บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อวิเคราะห์ดัชนีบ่งชี้ความเป็นเอกมิติของแบบสอบที่ได้จากการวิเคราะห์องค์ประกอบเชิงสำรวจ และการวิเคราะห์องค์ประกอบเชิงยืนยัน โดยใช้เมตริกซ์ 4 ชนิด ประกอบด้วย tetrachoric correlation matrix 3 ชนิด คือ เมตริกซ์แบบเดิม เมตริกซ์ที่มีการปรับเรียบข้อมูล เมตริกซ์ที่มีการแก้ค่าการเดา และ variance-covariance matrix ภายใต้การเปลี่ยนแปลงของจำนวนข้อสอบ จำนวนผู้สอบ และค่าความยากของข้อสอบ ตลอดจนเพื่อศึกษาถึงคุณภาพของดัชนีด้านความคงที่และด้านความไวในการตรวจสอบความเป็นเอกมิติของแบบสอบ โดยใช้แบบสอบคัดเลือกเข้ามหาวิทยาลัย วิชาภาษาอังกฤษ กข ปี 2536 แบบสอบวิชาภาษาอังกฤษและวิชาคณิตศาสตร์ ชั้นประถมปีที่ 5 ของ วรนุช แหยมแสง (2537)

การวิเคราะห์ข้อมูลประกอบด้วย

1 การเปรียบเทียบเมตริกซ์ที่ให้ค่าดัชนีคงที่เมื่อมีการเปลี่ยนแปลงจำนวนผู้สอบ จำนวนข้อสอบ และค่าความยากของข้อสอบ

2 การเปรียบเทียบความคงที่ของดัชนีแต่ละประเภท เมื่อมีการเปลี่ยนแปลงจำนวนผู้สอบ จำนวนข้อสอบ และค่าความยากของข้อสอบ

3 การเปรียบเทียบความไวของดัชนีต่อการเจือปนของข้อสอบในมิติอื่น

4 การวิเคราะห์คุณภาพของดัชนีแต่ละประเภท

5 การวิเคราะห์ผลของความเป็นเอกมิติในแบบสอบต่อการวิเคราะห์ข้อสอบ ตามแนว IRT ด้วยโปรแกรม BILOG

สรุปผลการวิจัย

1. ดัชนี ER, ERR, X^2, AGFI, RMR, NNFI และ CN มีค่าดัชนีคงที่มากที่สุดเมื่อใช้ variance-covariance matrix ในการคำนวณ

ส่วนดัชนี G^2 มีค่าดัชนีคงที่มากที่สุดเมื่อใช้ tetrachoric correlation matrix ที่มีการปรับเรียบข้อมูล และดัชนี RMR มีค่าดัชนีความคงที่มากที่สุดเมื่อใช้ tetrachoric correlation matrix แบบเดิม

2. คุณภาพของดัชนี

2.1 ดัชนี AGFI เป็นดัชนีที่มีความคงที่มากที่สุด เมื่อมีการเปลี่ยนแปลงจำนวนผู้สอบ จำนวนข้อสอบ และค่าความยาก รองลงมาคือ ดัชนี NNFI

2.2 ดัชนี ERR เป็นดัชนีที่มีความไวต่อการเจือปนของข้อสอบในมิติอื่นมากที่สุด รองลงมา คือ ดัชนี NNFI

คุณภาพของดัชนีที่ได้จากการศึกษา มีดังนี้

ดัชนีที่มีความคงที่และมีความไว คือ ดัชนี NNFI

ดัชนีที่ไม่มีความคงที่ แต่มีความไว คือ ดัชนี ERR

ดัชนีที่มีความคงที่ แต่ไม่มีความไว คือ ดัชนี AGFI และ ER

ดัชนีที่ไม่มีความคงที่และไม่มีความไว คือ ดัชนี G^2, χ^2, RMR และ CN

3. การวิเคราะห์ผลของความเป็นเอกมิติในแบบสอบต่อการวิเคราะห์ข้อสอบตามแนว IRT พบว่า มีการเปลี่ยนแปลงของค่าพารามิเตอร์ เมื่อมีการเจือปนข้อสอบในมิติอื่นเพื่อทำให้แบบสอบไม่เป็นเอกมิติ ทำให้

ค่าอำนาจจำแนก (a) โดยเฉลี่ยสูงขึ้น

ค่าความยาก (b) โดยเฉลี่ยลดลง

ค่าการเดา (c) โดยเฉลี่ยสูงขึ้น

ค่าสารสนเทศของข้อสอบ (IIF) โดยเฉลี่ยลดลง และ

ค่าสารสนเทศของแบบสอบ (TIF) โดยเฉลี่ยลดลง

แสดงว่าความเป็นเอกมิติของแบบสอบมีผลต่อการวิเคราะห์ข้อสอบตามแนว IRT โดยเฉพาะการประมาณค่าพารามิเตอร์ และค่าสารสนเทศที่ได้จากการวิเคราะห์ด้วยโปรแกรม BILOG

อภิปรายผล

1. การใช้ tetrachoric correlation matrix ที่มีการปรับเรียบข้อมูลและที่มีการแก้ค่าการเดา ซึ่งเป็นการแก้ปัญหา non-positively definite ของ tetrachoric correlation matrix แบบเดิม เพื่อคำนวณค่าดัชนี ER นั้น จากการศึกษาครั้งนี้พบว่าค่า correlation ของตัวแปรแต่ละคู่ในเมตริกซ์มีการเปลี่ยนแปลงไปจากเดิมค่อนข้าง

มาก โดยใน tetrachoric correlation matrix ที่มีการปรับเรียบข้อมูลพบว่า ร้อยละ 35 ของค่า correlation มีค่าเปลี่ยนไปจากเดิม ค่าที่เปลี่ยนมีค่าตั้งแต่ -0.35 ถึง 0.95

ส่วนใน tetrachoric correlation matrix ที่มีการแก้ค่าการเดา พบว่า ร้อยละ 40 ของค่า correlation มีค่าเปลี่ยนไปจากเดิม ค่าที่เปลี่ยนมีค่าตั้งแต่ -0.74 ถึง 0.93 โดยค่าที่เปลี่ยนไปมีช่วงกว้างมากกว่า tetrachoric correlation matrix ที่มีการปรับเรียบข้อมูล แม้ว่างานวิจัยต่อมาของ Carroll (1983 อ้างใน Tucker,1985:215) ยังคงยืนยันว่าการเดามีผลต่อค่า tetrachoric correlation ระหว่างข้อสอบแต่ละคู่ ในเรื่องนี้พิจารณาได้จากสูตร

$$r_{\phi} = \cos\left[\frac{180^{\circ}}{1+\sqrt{ad/bc}}\right]$$

เมื่อมีการเดาเกิดขึ้น ทำให้ผู้ที่มีความสามารถต่ำตอบข้อสอบได้ถูกต้อง เป็นผลให้ bc มีโอกาสเท่ากับ 0 หรือ ad/bc มีโอกาสเท่ากับ 1 ทำให้ r_{ϕ} มีค่าเป็น -1 หรือ 0 เป็นผลให้เมตริกซ์ที่ได้เป็น non-positively definite แต่การแก้ค่าการเดาด้วยวิธีที่ Carroll คิดขึ้นในปี 1945 ยังมีผลที่ไม่สามารถยืนยันได้แน่นอน โดยเฉพาะเมื่อใช้ใน CFA ดัชนีที่คำนวณได้ส่วนใหญ่รายงานผลว่า ข้อมูลที่ใช้มีความสอดคล้องกับโมเดลอย่างสมบูรณ์ (perfect fit) ซึ่งเป็นสิ่งที่เกิดขึ้นได้ค่อนข้างยากในสภาพความเป็นจริง

การแก้ไขปัญหา non-positively definite ของเมตริกซ์ด้วยการปรับเรียบข้อมูลและการแก้ค่าการเดาทำให้ correlation ของข้อสอบแต่ละคู่มีค่าเปลี่ยนไปจากเดิมค่อนข้างมากทำให้ไม่สามารถแน่ใจได้ว่าค่าดัชนีที่คำนวณได้เป็นค่าดัชนีของคะแนนที่ได้จากแบบสอบเดิมที่ต้องการตรวจสอบความเป็นเอกมิติหรือไม่

เมื่อ tetrachoric correlation matrix ซึ่งใช้ข้อมูลมีค่าเป็น 0 และ 1 ยังคงเป็นปัญหา ทำให้การใช้ variance-covariance matrix มีความเหมาะสมมากกว่า เพราะค่าที่ได้เป็นค่าของข้อมูลจริงไม่มีการแก้ไข และให้ผลที่น่าพอใจในด้านความคงที่ที่ได้ ซึ่งได้ผลตรงกับงานวิจัยของ Roznowski (1991) ที่ศึกษาด้วยการ simulation พบว่า variance-covariance matrix ให้ผลคงที่มากกว่า phi correlation matrix และ tetrachoric correlation matrix ในการวิเคราะห์องค์ประกอบ นอกจากนี้การใช้ variance-covariance matrix ยังไม่มีปัญหาเรื่องการเป็น non-positively definite ของเมตริกซ์

2. การใช้ ดัชนี NNFI ตรวจสอบความเป็นเอกมิติของแบบสอบ แม้ว่าดัชนี NNFI จะมีความคงที่ เมื่อมีการเปลี่ยนแปลงจำนวนผู้สอบ จำนวนข้อสอบ และค่าความยาก แต่การคำนวณดัชนี NNFI จะต้องคำนึงถึงความเหมาะสมของจำนวนผู้สอบและข้อสอบ โดยควรจะต้องคำนึงถึงความครอบคลุมของเนื้อหาควบคู่กับการกำหนดจำนวนข้อสอบ

3. ดัชนี ERR ที่ได้จากการวิเคราะห์องค์ประกอบแบบ EFA ยังคงสามารถใช้ในการตรวจสอบความเป็นเอกมิติได้ เพราะมีความไวต่อการเจือปนข้อสอบในมิติอื่น แม้จะมีปัญหาเกี่ยวกับความคงที่ ซึ่งหากพิจารณาดูจะพบว่า ความไม่คงที่เกิดขึ้นเมื่อมีการเปลี่ยนแปลงจำนวนผู้สอบแต่ไม่เกิดขึ้นเมื่อมีการเปลี่ยนแปลงจำนวนข้อสอบ การจะนำไปใช้จึงมีแนวทางแก้ไขปัญหานี้ได้ โดยการเพิ่มจำนวนตัวอย่างที่ใช้ในการวิเคราะห์ และควรใช้ไม่ต่ำกว่า 20 เท่าของจำนวนข้อสอบ

นอกจากนี้ การใช้ดัชนี ERR จะให้ผลดีกว่า ER ในเรื่องของความไวต่อการเจือปน ทั้งนี้เพราะในขณะที่มีการเจือปนข้อสอบในมิติอื่นเข้าไป มีผลให้ค่า eigen ที่ 1 มีค่าลดลงแต่ในขณะเดียวกันค่า eigen ที่ 2 และ 3 ก็มีการเปลี่ยนแปลงค่าด้วยเช่นกัน ซึ่งมีผลต่อการคำนวณค่าดัชนี ดังแสดงในตารางที่ 5.1

ตารางที่ 5.1 เปรียบเทียบความไวของดัชนี ER และ ERR

	จำนวนข้อสอบในมิติอื่นที่เจือปน					
	0	1	2	3	4	5
E1	2.44	2.43	2.23	2.16	2.07	1.99
E2	0.44	0.43	0.44	0.49	0.46	0.47
E3	0.35	0.30	0.33	0.35	0.26	0.26
ER=E1/E2	5.60	5.67	5.07	4.36	4.53	4.21
E2/E3	1.23	1.42	1.34	1.40	1.78	1.82
ERR=E1/E2/E2/E3	4.55	3.99	3.78	3.11	2.54	2.34

จากตารางจะเห็นได้ว่าเมื่อมีการเจือปนข้อสอบในมิติอื่นค่าไอเกนที่ 2 จะมีค่าเพิ่มขึ้น ส่วนค่าไอเกนที่ 3 จะมีค่าลดลง การใช้ค่า eigen ที่ 3 มาร่วมในการคำนวณค่าดัชนี ทำให้เห็นผลการเปลี่ยนแปลงที่ชัดเจนมากขึ้น ดังนั้น การใช้ดัชนีจากการวิเคราะห์องค์ประกอบแบบ EFA ตรวจสอบความเป็นเอกมิติของแบบสอบจึงควรใช้ ดัชนี ERR มากกว่า ดัชนี ER ในการตรวจสอบการเจือปนของข้อสอบในมิติอื่น แม้ว่า ERR มีความคงที่น้อยกว่าดัชนี ER

4. การใช้ค่า χ^2 จากโปรแกรม TESTFACT (G^2) เป็นวิธีการวิเคราะห์องค์ประกอบแบบ CFA ที่เรียกว่า Bock' full information factor analysis (Bock, Gibbon & Muraki,1985 อ้างใน Nandakumar,1994:17) ให้ผลเช่นเดียวกับ χ^2 ที่พัฒนาโดย Joreskog & Sorbom ในโปรแกรม LISREL คือมีค่าไม่คงที่เมื่อมีการเปลี่ยนแปลงจำนวนผู้สอบ จำนวนข้อสอบ และ ค่าความยาก ทำให้ไม่สามารถใช้ในการตรวจสอบความเป็นเอกมิติของแบบสอบ ส่วน CN ซึ่งเป็นค่า χ^2 ที่มีการแก้ไขอิทธิพลของกลุ่มตัวอย่าง แม้จะมีค่าคงที่มากขึ้นแต่ก็ให้ค่าผิดทิศทางเมื่อมีการเจือปนข้อสอบในมิติอื่น

5. การที่ดัชนีหลายตัวประกอบด้วย G^2, χ^2, AGFI, RMR, และ CN ไม่สามารถใช้ตรวจสอบความเป็นเอกมิติของแบบสอบไม่ได้หมายความว่าว่าดัชนีเหล่านี้ไม่มีคุณภาพ เพียงแต่ดัชนีดังกล่าวถูกพัฒนาขึ้นมา เพื่อใช้ตรวจสอบความสอดคล้องระหว่างข้อมูลกับโมเดลที่กำหนดใน covariance structure model ส่วนแบบสอบเป็นเพียงชุดของตัวบ่งชี้ (set of indicators) ของตัวแปรแฝง (latent variables) เท่านั้น จึงเป็นเพียง measurement model ซึ่งเป็นส่วนหนึ่งของ covariance structure model เท่านั้น

6. จากการวิจัยครั้งนี้ พบว่าความเป็นเอกมิติของแบบสอบมีผลต่อการประมาณค่าพารามิเตอร์ตามแนว IRT ด้วยโปรแกรม BILOG เท่านั้น ส่วนในโปรแกรมอื่นเช่น LOGIST มีงานวิจัยของ Harrison (1986) ที่ศึกษาเกี่ยวกับการฝ่าฝืนข้อตกลงเบื้องต้นของ IRT เรื่องความเป็นเอกมิติของแบบสอบ ด้วยการ simulation พบว่าการที่แบบสอบมีหลายองค์ประกอบ (factors) และองค์ประกอบเหล่านั้นมีความสัมพันธ์ซึ่งกันและกัน ไม่มีผลต่อการประมาณค่าพารามิเตอร์ในโปรแกรม LOGIST

ข้อเสนอแนะในการนำไปใช้

1. การเลือกใช้เมตริกซ์ในการวิเคราะห์องค์ประกอบเพื่อคำนวณค่าดัชนีที่ใช้ในการตรวจสอบความเป็นเอกมิติของแบบสอบ เมื่อข้อมูลมีค่าเป็น 0 และ 1 ควรเลือกใช้เมตริกซ์ที่เหมาะสมในการวิเคราะห์ โดย

ใน EFA ดัชนี ER, ERR ควรใช้ variance-covariance matrix

ใน CFA ดัชนี χ^2,AGFI,NNFI และ CN ควรใช้ variance-covariance matrix ดัชนี G^2 ควรใช้ tetrachoric correlation matrix ที่มีการปรับเรียบข้อมูล ส่วนดัชนี RMR ควรใช้ tetrachoric correlation matrix แบบเดิม

2. การนำดัชนีไปใช้การตรวจสอบความเป็นเอกมิติของแบบสอบ ดัชนีที่น่าสนใจ คือ

2.1 ดัชนี **NNFI** จากการวิเคราะห์องค์ประกอบแบบ CFA ด้วย variance-covariance matrix เนื่องจาก NNFI มีคุณภาพทั้งในเรื่องของความคงที่และความไว ทำให้สามารถใช้ได้ในกรณีที่มีการเปลี่ยนแปลงจำนวนผู้สอบ จำนวนข้อสอบ และค่าความยาก แต่อย่างไรก็ตาม จำนวนผู้สอบที่ใช้ในการวิเคราะห์ควรใช้ไม่ต่ำกว่า 200 ตัวอย่าง ตามคำแนะนำของ Boomsma (1982 อ้างใน Bearden, Sharma & Teel, 1982 : 429) ที่ได้ศึกษาเกี่ยวกับขนาดตัวอย่างที่เหมาะสมในการวิเคราะห์องค์ประกอบด้วยโปรแกรม LISREL ส่วนค่าดัชนีที่ได้ในเบื้องต้นนี้ควรมีค่าตั้งแต่ 0.85 ขึ้นไป

2.2 ดัชนี **ERR** จากการวิเคราะห์องค์ประกอบแบบ EFA ด้วย variance-covariance matrix แม้ว่าดัชนี ERR มีลักษณะไม่คงที่ แต่ก็มีความไวในการตรวจสอบความเป็นเอกมิติ ความไม่คงที่จะเกิดขึ้นชัดเจนเมื่อเปลี่ยนจำนวนผู้สอบ แต่ไม่ชัดเจนเมื่อเปลี่ยนจำนวน หากต้องการตรวจสอบความเป็นเอกมิติของแบบสอบด้วยดัชนี ERR โดยให้มีผลที่ค่อนข้างคงที่จึงควรวิเคราะห์ด้วยจำนวนผู้สอบค่อนข้างมากและต้องไม่ต่ำกว่า 20 เท่าของจำนวนข้อสอบ (Lindeman Merenda & Gold, 1980 : 155) ค่าดัชนีที่ได้ในเบื้องต้นนี้ควรมีค่าตั้งแต่ 4.00 ขึ้นไป

ข้อเสนอแนะในการทำการวิจัยครั้งต่อไป

1. ควรมีการศึกษาโดยใช้ข้อสอบฉบับเดียวกันในการศึกษาถึงความคงที่และความไวของดัชนี ควบคู่ไปกับการศึกษาถึงปฏิสัมพันธ์ระหว่างข้อสอบและค่าความยากของข้อสอบที่มีผลต่อค่าดัชนีที่คำนวณได้ ซึ่งจะทำให้การสรุปถึงคุณภาพของดัชนีมีความชัดเจนมากขึ้น

2. ควรมีการศึกษาถึงผลของความเป็นเอกมิติของแบบสอบที่มีต่อการประมาณค่าพารามิเตอร์ตามแนว IRT ที่วิเคราะห์ด้วยโปรแกรมต่าง ๆ เช่น LOGIST MULTILOG ด้วยข้อสอบจริง และเปรียบเทียบค่าพารามิเตอร์นี้กับค่าพารามิเตอร์ที่ได้จากโปรแกรม BILOG นอกจากนี้ควรมีการเปรียบเทียบผลที่ได้เมื่อเจือปนข้อสอบในมิติอื่นเข้าไปในแบบสอบ

3. ควรมีการศึกษาถึงแบบสอบมีหลายองค์ประกอบ (factors) และองค์ประกอบเหล่านั้นมีความสัมพันธ์ซึ่งกันและกัน ว่าจะมีผลอย่างไรต่อการประมาณค่าพารามิเตอร์ในโปรแกรม BILOG

4. ควรมีการศึกษาค่าสูงสุดและค่าต่ำสุดที่เป็นไปได้ของดัชนีแต่ละชนิดเพื่อกำหนดค่าของดัชนีที่เหมาะสมในการใช้เป็นเกณฑ์ตัดสินความเป็นเอกมิติของแบบสอบ ตามข้อตกลงเบื้องต้นของ IRT

5. ควรมีการศึกษาเปรียบเทียบการเปลี่ยนแปลงค่าดัชนี ในการวิเคราะห์องค์ประกอบแบบ CFA เมื่อใช้กับข้อมูลที่มีลักษณะเป็น covariance structure model กับข้อมูล measurement model ซึ่งมี model ที่มีความซับซ้อนน้อยกว่า เพื่อเป็นแนวทางกำหนดเกณฑ์การใช้ค่าดัชนีเหล่านี้กับ model ที่มีความซับซ้อนในลักษณะต่าง ๆ

6. ควรมีการศึกษาเปรียบเทียบค่าดัชนีที่ได้จากแบบสอบในวิชาต่าง ๆ ว่าเนื้อหาวิชามีผลต่อค่าดัชนีอย่างไร

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย