มุมวิจัยและสถิติ

การทดสอบความไม่แปรเปลี่ยนของการวัดระหว่างกลุ่มผู้ถูกวัด ด้วยการวิเคราะห์โมเดลสมการโครงสร้าง

Test of Measurement Invariance across Measured Groups Using Structural Equation Model Analysis

นงลักษณ์ วิรัชชัย *

ในช่วงสองทศวรรษที่ผ่านมานี้ เทคโนโลยีการวิเคราะห์ข้อมูลมีการพัฒนาอย่างรวดเร็ว โดยเฉพาะด้านการวิเคราะห์โมเดลสมการโครงสร้าง (structural equation model: SEM) ซึ่งมีเทคนิคการวิเคราะห์สำหรับการเปรียบเทียบความแตกต่างของโมเดลสมการโครงสร้างระหว่างกลุ่มประชากร ที่รู้จักกันในชื่อกลยุทธ์กลุ่มพหุ (multiple group strategy) หรือการวิเคราะห์กลุ่มพหุ (multi-sample analysis) นักวิชาการหลายสาขาได้นำเทคนิคกลยุทธ์กลุ่มพหุไปใช้ประโยชน์ในการตรวจสอบ 'ความไม่แปรเปลี่ยนของการวัด (measurement invariance) หรือ สมมูล/ความเท่าเทียมของการวัด (measurement equivalence)' อย่างกว้างขวาง เช่น DeShon (2004); Kuhn และ Holling (2009); Makikangas, Feldt และ Kinnunen (2006); Widaman, Ferrer และ Conger (2010) ในสาขาวิชาจิตวิทยา Ellis, Aguirre-Urret, Sun และ Marakas (2008) ในสาขาวิชาการสื่อสารสารสนเทศ McQueen, Tiro และ Vernon (2008) ในสาขาวิชาการแพทย์และสาธารณสุข Tucker, Ozer, Lyubomirsky และ Boehm (2006) ในสาขาวิชาสังคมวิทยา นอกจากนี้ Vandenberg และ Lance (2000) ได้สังเคราะห์งานวิจัยเกี่ยวกับความไม่แปรเปลี่ยนของการวัดจำนวน 80 เรื่อง Schmitt และ Kuljanin (2008) ได้สังเคราะห์งานวิจัยเกี่ยวกับความไม่แปรเปลี่ยนของการวัดจำนวน 88 เรื่อง ในสาขาวิชาวิทยาการจัดการเพื่อให้ได้กระบวนการในการวิเคราะห์ความแตกต่างของโมเดลการวัดระหว่างกลุ่มประชากร หลักฐานงานวิจัยที่ยกตัวอย่างนี้เป็นเพียงส่วนหนึ่งของงานวิจัยจำนวนมาก ที่แสดงให้เห็นความสำคัญของการวิจัยด้านความเหมาะสมในการใช้มาตรวัด หรือแบบวัดกับกลุ่มผู้ถูกวัดหลายกลุ่ม หรือการทดสอบความไม่แปรเปลี่ยนของการวัด ซึ่งทำได้หลายวิธี การเสนอสาระในตอนนี้ให้ความสำคัญกับการทดสอบด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน ซึ่งต้องใช้การวิเคราะห์โมเดลสมการโครงสร้างประเภทเทคนิคการวิเคราะห์กลุ่มพหุ เนื่องจากบทความนี้มุ่งนำเสนอมโนทัศน์เบื้องต้นโดยไม่ลงลึกถึงรายละเอียดด้านเทคนิคการวิเคราะห์ หัวข้อการนำ

* ศาสตราภิชาน คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

เสนอจึงนำเสนอสาระสำคัญเพียง 3 หัวข้อ คือ วัตถุประสงค์ หลักการ และวิธีการวิเคราะห์ ดังรายละเอียดแต่ละหัวข้อต่อไปนี้

วัตถุประสงค์

การทดสอบความไม่แปรเปลี่ยนของการวัด มีวัตถุประสงค์สำคัญรวม 4 ข้อ คือ 1) เพื่อตรวจสอบว่ากลุ่มผู้ถูกวัดที่มีลักษณะทางสังคม/ประชากร/วัฒนธรรมแตกต่างกัน มีการตีความการวัดด้วยแบบวัดตรงกันมากน้อยเพียงใด 2) เพื่อตรวจสอบว่าผู้ถูกวัดที่มีลักษณะเหมือนกัน และได้คะแนนจากแบบวัดเท่ากัน มีแหล่งที่มาของความแตกต่างของคะแนนในแต่ละมิติตรงกันมากน้อยเพียงใด 3) เพื่อตรวจสอบว่าความแตกต่างด้านเพศ เชื้อชาติ และภูมิหลังอื่นๆ ของผู้ถูกวัด ทำให้เกิดความแตกต่างในการตอบแบบวัดมากน้อยเพียงใด 4) เพื่อตรวจสอบว่ากระบวนการแทรกแซงหรือการจัดกระทำในการทดลอง มีผลทำให้เกิดการเปลี่ยนแปลงโมเดลการวัดมากน้อยเพียงใด

นักวิจัยตระหนักถึงความสำคัญของการทดสอบความไม่แปรเปลี่ยนของการวัดมานานแล้ว ในอดีตเมื่อยังไม่มีการพัฒนาการวิเคราะห์โมเดลสมการโครงสร้าง นักวิจัยใช้การตรวจสอบทางจิตมิติ ได้แก่ การประมาณค่าความเที่ยงและความตรงของตัวแปร ความเที่ยงที่ใช้กันมากเป็นความเที่ยงแบบความสอดคล้องภายใน (internal consistency) ซึ่งเป็นที่รู้จักกันในชื่อ ค่าสัมประสิทธิ์แอลฟ่าของ cronbach (cronbach's alpha coefficient) ค่าความเที่ยงที่สูง มีส่วนช่วยยืนยันว่าองค์ประกอบสร้างขึ้นมาจากตัวบ่งชี้ที่มีความสัมพันธ์กันจริง ส่วนความตรงที่ใช้กันมาก คือ ความตรงเชิงเกณฑ์สัมพันธ์ (criterion-related validity) ความตรงเชิงองค์ประกอบหรือความตรงแฟคทอเรียล (factorial validity) ความตรงเชิงจำแนก (discriminant validity) ความตรงเชิงโครงสร้างตามวิธีพหุลักษณะพหุวิธี (multi-trait multi-method construct validity) ค่าความตรงที่ได้ช่วยสนับสนุนว่าตัวแปรในการวิจัยเป็นผลการวัดที่ตรงกับสิ่งที่จะวัดจริง เมื่อได้ตรวจสอบแล้วว่าผลการวัดมีความตรงและความเที่ยงในระดับสูง การวิเคราะห์ข้อมูลเพื่อตอบคำถามวิจัย ได้ผลการวิจัยตรงตามวัตถุประสงค์วิจัยที่ต้องการ นั่นคือในการวิจัยเชิงทดลอง เมื่อนักวิจัยวิเคราะห์ข้อมูลด้วย t-test หรือ ANOVA ได้ผลการวิเคราะห์ข้อมูลที่บอกถึงความแตกต่างของผลการวัดที่เกิดจากการจัดกระทำระหว่างกลุ่มประชากรจริง มิใช่ความแตกต่างที่เกิดจากลักษณะของประชากรที่ต่างกัน และในการวิจัยเชิงสหสัมพันธ์ เมื่อนักวิจัยวิเคราะห์ข้อมูลด้วยการวิเคราะห์ถดถอยพหุคูณ (MRA) ได้ค่าขนาดอิทธิพลของปัจจัยทางสังคม จิตวิทยา เศรษฐกิจ การศึกษา และประชากร ที่อธิบายความแตกต่างในตัวแปรตามจริง มิใช่อิทธิพลที่เกิดจากลักษณะของประชากรที่ต่างกัน (Vandenberg & Lance, 2000) ความตรงประเภทนี้เรียกว่าเป็นความตรงของการวิจัยด้านการสรุปทางสถิติ (statistical conclusion validity) (Kirk, 1995) ด้วยเหตุนี้สถิติวิเคราะห์ประเภท ANOVA และ MRA รวมทั้งสถิติวิเคราะห์ขั้นสูงที่ต่อยอดจาก ANOVA

และ MRA จึงมีข้อตกลงเบื้องต้นว่าตัวแปรทุกตัวต้องมีความเที่ยงและความตรง (Kerlinger & Lee, 2000; นงลักษณ์ วิรัชชัย, 2542)

การทดสอบความไม่แปรเปลี่ยนของการวัด อาจทำได้หลายวิธีแตกต่างกัน วิธีการสำคัญที่ใช้กันในระยะแรกเป็นวิธีการทางด้านการวัดได้แก่ ทฤษฎีการตอบสนองรายข้อ (item response theory: IRT) และทฤษฎีการสรุปนัยทั่วไป (genralizability theory) ซึ่งใช้ได้ดีในกรณีที่ตัวแปรที่ต้องการวัดเป็นตัวแปรเอกมิติ (uni-dimensional scale) ส่วนกรณีที่ตัวแปรที่ต้องการวัดเป็นตัวแปรพหุมิติ (multi-dimensional scale) ในระยะแรกไม่สามารถทำได้ (Vandenberg & Lance, 2000)

เมื่อมีการพัฒนาการวิเคราะห์โมเดลสมการโครงสร้าง ซึ่งสามารถใช้วิเคราะห์ข้อมูลที่เป็นตัวแปรพหุมิติได้ การทดสอบความไม่แปรเปลี่ยนของการวัดในระยะหลังส่วนใหญ่จึงใช้การวิเคราะห์โมเดลสมการโครงสร้าง เพราะทำให้นักวิจัยมีความเข้าใจที่ชัดเจนเรื่องความไม่แปรเปลี่ยนของการวัด (MI) มากขึ้น และเห็นว่าข้อตกลงเบื้องต้นเกี่ยวกับความเที่ยงและความตรงของผลการวัด ยังไม่สามารถประกันได้ว่าผลการวิจัยจะมีความตรงด้านการสรุปเชิงสถิติ และได้เพิ่มข้อตกลงเบื้องต้นที่เกี่ยวข้องกับความไม่แปรเปลี่ยนของการวัด (MI) หรือความไม่แปรเปลี่ยนของโมเดลการวัด (invariance of measurement model) ระหว่างกลุ่มประชากร ที่ควรต้องทดสอบก่อนที่จะวิเคราะห์ข้อมูลเพื่อตอบคำถามวิจัยอีก 3 ประการ (Bollen, 1989; Vandenberg & Lance, 2000) ดังนี้

ก. ข้อตกลงเบื้องต้นด้านความเท่าเทียมหรือสมมูลของคะแนนจริง (true score equivalence) หรือคะแนนองค์ประกอบ หรือคะแนนตัวแปรแฝงในโมเดลการวัดระหว่างกลุ่มประชากร

ข. ข้อตกลงเบื้องต้นด้านความเท่าเทียมหรือสมมูลของน้ำหนักองค์ประกอบ (factor loading equivalence) ในโมเดลการวัดระหว่างกลุ่มประชากร

ค. ข้อตกลงเบื้องต้นด้านความเท่าเทียมหรือสมมูลของความคลาดเคลื่อนหรือองค์ประกอบเฉพาะ (unique factor equivalence) และ/หรือความแปรปรวนของความคลาดเคลื่อนในโมเดลการวัดระหว่างกลุ่มประชากร

กล่าวโดยสรุป การตรวจสอบความไม่แปรเปลี่ยนของการวัดจึงมีความสำคัญ และเป็นสิ่งที่นักวิจัยเชิงปริมาณทั้งในการวิจัยเชิงทดลองและเชิงสหสัมพันธ์ ต้องตรวจสอบก่อนว่า ผลการวัดมีคุณสมบัติไม่แปรเปลี่ยน (invariance) หรือมีความเท่าเทียมกันหรือสมมูล (equivalence) ในทุกกลุ่มประชากร จึงจะสามารถดำเนินการวิเคราะห์เพื่อตอบคำถามวิจัยให้ได้ผลการวิจัยตรงตามวัตถุประสงค์วิจัยของตนต่อไป

หลักการ

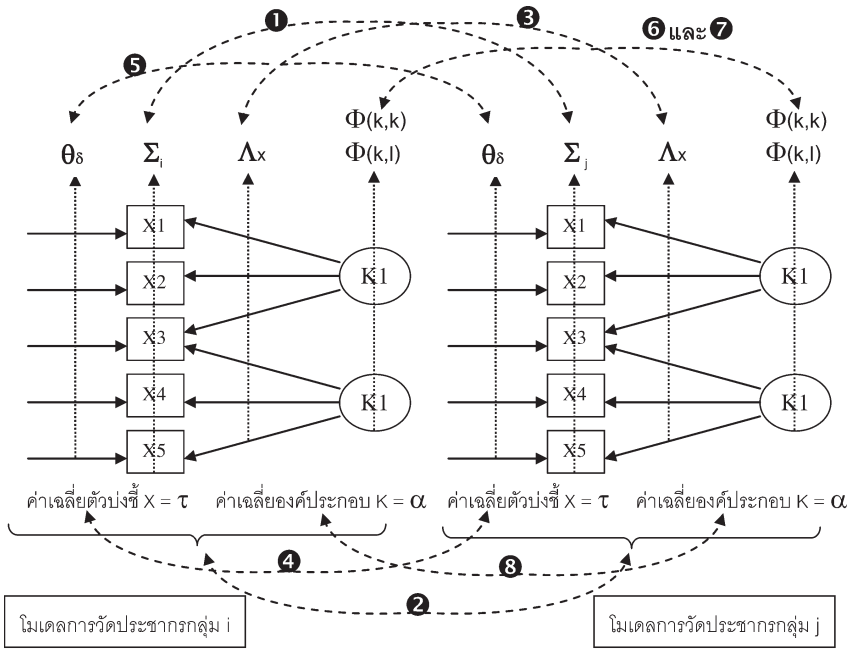
การวัดหรือคะแนนจากแบบวัดฉบับหนึ่งมีคุณสมบัติไม่แปรเปลี่ยนก็ต่อเมื่อผู้ถูกวัดที่มีคุณสมบัติที่ต้องการวัดตรงกัน แม้ว่าจะมาจากประชากรต่างกลุ่ม ควรต้องได้รับคะแนนการวัดจากแบบวัดนั้นตรงกัน ในกรณีที่ผู้ถูกวัดที่มีคุณสมบัติที่ต้องการวัดตรงกัน แต่เมื่อมา

จากประชากรต่างกลุ่ม กลับได้รับคะแนนการวัดจากแบบวัดนั้นแตกต่างกัน แสดงว่าการวัดหรือคะแนนจากแบบวัดฉบับหนึ่งมีคุณสมบัติแปรเปลี่ยน Vandenberg และ Lance (2000) จัดประเภทของความไม่แปรเปลี่ยนของการวัดตามลักษณะการทดสอบชุดสมมุติฐานในการวิเคราะห์โมเดลสมการโครงสร้างไว้หลายแบบ Schmitt และ Kuljanin (2008) นำมาจัดหมวดหมู่ตามลักษณะการทดสอบชุดสมมุติฐานในการวิเคราะห์โมเดลสมการโครงสร้างได้ 7 แบบ แต่ละแบบแสดงสมมุติฐานหลักทางสถิติเฉพาะกรณีประชากรสองกลุ่มเท่านั้น (หากมีมากกว่าสองกลุ่มเพิ่มค่าพาราเตอร์ตามจำนวนกลุ่มประชากร) ดังสาระและภาพโมเดลการวัดระหว่างประชากรสองกลุ่มต่อไปนี้

1. การทดสอบ**ความไม่แปรเปลี่ยนด้านเมทริกซ์ความแปรปรวน-ความแปรปรวนร่วม** (variance-covariance matrix) ระหว่างกลุ่มประชากร (1) สมมุติฐานหลักทางสถิติที่ทดสอบ คือ $H_o: \Sigma_i = \Sigma_j$ ถ้าไม่ปฏิเสธสมมุติฐาน สรุปได้ว่าการวัดไม่แปรเปลี่ยนระหว่างกลุ่มประชากร

2. การทดสอบ**ความไม่แปรเปลี่ยนด้านโครงแบบ** (configural invariance: COI) เป็นการทดสอบว่ารูปแบบของโมเดล (model form) ไม่แปรเปลี่ยนระหว่างกลุ่มประชากร (2) สมมุติฐานหลักทางสถิติ คือ $H_o: Form_i = Form_j$ แต่สมมุติฐานข้อ 2 ในชุดสมมุติฐาน เป็นสมมุติฐานรวมสมมุติฐานข้อ 1 ด้วย คือ Ho: $(\Sigma, Form)_i = (\Sigma, Form)_j$ ดังนั้นถ้าไม่ปฏิเสธสมมุติฐาน แสดงว่าไม่ปฏิเสธสมมุติฐานทั้งชุดคือทั้งข้อ 1 และ ข้อ 2 จึงสรุปได้ว่าโมเดลการวัดไม่แปรเปลี่ยนทั้งด้านเมทริกซ์ ความแปรปรวน-ความแปรปรวนร่วม และด้านรูปแบบ คือ โมเดลการวัด นอกจากจะมีเมทริกซ์ความแปรปรวน-ความแปรปรวนร่วมเหมือนกันแล้ว ยังมีโครงสร้างองค์ประกอบและตัวบ่งชี้เหมือนกัน แต่โมเดลอาจมีค่าน้ำหนักองค์ประกอบและค่าสถิติอื่นๆ แตกต่างกันระหว่างกลุ่มประชากรได้ และสรุปได้ว่าโมเดลการวัดมีความไม่แปรเปลี่ยนน้อยหรืออ่อน (weak invariance) ผลการทดสอบสมมุติฐานข้อ 2 นี้ เมื่อวิเคราะห์แยกทีละโมเดล (separate analysis) นอกจากจะได้โมเดลพื้นฐานสำหรับการตัดสินใจปฏิเสธ/ยอมรับสมมุติฐานทางสถิติในข้อต่อไปแล้ว ยังได้โมเดลการวัดสำหรับกลุ่มประชากรแต่ละกลุ่มด้วย

3. การทดสอบ**ความไม่แปรเปลี่ยนด้านเมตริก** (metric invariance: MEI) (3) เป็นการทดสอบว่าค่าน้ำหนักองค์ประกอบ (factor loadings) ของตัวบ่งชี้แต่ละตัวที่มีต่อองค์ประกอบในโมเดลการวัดไม่แปรเปลี่ยนระหว่างกลุ่มประชากร สมมุติฐานหลักทางสถิติที่ทดสอบ คือ $H_o: \Lambda x_i = \Lambda x_j$ และ $\Lambda y_i = \Lambda y_j$ แต่สมมุติฐานข้อ 3 ในชุดสมมุติฐานนี้ เป็นสมมุติฐานรวมสมมุติฐานข้อ 1-2 ด้วย คือ Ho: $(\Sigma, Form, \Lambda)_i = (\Sigma, Form, \Lambda)_j$ ดังนั้นถ้าไม่ปฏิเสธสมมุติฐานข้อ 3 แสดงว่าไม่ปฏิเสธสมมุติฐานทั้งชุดคือทั้งข้อ 1-3 จึงสรุปได้ว่าโมเดลการวัดไม่แปรเปลี่ยนระหว่างกลุ่มประชากรทั้งด้านเมทริกซ์ความแปรปรวน-ความแปรปรวนร่วม ด้านรูปแบบ และด้านน้ำหนักองค์ประกอบ หรือสรุปว่าโมเดลการวัดไม่แปรเปลี่ยนด้าน



รูปที่ 1 สมมุติฐานทางสถิติในการทดสอบความไม่แปรเปลี่ยนของโมเดลการวัดประชากร 2 กลุ่ม

เมตริก และถือว่าโมเดลการวัดมีความไม่แปรเปลี่ยนสูงหรือเข้ม (strong invariance)

4. การทดสอบ**ความไม่แปรเปลี่ยนด้านสเกลาร์** (scalar invariance: SCI) ❹ เป็นการทดสอบว่าค่าเฉลี่ยของตัวบ่งชี้ ขององค์ประกอบในโมเดลการวัดไม่แปรเปลี่ยนระหว่างกลุ่มประชากร เมื่อ τx, τy คือ เมทริกซ์ค่าเฉลี่ยตัวแปรสังเกตได้ภายนอก x และเมทริกซ์ค่าเฉลี่ยตัวแปรสังเกตได้ภายใน y ตามลำดับ สมมุติฐานหลักทางสถิติที่ทดสอบ คือ H_o: $\tau x_i = \tau x_j$ และ $\tau y_i = \tau y_j$ แต่สมมุติฐานข้อ 4 ในชุดสมมุติฐานนี้ เป็นสมมุติฐานรวมสมมุติฐานข้อ 1-3 ด้วย นั่นคือ H_o: $(\Sigma, \text{Form}, \Lambda, \tau)_i = (\Sigma, \text{Form}, \Lambda, \tau)_j$ ดังนั้นถ้าไม่ปฏิเสธสมมุติฐานข้อ 4 แสดงว่าไม่ปฏิเสธสมมุติฐานทั้งชุดคือทั้งข้อ 1-4 จึงสรุปได้ว่าโมเดลการวัดไม่แปรเปลี่ยนระหว่างกลุ่มประชากรทั้งด้านเมทริกซ์ความแปรปรวน-ความแปรปรวนร่วม ด้านรูปแบบ ด้านน้ำหนักองค์ประกอบ และด้านสเกลาร์ หรือสรุปว่าโมเดลการวัดไม่แปรเปลี่ยนด้านเมตริก

5. การทดสอบ**ความไม่แปรเปลี่ยนด้านความเฉพาะตัว** (uniqueness invariance or equal residuals: ER) ❺ เป็นการทดสอบความไม่แปรเปลี่ยนระหว่างกลุ่มประชากรที่เข้มงวด (strict invariance) เพราะมีโอกาสน้อยมากในการไม่ปฏิเสธสมมุติฐาน มีชื่อเรียกอีกชื่อหนึ่งในกรณีที่ความแปรปรวนขององค์ประกอบไม่แปรเปลี่ยนว่า **การทดสอบความไม่แปรเปลี่ยนด้านความเที่ยง** (reliability invariance) การทดสอบในขั้นตอนนี้ เป็นการทดสอบว่า เทอมความคลาดเคลื่อนของตัวบ่งชี้ในโมเดลการวัดไม่แปรเปลี่ยนระหว่างกลุ่มประชากร เมื่อ θ คือ เมทริกซ์ความแปรปรวน-

ความแปรปรวนร่วมของความคลาดเคลื่อน สมมุติฐานทางสถิติ คือ $H_o: \theta_i = \theta_j$ แต่สมมุติฐานข้อ 5 ในชุดสมมุติฐานนี้ เป็นสมมุติฐานรวมสมมุติฐานข้อ 1-4 ด้วย นั่นคือ Ho: $(\Sigma, \text{Form}, \Lambda, \tau, \theta)_i = (\Sigma, \text{Form}, \Lambda, \tau, \theta)_j$ ดังนั้นถ้าไม่ปฏิเสธสมมุติฐานข้อ 5 แสดงว่าไม่ปฏิเสธสมมุติฐานทั้งชุดคือทั้งข้อ 1-5 จึงสรุปได้ว่าโมเดลการวัดไม่แปรเปลี่ยนระหว่างกลุ่มประชากร ทั้งด้านเมทริกซ์ความแปรปรวน-ความแปรปรวนร่วม ด้านรูปแบบ น้ำหนักองค์ประกอบ สเกลาร์ และความคลาดเคลื่อน หรือสรุปว่าโมเดลการวัดไม่แปรเปลี่ยนด้านเมตริก

การทดสอบสมมุติฐานทั้ง 5 ข้อข้างต้นนี้ เป็นการทดสอบชุดสมมุติฐานสำหรับการทดสอบ**ความไม่แปรเปลี่ยนของการวัด** (measurement invariance) ระหว่างกลุ่มประชากร และเป็นสมมุติฐานสำหรับการตรวจสอบข้อตกลงเบื้องต้นที่เพิ่มขึ้นอีก 3 ประการที่กล่าวไว้ข้างต้น สำหรับการทดสอบสมมุติฐานที่เหลืออีก 3 ข้อ เป็นการทดสอบชุดสมมุติฐานเกี่ยวกับความสัมพันธ์ระหว่างองค์ประกอบ สำหรับการทดสอบ**ความไม่แปรเปลี่ยนเชิงโครงสร้าง** (structural invariance)

6. การทดสอบ**ความไม่แปรเปลี่ยนของความแปรปรวนในองค์ประกอบ** (invariance of factor variances)[6] เป็นการทดสอบว่าความแปรปรวนขององค์ประกอบในโมเดลการวัด ไม่แปรเปลี่ยนระหว่างกลุ่มประชากร สมมุติฐานหลักทางสถิติที่ทดสอบ คือ $H_o: \Phi(k,k)_i = \Phi(k,k)_j$ เมื่อ k เป็นรหัสแทนองค์ประกอบแต่ละองค์ประกอบ แต่สมมุติฐานข้อ 6 ในชุดสมมุติฐานนี้ เป็นสมมุติฐานรวมสมมุติฐานข้อ 1-5 ด้วย คือ Ho: $(\Sigma, \text{Form}, \Lambda, \tau, \theta, \Phi(k,k))_i = (\Sigma, \text{Form}, \Lambda, \tau, \theta, \Phi(k,k))_j$ ดังนั้นถ้าไม่ปฏิเสธสมมุติฐานข้อ 6 แสดงว่าไม่ปฏิเสธสมมุติฐานทั้งชุดคือทั้งข้อ 1-6 จึงสรุปได้ว่าโมเดลการวัดไม่แปรเปลี่ยนระหว่างกลุ่มประชากร ทั้งด้านเมทริกซ์ความแปรปรวน-ความแปรปรวนร่วม รูปแบบ น้ำหนักองค์ประกอบ ด้านสเกลาร์ ด้านความคลาดเคลื่อน และด้านความแปรปรวนขององค์ประกอบในโมเดลการวัด

7. การทดสอบ**ความไม่แปรเปลี่ยนของความแปรปรวนร่วมในองค์ประกอบ** (invariance of factor covariances)[7] เป็นการทดสอบว่าความแปรปรวนร่วมขององค์ประกอบในโมเดลการวัด ไม่แปรเปลี่ยนระหว่างกลุ่มประชากร สมมุติฐานหลักทางสถิติที่ทดสอบ คือ $H_o: \Phi(k,l)_i = \Phi(k,l)_j$ เมื่อ k, l เป็นรหัสแทนองค์ประกอบแต่ละองค์ประกอบ แต่สมมุติฐานข้อ 7 ในชุดสมมุติฐานนี้ เป็นสมมุติฐานรวมสมมุติฐานข้อ 1-6 ด้วย คือ $H_o: (\Sigma, \text{Form}, \Lambda, \tau, \theta, \Phi(k,k), \Phi(k,l))_i = (\Sigma, \text{Form}, \Lambda, \tau, \theta, \Phi(k,k), \Phi(k,l))_j$ ดังนั้นถ้าไม่ปฏิเสธสมมุติฐานข้อ 7 แสดงว่าไม่ปฏิเสธสมมุติฐานทั้งชุดคือทั้งข้อ 1-7 จึงสรุปได้ว่าโมเดลการวัดไม่แปรเปลี่ยนระหว่างกลุ่มประชากร ทั้งด้านเมทริกซ์ความแปรปรวน-ความแปรปรวนร่วม ด้านรูปแบบ ด้านน้ำหนักองค์ประกอบ ด้านสเกลาร์ ด้านความคลาดเคลื่อน ด้านความแปรปรวนและความแปรปรวนร่วมขององค์ประกอบในโมเดลการวัด

8. การทดสอบ**ความไม่แปรเปลี่ยนของค่าเฉลี่ยองค์ประกอบ** (invariance of factor means) เป็นการทดสอบว่าความแตกต่างของค่าเฉลี่ยองค์ประกอบระหว่างกลุ่มประชากร 8 ซึ่งเป็นวัตถุประสงค์สำคัญของการวิจัยเชิงทดลอง สมมุติฐานหลักทางสถิติที่ทดสอบ คือ H_o: $\alpha_i = \alpha_j$; $\kappa_i = \kappa_j$ เมื่อ α (alpha) และ κ (kappa) เป็นค่าเฉลี่ยของตัวแปรแฝงภายนอกและภายในตามลำดับ แต่สมมุติฐานข้อ 8 ในชุดสมมุติฐานนี้ เป็นสมมุติฐานรวมสมมุติฐานข้อ 1-7 ด้วย คือ H_o: (Σ, Form, Λ, τ, θ, Φ(k,k), Φ(k,l), α, κ)$_i$ = (Σ, Form, Λ, τ, θ, Φ(k,k), Φ(k,l), α, κ)$_j$ ดังนั้นถ้าไม่ปฏิเสธสมมุติฐานข้อ 8 แสดงว่าไม่ปฏิเสธสมมุติฐานทั้งชุดคือทั้งข้อ 1-8 และสามารถสรุปได้ว่าโมเดลการวัดไม่แปรเปลี่ยนระหว่างกลุ่มประชากร ทั้งด้านเมทริกซ์ ความแปรปรวน-ความแปรปรวนร่วม ด้านรูปแบบ ด้านน้ำหนักองค์ประกอบ ด้านสเกลาร์ ด้านความคลาดเคลื่อน ด้านความแปรปรวนและความแปรปรวนร่วมขององค์ประกอบ และด้านค่าเฉลี่ยองค์ประกอบในโมเดลการวัด

วิธีการวิเคราะห์ข้อมูลเพื่อทดสอบชุดสมมุติฐานทางสถิติทั้ง 8 ข้อนี้ ดำเนินการตามขั้นตอนการวิเคราะห์ข้อมูลโมเดลสมการโครงสร้าง กลยุทธ์กลุ่มพหุ (Joreskog & Sorbom, 1996; Jaccard & Wan, 1996; Bollen, 1989; นงลักษณ์ วิรัชชัย, 2542) ซึ่งมีขั้นตอนการวิเคราะห์ข้อมูล 2 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 การวิเคราะห์กลุ่มพหุ

การวิเคราะห์กลุ่มพหุ เป็นการวิเคราะห์ข้อมูลตามชุดสมมุติฐานทางสถิติทั้ง 8 ข้อ มีการดำเนินงานได้เป็น 2 แบบ ดังนี้

แบบที่ 1 การวิเคราะห์กลุ่มพหุเริ่มต้นจากการไม่กำหนดเงื่อนไขบังคับ

การวิเคราะห์แบบนี้ เป็นการวิเคราะห์ข้อมูล เริ่มต้นจากการประมาณค่าพารามิเตอร์ในโมเดลสมการโครงสร้างของแต่ละกลุ่มประชากรแยกกัน (separate analysis) ทำให้ได้โมเดลที่มีลักษณะแตกต่างกันจากนั้นจึงค่อย ๆ เพิ่มข้อกำหนดเงื่อนไขบังคับ (constraint) ให้พารามิเตอร์เท่ากันในการทดสอบ เป็นการทดสอบสมมุติฐานทางสถิติเรียงจากข้อ 1, 2, 3,... 7 ตามลำดับ โดยที่สมมุติฐานทางสถิติข้อ 1 มีข้อกำหนดเงื่อนไขบังคับน้อยที่สุด และค่อย ๆ เพิ่มมากขึ้นจนถึงข้อที่ 7 ซึ่งแสดงถึงความไม่แปรเปลี่ยนของโมเดลการวัดโดยสมบูรณ์ ผลการวิเคราะห์ด้วยโปรแกรมลิสเรลจะได้รายงานค่าดัชนีวัดระดับความสอดคล้อง (goodness of fit indices) ของการทดสอบสำหรับสมมุติฐานทุกข้อ แต่ละข้อประกอบด้วย ค่าดัชนีการวัดความสอดคล้องทุกกลุ่มประชากรในภาพรวม (overall goodness of fit) ซึ่งได้มาจากดัชนีวัดระดับความสอดคล้องจากกลุ่มประชากรแต่ละกลุ่มรวมกัน ถ้าพบว่าผลการวิเคราะห์ข้อมูลในสมมุติฐานข้อใดได้ค่าไค-สแควร์รวมไม่มีนัยสำคัญทางสถิติ แสดงว่าโมเดลแต่ละกลุ่มประชาการสอดคล้องกับข้อมูลเชิงประจักษ์ทุกกลุ่ม สรุปว่าโมเดลไม่แปรเปลี่ยนระหว่างประชากร ถ้าได้ค่าไค-สแควร์รวมมีนัยสำคัญ

ทางสถิติ แสดงว่ามีโมเดลของประชากรอย่างน้อยหนึ่งกลุ่มที่ไม่สอดคล้องกับข้อมูลเชิงประจักษ์ สรุปว่าโมเดลแปรเปลี่ยนระหว่างประชากร ให้เสนอผลการวิเคราะห์แสดงให้เห็นถึงความแตกต่างของพารามิเตอร์ในโมเดลระหว่างประชากรต่อไป

แบบที่ 2 การวิเคราะห์กลุ่มพหุเริ่มต้นจากการกำหนดเงื่อนไขบังคับ

การวิเคราะห์แบบนี้ เป็นการวิเคราะห์ข้อมูลเริ่มต้นจากการประมาณค่าพารามิเตอร์ในโมเดลสมการโครงสร้างของแต่ละกลุ่มประชากร โดยกำหนดเงื่อนไขบังคับว่า พารามิเตอร์ทั้งหมดในโมเดลของประชากรแต่ละกลุ่มเท่ากัน ทำให้ได้โมเดลที่มีลักษณะไม่แปรเปลี่ยน จากนั้นจึงค่อย ๆ ลดข้อกำหนดเงื่อนไขบังคับ (constraint) ให้พารามิเตอร์เท่ากันในการทดสอบ เป็นการทดสอบสมมุติฐานทางสถิติเรียงจากข้อ 8, 7, 6,... 1 ตามลำดับ โดยที่สมมุติฐานทางสถิติข้อ 8 มีข้อกำหนดเงื่อนไขบังคับมากที่สุด และค่อย ๆ เพิ่มมากขึ้นจนถึงข้อที่ 1 ซึ่งแสดงถึงความไม่แปรเปลี่ยนของโมเดลการวัดโดยสมบูรณ์ ผลการวิเคราะห์ด้วยโปรแกรมลิสเรลจะได้รายงานค่าดัชนีวัดระดับความสอดคล้อง (goodness of fit indices) ของการทดสอบสำหรับสมมุติฐานทุกข้อ แต่ละข้อประกอบด้วย ค่าดัชนีการวัดความสอดคล้องทุกกลุ่มประชากรในภาพรวม (overall goodness of fit) ซึ่งได้มาจากดัชนีวัดระดับความสอดคล้องจากกลุ่มประชากรแต่ละกลุ่มรวมกัน ถ้าพบว่าผลการวิเคราะห์ข้อมูลในสมมุติฐานข้อใดได้ค่าไค-สแควร์รวมไม่มีนัยสำคัญทางสถิติ แสดงว่าโมเดลแต่ละกลุ่มประชากรสอดคล้องกับข้อมูลเชิงประจักษ์ทุกกลุ่ม สรุปว่าโมเดลไม่แปรเปลี่ยนระหว่างประชากร ถ้าได้ค่าไค-สแควร์รวมมีนัยสำคัญทางสถิติ แสดงว่ามีโมเดลของประชากรอย่างน้อยหนึ่งกลุ่มที่ไม่สอดคล้องกับข้อมูลเชิงประจักษ์ สรุปว่าโมเดลแปรเปลี่ยนระหว่างประชากร ให้เสนอผลการวิเคราะห์แสดงให้เห็นถึงความแตกต่างของพารามิเตอร์ในโมเดลระหว่างประชากรต่อไป

ในการวิเคราะห์ข้อมูลนี้ หากผลการวิเคราะห์ในตอนเริ่มต้นพบว่าโมเดลไม่สอดคล้องกับข้อมูลเชิงประจักษ์ ต้องปรับโมเดลแล้ววิเคราะห์ใหม่ เพื่อให้ได้โมเดลปรับแก้ที่มีลักษณะโครงสร้างความสัมพันธ์ระหว่างตัวแปรสอดคล้องกับข้อมูลเชิงประจักษ์ การปรับแก้โมเดลนี้ทำได้โดยใช้ข้อเสนอแนะที่โปรแกรมรายงานในส่วนของดัชนีดัดแปลง (modification indices) หรือปรับแก้ตามข้อสังเกตของนักวิจัยบนพื้นฐานทางทฤษฎี เช่นเดียวกับการปรับแก้โมเดลในการวิเคราะห์โมเดลสมการโครงสร้างทั่วไป เมื่อโมเดลสอดคล้องกับข้อมูลเชิงประจักษ์แล้ว จึงดำเนินการวิเคราะห์สมมุติฐานข้อต่อไปตามลำดับ

ขั้นตอนที่ 2 การวิเคราะห์สรุป

การวิเคราะห์ในขั้นตอนนี้ เป็นการนำผลการวิเคราะห์กลุ่มพหุในการทดสอบสมมุติฐานทั้ง 8 ข้อ มาคำนวณหาผลต่างของดัชนีวัดระดับความสอดคล้อง ที่ได้จากการทดสอบสมมุติฐานในขั้นตอนที่ 1 โดยคำนวณหาผลต่างของดัชนีระหว่างคู่ที่มีเงื่อนไขบังคับ

น้อยกับมีเงื่อนไขบังคับมาก แล้วแปลความหมายสรุปผลการวิเคราะห์เกี่ยวกับโมเดลกลุ่มพหุทั้งหมด โดยเน้นการสรุปความหมายจากผลการทดสอบสมมุติฐานว่าโมเดลที่ไม่แปรเปลี่ยนมีลักษณะอย่างไร พารามิเตอร์ใดที่แปรเปลี่ยนและพารามิเตอร์ใดไม่แปรเปลี่ยนระหว่างกลุ่มประชากร ทั้งนี้อาจแปลความหมายเพิ่มเติมว่าความไม่แปรเปลี่ยนเกิดจากอิทธิพลของตัวแปรปรับ/ตัวแปรกำกับ (moderator) ตัวใดด้วย

ในการวิเคราะห์ความไม่แปรเปลี่ยนของการวัด นักวิจัยยังพลิกแพลงวิธีการทดสอบได้อีกหลายแบบ แบบที่หนึ่ง การข้าม/งดเว้นการทดสอบสมมุติฐานทางสถิติบางข้อ เช่น ข้อ 4 ความไม่แปรเปลี่ยนด้านสเกลาร์ ซึ่งนักวิจัยไม่สนใจก็ได้ แบบที่สอง การขยายการทดสอบจากระดับเมทริกซ์พารามิเตอร์ เป็นการทดสอบสมมุติฐานย่อยแยกเป็นระดับพารามิเตอร์ก็ได้ แบบที่สาม การวิเคราะห์รวมการทดสอบสมมุติฐานหลายข้อพร้อมกัน เช่น มีการทดสอบสมมุติฐานแยกเป็นรายข้อ แล้วมีการทดสอบสมมุติฐานรวมข้อ 1-5 เป็นการทดสอบรวม 'ความไม่แปรเปลี่ยนของการวัด' ทั้งชุดก็ได้ (Kuhn & Holling, 2009; Joreskog & Sorbom, 2001; Bollen, 1989; นงลักษณ์ วิรัชชัย, 2542)

ตัวอย่างงานวิจัยทดสอบความไม่แปรเปลี่ยนของการวัด

ในที่นี้ผู้เขียนขอยกตัวอย่างผลการวิเคราะห์สรุปจากงานวิจัยของ Tucker, Ozer, Lyubomirsky และ Boehm (2006) ซึ่งทำวิจัยเพื่อทดสอบความไม่แปรเปลี่ยนของการวัดด้วยมาตรความพึงพอใจต่อชีวิต (satisfaction with life scale: SWLS) ระหว่างกลุ่มประชากรชาวรัสเซีย และอเมริกัน โดยทดสอบสมมุติฐานทางสถิติเพียง 3 ข้อ คือ 1) การทดสอบความไม่แปรเปลี่ยนด้านโครงแบบ (configural invariance: COI) หรือการทดสอบโมเดลเส้นฐาน (baseline models) สำหรับการเปรียบเทียบ 2) การทดสอบความไม่แปรเปลี่ยนด้านเมตริก (metric invariance: MEI) หรือการทดสอบโมเดลความไม่แปรเปลี่ยนของการวัดแบบอ่อน (weak measurement invariance model) และ 3) การทดสอบความไม่แปรเปลี่ยนด้านค่าเฉลี่ยตัวบ่งชี้หรือด้านสเกลาร์ (scalar invariance: SCI) หรือการทดสอบโมเดลความไม่แปรเปลี่ยนของการวัดแบบเข้ม (strong measurement invariance model) ไม่มีการทดสอบสมมุติฐานทางสถิติตามข้อ 1, 5-8 เหตุผลในการไม่ทดสอบข้อ 1 เพราะนักวิจัยมีความเชื่อมั่นว่ารูปแบบ (form) ของโมเดลการวัดไม่แปรเปลี่ยนจึงไม่ทดสอบ ส่วนสมมุติฐานทางสถิติข้อ 5-8 ไม่มีการทดสอบ ด้วยเหตุผลสองประการ คือ ประการแรก สมมุติฐานข้อ 4 ได้ผลว่าปฏิเสธสมมุติฐานแล้วไม่จำเป็นต้องทดสอบต่อ ประการที่สอง นักวิจัยต้องการปล่อยให้ค่าสถิติอื่น ๆแปรเปลี่ยนได้ หรือแปรผันได้โดยอิสระ

ผลการวิเคราะห์สรุปได้ว่า การทดสอบโมเดลเส้นฐานตามสมมุติฐานข้อแรก และการทดสอบโมเดลความไม่แปรเปลี่ยนของการวัดแบบอ่อน พบว่าไม่ปฏิเสธสมมุติฐาน แต่การ

ทดสอบโมเดลความไม่แปรเปลี่ยนของการวัดแบบเข้ม พบว่าปฏิเสธสมมุติฐาน และเมื่อพิจารณาค่าผลต่างไค-สแควร์ พบว่าโมเดลที่ 1 และ 2 แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ แต่โมเดลที่ 2 และ 3 แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ซึ่งแสดงว่าผลการตรวจสอบความไม่แปรเปลี่ยนของการวัดในที่นี้ ได้โมเดลที่มีความไม่แปรเปลี่ยนของการวัดแบบอ่อน คือ ยังมีความแตกต่างระหว่างกลุ่มประชากรในด้านค่าเฉลี่ยตัวบ่งชี้ ด้านความคลาดเคลื่อน ด้านความแปรปรวนและความแปรปรวนร่วมขององค์ประกอบ และด้านค่าเฉลี่ยองค์ประกอบในโมเดลการวัด ข้อค้นพบนี้แสดงให้เห็นว่าการใช้มาตรวัด SWLS สำหรับกลุ่มประชากรที่มีความแตกต่างทางวัฒนธรรมยังมีปัญหาในเรื่องความไม่แปรเปลี่ยนของการวัด ดังผลการวิเคราะห์ในตารางที่ 1

ตารางที่ 1 ผลการวิเคราะห์สรุปสำหรับการทดสอบความไม่แปรเปลี่ยนของการวัด

Model	c^2	df	GFI	RMR	c^2 change	Df change
Baseline	20.10	10	0.98	0.09	-	
Weak MI	26.33	14	0.98	0.09	6.23	4
Strong MI	57.66	18	0.96	0.13	31.33*	4

หมายเหตุ ** p<.05*
ที่มา *: Tucker, Ozer, Lyubomirsky, & Boehm (2006)*

ผลการวิเคราะห์ข้างต้นสรุปได้ว่า การนำเครื่องมือวัดไปใช้กับประชากรต่างกลุ่ม ทำให้ได้โมเดลการวัดและผลการวัดแตกต่างกัน ดังนั้นเมื่อนักวิจัยนำเครื่องมือวัดที่สร้างขึ้นสำหรับประชากรในประเทศทางตะวันตกนี้ไปใช้กับประชากรในประเทศไทย ในการทดลองเปรียบเทียบค่าเฉลี่ยระหว่างกลุ่มที่ได้รับตัวแปรจัดกระทำต่างกัน จะเกิดปัญหาเพราะความแตกต่างในค่าเฉลี่ยผลการวัดเกิดจากทั้งตัวแปรจัดกระทำและจากความแปรเปลี่ยนของผลการวัด ซึ่งนักวิจัยต้องหามาตรการในการแก้ปัญหาต่อไป

สรุปสาระ

บทความนี้นำเสนอสาระเกี่ยวกับวัตถุประสงค์ หลักการ และวิธีการในการตรวจสอบความไม่แปรเปลี่ยนของการวัดด้วยการวิเคราะห์โมเดลสมการโครงสร้าง อันเป็นกระบวนการวิเคราะห์ข้อมูลที่มีความสำคัญ และนักวิจัยปัจจุบันนี้ยอมรับว่าควรต้องมีการวิเคราะห์ก่อนที่จะวิเคราะห์เพื่อตอบคำถามวิจัย หากผลการวิเคราะห์พบว่าเป็นโมเดลความไม่แปรเปลี่ยนของการวัดแบบอ่อน แสดงว่า ยังมีความแตกต่างระหว่างกลุ่มประชากรในด้านค่าเฉลี่ยตัวบ่งชี้ ความคลาดเคลื่อน ความแปรปรวนและความแปรปรวนร่วมขององค์ประกอบ และค่าเฉลี่ยองค์ประกอบในโมเดลการวัด ซึ่งเป็นการพิสูจน์ว่าเครื่องมือวัด/มาตรวัดที่เหมาะสมกับประชากรกลุ่มหนึ่ง ไม่เหมาะสมกับ

ประชากรอีกกลุ่มหนึ่งที่มีวัฒนธรรมต่างกัน

ข้อสรุปดังกล่าว มีประเด็นที่น่าสนใจ 2 ประการ ประการแรก ข้อสรุปจากการวิเคราะห์ความไม่แปรเปลี่ยนของการวัดแสดงให้เห็นว่าการใช้มาตรวัดจากต่างประเทศมาใช้วัดกลุ่มตัวอย่างในไทย อาจเกิดปัญหาความไม่แปรเปลี่ยนของการวัด และนักวิจัยควรต้องพิจารณาปรับปรุงเครื่องมือวัดให้เหมาะสมกับบริบทของไทยก่อนนำไปใช้ ประการที่สอง แม้ว่าจะมีการปรับปรุงมาตรวัดจากต่างประเทศให้เหมาะสมกับบริบทของไทยแล้ว แต่ในการวิจัยมีการเปรียบเทียบระหว่างกลุ่มตัวอย่างที่มีภูมิหลัง (เพศ อายุ การอบรมเลี้ยงดู ฯลฯ) ต่างกัน นักวิจัยควรต้องทดสอบความไม่แปรเปลี่ยนของการวัดระหว่างกลุ่มประชากรก่อนที่จะวิเคราะห์เพื่อตอบคำถามวิจัยด้วย เพราะความแตกต่างที่เกิดขึ้นในการเปรียบเทียบค่าเฉลี่ยระหว่างกลุ่มประชากรอาจจะมิใช่ความแตกต่างจากตัวแปรจัดกระทำ แต่เป็นความแตกต่างจากความแปรเปลี่ยนของโมเดลการวัดระหว่างกลุ่มประชากรได้

เอกสารอ้างอิง

นงลักษณ์ วิรัชชัย. (2542). *โมเดลลิสเรล: สถิติวิเคราะห์สำหรับการวิจัย*. กรุงเทพฯ : โรงพิมพ์ จุฬาลงกรณ์มหาวิทยาลัย.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York : John Wiley Sons.

DeShon, R. (2004). Measures are not invariant across groups without error variance homogeneity. In *Psychological Sciences*, 46(1), 137-149.

Ellis, M. E., Aguirre-Urreta, M. I., Sun, W. N., & Marakas, G. M. (2008). *Establishing the need for measuring invariance in information system research: A step-by step example using technology acceptance research. Decision sciences proceeding*. Retrieved May, 2010, from http://www.decision science.org/proceedings/DSI2008/docs/446_6332.pdf.

Jaccard, J., & Wan, C. K. (1996). *Lisrel approach to interaction effects in multiple regression*. Thousand Oaks: Sage Publication.

Joreskog, K. G., & Sorbom, D. (1996). *Lisrel 8: Structural equation modeling with the SIMPLIS command language*. Chicago: Scientific Software International.

Joreskog, K. G., & Sorbom, D. (2001). *Lisrel 8 user's reference guide*. Chicago: Scientific Software International.

Kerliger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research*. Orlando: Harcourt College Publishers.

Kirk, R.E. (1995). *Experimental design : procedures for the behavioral*

sciences. (3rd ed). CA : Brooks/ Cole.

Kuhn, J. T., & Holling, H. (2009). Measurement invariance of divergent thinking across gender, age, and school forms. *European Journal of Psychological Assessment*, 25(1), 1-7.

Makikangas, A., Feldt, T., & Kinnunen, U. (2006). The factor structure and factorial invariance of the 12-item general health questionnaire (GHQ-12) across time: Evidence from two community-based samples. *Psychological Assessment*, 18(4), 444-451.

McQueen, A., Tiro, J. A., & Vernon, S. W. (2008). Construct validity and invariance of four factors associated with colorectal cancer screening across gender, race and prior screening. *Cancer Epidemiol Biomarkers Prev*. 17(9), 2231-2237.

Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resources Management Review*. 18, 210-222.

Tucker, K. L., Ozer, D. J., Lyubomirsky, S., & Boehm, J.K. (2006). Testing for measurement invariance in the satisfaction with life scale: A comparison of Russians and North American. *Social Indicator Research*, 78, 341- 360.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestion, practices, and recommendations for organizational research. *Organizational Research Methods*. 3(1), 4-70.

Widaman, K. F., Ferrer, E., & Conger, R.D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives*. 4(1): 10-18.