั้งชื่อโดยอาศัยตัวแปรที่มีค่าสหสัมพันธ์สูงสุด ซึ่งมีส่วนในการอธิบายองค์ประกอบนั้นมากที่สุด (Bernstein, Garbin & Teng, 1988 : 164-165)

ส่วน CFA เป็นการวิเคราะห์องค์ประกอบ เมื่อผู้วิจัยมีสมมุติฐานที่แน่นอน โดยมีตัวแปรแฝง (latent variable) ระหว่างกลุ่มตัวแปรที่ทำการศึกษา และใช้ความรอบคอบในการเลือกตัวแปรมาวิเคราะห์องค์ประกอบ เพื่อเปิดเผยตัวแปรแฝงนั้นให้ชัดเจนเท่าที่จะทำได้ (Mulaik, 1972 : 362) องค์ประกอบที่ได้จึงเป็นผลจากการกำหนดตามทฤษฎีที่มีอยู่ (Bernstein, Garbin & Teng, 1988 : 164-165) CFA เป็นเทคนิคที่อาศัยหลักการของ covariance structure analysis เพื่อตรวจสอบความสอดคล้องระหว่างข้อมูลและโมเดลทางทฤษฎี

การตรวจสอบความเป็นเอกมิติของแบบสอบวัดผลทางการศึกษาตามแนวของ IRT เมื่อพิจารณาถึงความหมายของการวิเคราะห์องค์ประกอบทั้งสองชนิดข้างต้น การใช้ CFA มาตรวจสอบความเป็นเอกมิติ น่าจะเหมาะสมมากกว่า EFA

การพิจารณาดัชนีที่ใช้สำหรับตรวจสอบความเป็นเอกมิติของแบบสอบ นอกเหนือจากวิธีการวิเคราะห์องค์ประกอบแต่ละวิธี ยังต้องพิจารณาถึงเมตริกซ์ที่ใช้วิเคราะห์องค์ประกอบและดัชนีที่ใช้บ่งชี้ความเป็นเอกมิติ นั้นด้วย

จากงานวิจัยที่เกี่ยวกับการใช้เมตริกซ์ในการวิเคราะห์องค์ประกอบแบบ EFA เริ่มจาก ในปี 1945 Carroll ได้ศึกษาถึงปัญหาการเดา (c) ที่มีผลต่อ tetrachoric correlation และได้ใช้วิธีแก้ค่าการเดาในตาราง contigency ขนาด 2x2 ก่อนที่จะคำนวณค่า tetrachoric correlation เพื่อนำไปใช้ในการวิเคราะห์องค์ประกอบ

ต่อมา Drasgow & Lissak (1982) ได้ศึกษาการใช้ tetrachoric correlation แล้วใช้ค่า correlation coefficient สูงสุดในตัวแปรนั้นแทนค่าในแนวทแยง ในการวิเคราะห์องค์ประกอบเพื่อตรวจสอบความเป็นเอกมิติ

นอกจากนี้ Roznowski (1991) พบว่าการใช้ variance-covariance matrix ในการวิเคราะห์องค์ประกอบจะได้ผลที่คงที่มากกว่าการใช้ phi correlation matrix และ tetrachoric correlation matrix

ส่วนงานวิจัยเกี่ยวกับดัชนีจากการวิเคราะห์แบบ EFA จากผลการวิจัยของ วรนุช แหยมแสง (2537) พบว่า ดัชนีอัตราส่วนค่าไอเกน (Eigen Ratio:ER) ซึ่งเป็นอัตราส่วนระหว่างค่า eigen ขององค์ประกอบที่ 1 และค่า eigen ขององค์ประกอบที่ 2 มีความเหมาะสมมากที่สุดในการบ่งชี้ความเป็นเอกมิติ ในกรณีที่ใช้ tetrachoric correlation matrix

แต่เมื่อพิจารณาค่า eigen ที่ได้จากการวิเคราะห์องค์ประกอบแบบ EFA หากแบบสอบมีความเป็นเอกมิติลดลง จะพบว่าค่า eigen ขององค์ประกอบที่ 2 และ 3 จะมีค่าเพิ่มขึ้น ในขณะที่ค่า eigen ขององค์ประกอบที่ 1 ลดลง จึงน่าจะนำค่าดังกล่าวมาใช้คำนวณเป็นดัชนี โดยนำอัตราส่วนของ eigen ขององค์ประกอบที่ 1 และค่า eigen ขององค์ประกอบที่ 2 หารด้วยอัตราส่วนของ eigen ขององค์ประกอบที่ 2 และค่า eigen ขององค์ประกอบที่ 3 ดัชนีที่ได้จึงเป็นค่าของอัตราส่วนของอัตราส่วนค่าไอเกน (Ratio of Eigen Ratio:ERR)

ในด้านงานวิจัยที่เกี่ยวกับเมตริกซ์และดัชนีในการวิเคราะห์องค์ประกอบแบบ CFA มีดังนี้

Wilson และคณะ (1991) ได้พัฒนาโปรแกรม TESTFACT เพื่อใช้ในการวิเคราะห์ข้อสอบ และในการวิเคราะห์องค์ประกอบของแบบสอบ โดยใช้ tetrachoric correlation

โดยมีค่าสั่งย่อยช่วยในการปรับเรียบข้อมูลเพื่อแก้ปัญหา non-positively definite ก่อนที่จะวิเคราะห์องค์ประกอบ ดัชนีจากการวิเคราะห์องค์ประกอบแบบ CFA ด้วยการกำหนดให้ข้อมูลมีเพียงองค์ประกอบเดียวและทดสอบจำนวนองค์ประกอบด้วย chi-square approximation for the likelihood ratio test (G^2) ใช้ได้เมื่อกลุ่มตัวอย่างขนาด 2^n โดย n คือ จำนวนข้อสอบ ตามวิธีการ Bock's full information factor analysis (Wilson และคณะ, 1991 : 15)

Jöreskog & Sörbom (1989) ได้พัฒนาโปรแกรม LISREL และใช้ variance-covariance matrix วิเคราะห์องค์ประกอบและใช้ดัชนีที่ได้จากการวิเคราะห์ตรวจสอบความเป็นเอกมิติของแบบสอบ ในโปรแกรม LISREL ของ Jöreskog & Sörbom ได้รวบรวมดัชนีต่าง ๆ ที่มีผู้พัฒนาเพื่อใช้ตรวจสอบความสอดคล้องระหว่างข้อมูลและโมเดลที่กำหนด หรือที่เรียกว่า fit index แบ่งเป็น 3 ชุด แต่ละชุดมีดัชนีที่สำคัญ คือ

1. การทดสอบแบบ overall fit ประกอบด้วย χ^2, Root Mean Square Residuals (RMR)

2. การทดสอบด้วย incremental fit index ประกอบด้วย Non-normed Fit Index (NNFI)

3. การทดสอบด้วย overall fit index ประกอบด้วย Adjusted Goodness of Fit Index (AGFI), Critical N (CN)

การพิจารณาเทคนิคทางสถิติที่ใช้ตรวจสอบเป็นเพียงส่วนหนึ่งของการตรวจสอบความเป็นเอกมิติเท่านั้น ในการตรวจสอบความเป็นเอกมิติของแบบสอบ จำเป็นจะต้องนำแบบสอบที่ต้องการตรวจสอบไปทดลองใช้กับกลุ่มผู้สอบ ซึ่งมีจำนวนแตกต่างกันไปในแต่ละสถานการณ์ของการสอบ ส่วนจำนวนข้อสอบ ก็แตกต่างกันไปตามเนื้อหาที่ใช้ในการสร้าง และเวลาที่ใช้ในการสอบ นอกจากนี้จำนวนข้อสอบยังเป็นปัญหาต่อการวิเคราะห์องค์ประกอบอีกด้วย (Hair และคณะ, 1988 : 248) เมื่อพิจารณาควบคู่ไปกับ IRT ทั้งจำนวนข้อสอบ และจำนวนผู้สอบล้วนมีผลต่อการประมาณค่าพารามิเตอร์ ได้แก่ ค่าความยากของข้อสอบ ค่าอำนาจจำแนก และค่าการเดา โดยเฉพาะค่าความยากของข้อสอบซึ่งมีผลต่อ EFA ที่ใช้ tetrachoric correlation และทำให้เกิดการเดาในกลุ่มที่มีความสามารถต่ำ เป็นผลให้ค่าอำนาจจำแนกของข้อสอบข้อนั้นต่ำไปด้วย แต่การสร้างแบบสอบที่จำกัดช่วงความยากยังคงเป็นสิ่งที่ทำได้ยาก และหากว่า "**ความเป็นเอกมิติจะต้องเป็น**

คุณสมบัติของทั้งผู้สอบและแบบสอบ" (Reckase, 1990) แล้ว การที่แบบสอบมีความเป็นเอกมิติย่อมส่งผลต่อคุณภาพของแบบสอบในเรื่องของความตรงเชิงทฤษฎี (construct validity)

เมื่อเป็นเช่นนี้ การพิจารณาดัชนีที่ได้จากการวิเคราะห์แบบ EFA และ CFA จึงควรพิจารณาคุณสมบัติของดัชนีในแง่ความคงที่เมื่อมีการเปลี่ยนแปลงจำนวนข้อสอบ จำนวนผู้สอบ และค่าความยากของข้อสอบ และในขณะเดียวกันควรมีความไวต่อการเจือปนของข้อสอบในมิติอื่นประกอบด้วย จึงจะเหมาะสม

จากแนวคิดข้างต้นทั้งหมด ผู้วิจัยจึงได้ตั้งสมมุติฐานการวิจัยดังนี้

1. เมื่อจำนวนข้อสอบ จำนวนผู้สอบ และค่าความยากของข้อสอบ เปลี่ยนแปลงไป การใช้ variance-covariance matrix ใน EFA และ CFA น่าจะให้ค่าดัชนีทุกประเภทมีความคงที่มากกว่าค่าดัชนีที่ใช้ tetrachoric correlation matrix แบบเดิม tetrachoric correlation matrix ที่มีการปรับเรียบข้อมูล และ tetrachoric correlation matrix ที่มีการแก้ค่าการเคาในการวิเคราะห์

2. เมื่อพิจารณาคุณสมบัติของดัชนีที่ได้จากการวิเคราะห์ น่าจะแบ่งดัชนีได้ 4 กลุ่มด้วยกัน คือ

2.1 ดัชนีที่มีความไวและความคงที่ ประกอบด้วย NNFI และ CN

2.2 ดัชนีที่มีความไวแต่ไม่คงที่ ประกอบด้วย AGFI

2.3 ดัชนีที่มีความคงที่แต่ไม่มีความไว ประกอบด้วย RMR

2.4 ดัชนีที่ไม่มีความไวและไม่มีความคงที่ ประกอบด้วย ER, ERR, G^2 และ χ^2

ขอบเขตของการวิจัย

1. การวิจัยครั้งนี้มุ่งตรวจสอบความเป็นเอกมิติของแบบสอบวัดผลทางการเรียนที่มีการให้คะแนนแบบ 0 และ 1 (binary-scored item) ภายใต้ข้อตกลงเบื้องต้นของ IRT เกี่ยวกับความเป็นเอกมิติในลักษณะมิติที่เด่นมิติเดียว (dominant dimension)

2. ข้อสอบที่ใช้ในการศึกษา เป็น

2.1 ข้อสอบคัดเลือกเข้ามหาวิทยาลัยปี 2536 วิชาภาษาอังกฤษ กข 100 ข้อ

2.2 ข้อสอบวิชาภาษาอังกฤษ และข้อสอบวิชาคณิตศาสตร์ ชั้นประถมปีที่ 5 ของวรนุช แหยมแสง (2537)

นำข้อมูลการตอบข้อสอบมาวิเคราะห์ตามแนว IRT ด้วยโปรแกรม BILOG โดยคัดเลือกข้อสอบที่เหมาะกับโมเดล 3 พารามิเตอร์

3. การตรวจสอบความเป็นเอกมิติของแบบสอบใช้ดัชนีที่ได้จากการวิเคราะห์องค์ประกอบ โดยใช้วิธีการดังนี้

3.1 การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory factor analysis) สำหรับสร้างดัชนี ER และ ERR ใช้ matrix ในการวิเคราะห์ 4 ชนิด

ก. Tetrachoric corretation matrix แทนค่าในแนวทแยงด้วย correlation coefficient ที่มีค่าสูงสุดในตัวแปรนั้น วิเคราะห์ด้วยโปรแกรม SPSS/PC+

ข. Tetrachoric corretation matrix มีการปรับเรียบข้อมูลเพื่อแก้ปัญหาการที่ matrix เป็น non-positively definite วิเคราะห์ด้วยโปรแกรม SPSS/PC+

ค. Tetrachoric corretation matrix ที่มีการปรับแก้ค่าการเดา วิเคราะห์ด้วยโปรแกรม SPSS/PC+

ง. Variance-covariance matrix วิเคราะห์ด้วยโปรแกรม SPSS/PC+

3.2 การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory factor analysis) สำหรับสร้างดัชนี G^2 ใช้ matrix ในการวิเคราะห์ 2 ชนิด

ก. Tetrachoric corretation matrix มีการปรับเรียบข้อมูลเพื่อแก้ปัญหาการที่ matrix เป็น non-positively definite วิเคราะห์ด้วยโปรแกรม TESTFACT

ข. Tetrachoric corretation matrix ที่มีการปรับแก้ค่าการเดา วิเคราะห์ด้วยโปรแกรม TESTFACT

3.3 การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory factor analysis) สำหรับสร้างดัชนี χ^2, AGFI, RMR, NNFI, และ CN ใช้ matrix ในการวิเคราะห์ 4 ชนิด

ก. Tetrachoric corretation matrix แทนค่าในแนวทแยงด้วย correlation coefficient ที่มีค่าสูงสุดในตัวแปรนั้น วิเคราะห์ด้วยโปรแกรม LISREL

ข. Tetrachoric corretation matrix มีการปรับเรียบข้อมูลเพื่อแก้ปัญหาการที่ matrix เป็น non-positively definite วิเคราะห์ด้วยโปรแกรม LISREL

ค. Tetrachoric corretation matrix ที่มีการปรับแก้ค่าการเดา วิเคราะห์ด้วยโปรแกรม LISREL

ง. Variance-covariance matrix วิเคราะห์ด้วยโปรแกรม LISREL

ข้อจำกัดในการวิจัย

1. ข้อสอบภาษาอังกฤษที่ใช้วิเคราะห์เป็นข้อสอบคัดเลือกเข้ามหาวิทยาลัยประกอบด้วยข้อคำถามเกี่ยวกับความรู้ในด้าน grammar comprehension และ vocabulary ความเป็นเอกมิติของแบบสอบในการวิจัยครั้งนี้จึงพิจารณาความสามารถในวิชาภาษาอังกฤษในลักษณะของ general factor
2. ข้อสอบที่ใช้ในการวิเคราะห์เป็นข้อสอบสอบเข้ามหาวิทยาลัย ซึ่งเป็นข้อสอบที่ใช้ในการคัดเลือกบุคคล และเป็นการประเมินผลแบบอิงกลุ่ม ส่วนข้อสอบวิชาภาษาอังกฤษและคณิตศาสตร์ของ วรนุช แหยมแสง เป็นข้อสอบที่ประเมินผลแบบอิงเกณฑ์ ซึ่งการศึกษาครั้งนี้ไม่ได้มีการพิจารณาถึงความแตกต่างของการประเมินผลแบบอิงกลุ่มและอิงเกณฑ์ ที่มีผลต่อการวิเคราะห์ข้อสอบตามแนว IRT
3. ในการวิเคราะห์องค์ประกอบเชิงสำรวจที่คำนวณจาก tetrachoric correlation matrix แบบเดิม มีปัญหาเกี่ยวกับในเรื่อง non-positively definite ของเมตริกซ์ ซึ่งเป็นข้อจำกัดของโปรแกรม SPSS/PC^{+} ที่ไม่สามารถคำนวณค่าสถิติได้ครบทุกขั้นตอน ค่าไอเกนที่ผู้วิจัยนำมาใช้คำนวณค่าดัชนีจึงเป็นค่าสถิติเริ่มต้น (initial statistics) เท่านั้น

นิยามศัพท์

ความเป็นเอกมิติของแบบสอบ หมายถึง การที่แบบสอบมีคุณลักษณะแฝง (latent trait) ที่เด่นที่สุด (domonant) เพียงคุณลักษณะเดียว

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ดัชนีบ่งชี้ความเป็นเอกมิติ หมายถึง ดัชนีบ่งชี้ความเป็นเอกมิติของแบบสอบจากการวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) และการวิเคราะห์องค์ประกอบเชิงยืนยัน (CFA)

การวิเคราะห์องค์ประกอบเชิงสำรวจ (exploratory factor analysis : EFA) หมายถึง การวิเคราะห์องค์ประกอบด้วยโมเดลเชิงสำรวจ ซึ่งมีข้อตกลงเบื้องต้นว่า (Long,1982:12 อ้างใน ศิริชัย กาญจนวาสี, 2532:13)

1. ตัวประกอบร่วมทุกตัวมีความสัมพันธ์กัน (oblique rotation) หรือตัวประกอบร่วมทุกตัวเป็นอิสระต่อกัน (orthogonal rotation)

2. ตัวแปรที่สังเกตค่าได้ทุกตัวได้รับอิทธิพลโดยตรงจากทุกตัวประกอบ

3. ตัวแปรที่สังเกตค่าได้ทุกตัวได้รับอิทธิพลจากตัวประกอบเฉพาะหรือความคลาดเคลื่อนเพียงตัวเดียว

4. ความคลาดเคลื่อนทุกตัวเป็นอิสระต่อกัน และเป็นอิสระจากตัวประกอบทุกตัว เมื่อผู้วิจัยต้องการศึกษากลุ่มตัวแปร โดยยังไม่มีความรู้ในเรื่องนั้นมาก่อน

การวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor analysis : CFA) หมายถึง การวิเคราะห์องค์ประกอบด้วยโมเดลเชิงยืนยัน ซึ่งมีการผ่อนคลายข้อตกลงเบื้องต้นของ EFA และผู้วิจัยสามารถเพิ่มข้อจำกัดบางประการที่สอดคล้องกับแนวคิด/ทฤษฎีที่ต้องการทดสอบได้ เช่น ผู้วิจัยสามารถวางเงื่อนไขให้ ตัวประกอบบางคู่มีความสัมพันธ์กัน เลือกตัวแปรที่สังเกตค่าได้บางตัวให้ได้รับอิทธิพลโดยตรงจากเพียงบางตัวประกอบ เลือกตัวแปรที่สังเกตได้เพียงบางตัวที่ได้รับอิทธิพลจากความคลาดเคลื่อน หรือกำหนดให้ความคลาดเคลื่อนของตัวแปรบางคู่มีความสัมพันธ์กัน เป็นต้น และสามารถตรวจสอบความสอดคล้องระหว่างข้อมูลและโมเดลทางทฤษฎี

ดัชนีบ่งชี้ความเป็นเอกมิติจาก EFA หมายถึง ดัชนีอัตราส่วนไอเกน (Eigen Ratio: ER) และดัชนีอัตราส่วนของอัตราส่วนไอเกน (Eigen Ratio Ratio:ERR)

ค่าไอเกน (Eigen Value) หมายถึง ผลรวมของความแปรปรวนของตัวแปรแต่ละตัว ที่มีส่วนในการอธิบายองค์ประกอบนั้น คำนวณจากผลบวกของกำลังสองของน้ำหนักองค์ประกอบของตัวแปรแต่ละตัว ในองค์ประกอบหนึ่ง ๆ ก่อนที่จะมีการหมุนแกน

ดัชนีอัตราส่วนไอเกน (Eigen Ratio:ER) หมายถึง ค่าอัตราส่วนของไอเกนขององค์ประกอบที่ 1 และไอเกนขององค์ประกอบที่ 2

ดัชนีอัตราส่วนของอัตราส่วนไอเกน (Ratio of Eigen Ratio : ERR) หมายถึง ค่าอัตราส่วนของอัตราส่วนของค่าไอเกนขององค์ประกอบที่ 1 และค่าไอเกนขององค์ประกอบที่ 2 และค่าอัตราส่วนของค่าไอเกนขององค์ประกอบที่ 2 และค่าไอเกนขององค์ประกอบที่ 3

ดัชนีบ่งชี้ความเป็นเอกมิติจาก CFA หมายถึง ดัชนี Likelihood ratio chi-square (G^2) ในโปรแกรม Testfact,Chi-square (χ^2) ในโปรแกรม LISREL,Adjusted Goodness of Fit Index (AGFI),Root Mean Square Residual (RMR), Non-Normed Fit Index (NNFI), และ Critical N (CN)

ดัชนีจากการทดสอบ chi-square สำหรับ likelihood ratio (G^2) หมายถึง ดัชนีที่ใช้ในการทดสอบจำนวนองค์ประกอบของชุดข้อมูล เป็นการใช้การทดสอบ chi-square ประมาณค่า likelihood ratio เพื่อทดสอบความเหมาะสมของโมเดลเมื่อมีการกำหนดจำนวนองค์ประกอบไว้ล่วงหน้า เมื่อค่า G^2 มีนัยสำคัญแสดงว่าข้อมูลมีจำนวนองค์ประกอบเท่าที่กำหนดในการทดสอบ คำนวนได้จากสูตร

$$G^2 = 2 \sum_{I=1}^{2^n} rl \cdot \ln \frac{rl}{N\tilde{P}l}$$

rl = จำนวนความถี่ใน pattern l

$\tilde{P}l$ = item parameter ได้จากการประมาณค่าด้วยวิธี maximum likelihood

Σrl = จำนวนข้อมูล

ค่า degrees of freedom เท่ากับ $2^n(m+1)+m(m-1)/2$

m = จำนวนองค์ประกอบ

n = จำนวนข้อสอบ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ดัชนีความสอดคล้องระหว่างข้อมูลและโมเดลที่กำหนด หมายถึง ดัชนีที่ใช้ในการพิจารณาความสอดคล้องระหว่างโมเดลทางทฤษฎีที่เกี่ยวกับตัวแปรแฝงที่ศึกษากับข้อมูลที่เก็บรวบรวมจริงจากตัวแปรต่าง ๆ เหล่านั้น ในการวิเคราะห์องค์ประกอบเชิงยืนยัน ประกอบด้วย

1. Chi-square test ใช้ทดสอบจำนวนองค์ประกอบ คำนวณจากสูตร (นงลักษณ์ วิรัชชัย,2537:48)

$$\chi^2 = (n-1)F[s,\Sigma(\theta)]; \quad df=\{(k)(k+1)/2\}-t$$

F = F[s,Σ(θ)] = ค่าต่ำสุดของฟังก์ชั่นความกลมกลืนของโมเดลจากพารามิเตอร์ θ

n = ขนาดของกลุ่มตัวอย่าง

k = จำนวนตัวแปรที่สังเกตได้

t = จำนวนพารามิเตอร์อิสระ

2. Adjusted Goodness of Fit Index (AGFI) คำนวณจากค่า GFI และปรับแก้ด้วยค่า degree of freedom เพื่อแก้ไขค่าที่ได้อันเนื่องจากขนาดของกลุ่มตัวอย่าง ดังสูตร

$$AGFI = 1 - \frac{(p+q)(p+q+1)}{2d}(1 - GFI)$$

p = จำนวน observed variables ในที่หมายถึงจำนวนข้อสอบ

q = จำนวน predictor variables ในการศึกษาความเป็นเอกมิติของแบบสอบไม่ได้กำหนดในโมเดล ดังนั้น q = 0

d = degrees of freedom ของโมเดล

$$GFI = 1 - \frac{(s - \hat{\sigma})' W^{-1} (s - \hat{\sigma})}{s' W^{-1} s}$$

s = variance-covariance matrix ของกลุ่มตัวอย่าง

$\hat{\sigma}$ = variance-covariance matrix ของประชากรตามทฤษฎี

W = เมตริกซ์น้ำหนักที่ใช้ปรับค่าในการคำนวณ

3. Root Mean Square Residual (RMR) เป็นค่าที่วัดความคลาดเคลื่อนเฉลี่ยของข้อมูลจากกลุ่มตัวอย่างที่คลาดเคลื่อนไปจากโมเดลทางทฤษฎี (Average of the fitted residuals) คำนวณจากสูตร

$$RMR = \left[2 \sum_{i=1}^{p+q} \sum_{j=1}^{i} (S_{ij} - \hat{\sigma}_{ij})^2 / (p+q)(p+q+1) \right]^{1/2}$$

4. Non-normed fit index (NNFI) คำนวณจากสูตร

$$NNFI = (f_i - f)/(f_i - 1)$$

$f_i = nF_b/d_b$

$f = nF_m/d_m$

F_m = ค่าต่ำสุดของ fit function หรือ $F[s, \Sigma(\theta)]$ สำหรับโมเดลที่ถูกประมาณค่า

F_b = ค่าต่ำสุดของ fit function หรือ $F[s, \Sigma(0)]$ สำหรับโมเดลที่เป็น baseline

5. Critical N (CN) คำนวณจากสูตร

$$CN = \frac{\chi^2_{1-\alpha}}{F} + 1$$

คุณภาพของดัชนี หมายถึง การที่ดัชนีมีคุณสมบัติ 2 ประการ คือ

1. มีความคงที่ภายใต้การเปลี่ยนแปลงของจำนวนข้อสอบ จำนวนผู้สอบ ค่าความยากของข้อสอบ
2. มีความไวในการตรวจสอบความเป็นเอกมิติ เมื่อเจือปนด้วยข้อสอบมิติอื่น

ความคงที่ หมายถึง การที่ดัชนีแต่ละประเภทมีอัตราการเปลี่ยนแปลงต่ำ หรือค่อนข้างต่ำเมื่อมีการเปลี่ยนแปลงจำนวนข้อสอบ จำนวนผู้สอบ และค่าความยากของข้อสอบ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ความไวของการตรวจสอบ หมายถึง อัตราการเปลี่ยนแปลงของค่าดัชนีในการตรวจสอบความเป็นเอกมิติเมื่อลดความเป็นเอกมิติของแบบสอบ ด้วยการเพิ่มข้อสอบที่วัดความสามารถด้านอื่นเข้าไปในแบบสอบเดิมทีละข้อ

ประโยชน์ที่คาดว่าจะได้รับ

จะได้วิธีการที่มีคุณภาพและเหมาะสมในการตรวจสอบความเป็นเอกมิติ ในแบบสอบวัดผลทางการศึกษาที่มีการให้คะแนนแบบ 0 และ 1 โดยดัชนีที่ได้จะมีความไวในการตรวจสอบและมีความคงที่เมื่อมีการเปลี่ยนแปลงจำนวนข้อสอบ จำนวนผู้สอบ และมีการกระจายค่าความยากของข้อสอบที่แตกต่างกัน นอกจากนี้จะเป็นแนวทางที่มีประโยชน์ต่อการตรวจสอบความเป็นเอกมิติของชุดข้อสอบที่ได้มาจากคลังข้อสอบขนาดใหญ่ ซึ่งไม่สามารถตรวจสอบความเป็นเอกมิติของข้อสอบทั้งหมดได้ในคราวเดียวกัน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย