

การแสดงผลหลักฐานความเที่ยงตรงของโครงสร้าง
องค์ประกอบอันดับหนึ่งและอันดับสูงของแบบ
ประเมินการสอนโดยผู้เรียนของมาร์ชฉบับภาษาไทย
The Validity Evidence of First order and Higher Order
Factor Structure of Thai Version of Marsh's Students'
Evaluation of Educational Quality

★ अनु เจริญวงศ์ระยับ¹

รองศาสตราจารย์นิภา ศรีไพโรจน์²

ผู้ช่วยศาสตราจารย์ ดร.องอาจ นัยพัฒน์²

บทคัดย่อ

แบบประเมินการสอนโดยผู้เรียนของมาร์ชเป็นเครื่องมือที่ใช้ประเมินการสอนของอาจารย์ระดับอุดมศึกษาอย่างแพร่หลาย เพราะเนื้อหาของแบบประเมินแสดงให้เห็นถึงกระบวนการจัดการเรียนการสอนได้อย่างครอบคลุม จากข้อค้นพบที่ผ่านมาแสดงให้เห็นว่าสารสนเทศการประเมินที่ได้จากผู้เรียนสามารถช่วยให้นักอาจารย์ในสถาบันอุดมศึกษาพัฒนาการจัดการเรียนการสอนให้ดีขึ้น อย่างไรก็ตามการนำแบบประเมินการสอนโดยผู้เรียนของมาร์ชมาใช้ในแต่ละบริบทของสังคมควรต้องมีการแสดงผลหลักฐานก่อนว่าแบบประเมินนั้นมีความเที่ยงตรงหรือไม่เมื่อนำมาใช้ในบริบทที่เปลี่ยนไป การแสดงคุณลักษณะที่ดีของแบบประเมินการสอนโดยผู้เรียนของมาร์ชได้มีการศึกษามาเป็นอย่างดีในหลายประเทศ อย่างไรก็ตามในวัฒนธรรมไทยยังไม่มีศึกษาเชิงประจักษ์ว่าแบบประเมินการสอนโดยผู้เรียนของมาร์ชสามารถนำมาใช้ได้อย่างมีความเที่ยงตรง วัตถุประสงค์ในการวิจัยครั้งนี้เป็นการแสดงผลหลักฐานความเที่ยงตรงเชิงโครงสร้างองค์ประกอบของแบบประเมินการสอน

¹ นิสิตปริญญาโท สาขาวิชาการวัดผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

² อาจารย์ประจำภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

โดยผู้เรียนของมาร์ช ทั้งองค์ประกอบอันดับหนึ่งจำนวน 9 องค์ประกอบ องค์ประกอบอันดับสองจำนวน 2 โมเดลคือ โมเดล 2 องค์ประกอบ และโมเดล 4 องค์ประกอบ และ องค์ประกอบอันดับสามโมเดล 1 องค์ประกอบ

ผู้ให้ข้อมูลในการศึกษาเป็นผู้เรียนระดับอุดมศึกษา จำนวน 6,432 คน จาก 2 มหาวิทยาลัย เป็นผู้เรียนใน มหาวิทยาลัยศรีนครินทรวิโรฒร้อยละ 59.81 และผู้เรียนใน มหาวิทยาลัยเกษตรศาสตร์ร้อยละ 40.19

ผลการวิจัยพบว่า องค์ประกอบอันดับหนึ่งโมเดล 9 องค์ประกอบ องค์ประกอบอันดับสองโมเดล 2 องค์ประกอบ และโมเดล 4 องค์ประกอบ และองค์ประกอบอันดับสามโมเดล ภาพรวมการประเมินการสอน มีความกลมกลืนกับข้อมูลเชิงประจักษ์ในระดับดี นอกจากนี้พบว่าสัมประสิทธิ์องค์ประกอบ และค่า R^2 ของแต่ละตัวบ่งชี้มีความเที่ยงตรงเชิงลู่เข้าใน องค์ประกอบแต่ละองค์ประกอบ ยกเว้นตัวบ่งชี้จำนวนชั่วโมง ต่อสัปดาห์ที่ศึกษานอกชั้นเรียนเท่านั้นที่มีความเที่ยงตรงเชิงลู่ เข้าในระดับต่ำ

The Validity Evidence of First order and Higher Order Factor Structure of Thai Version of Marsh's Students' Evaluation of Educational Quality

ABSTRACT

Marsh's Students' Evaluation of Educational Quality (SEEQ) is well known objective instrument to teacher evaluation practice in higher education institutions because of its multi-factor that represent the whole process of class room matters. Past findings showed that the student evaluation feedback information can help the faculty improve their teaching behaviors. Before using SEEQ in each culture, the researcher should show its validity evidence. SEEQ psychometric properties were well done in western countries and some parts of eastern countries. Anyway, in Thai culture, there is still no evidence to show that SEEQ can be used in Thai society. The purpose of this research was to show evidence of factorial validity of the Thai version of SEEQ, a tradition 9 first order factors model, two second order

2-factors and 4-factors models, and two third order 1-factor models of teacher evaluation in higher education model were proposed to test their construct validity.

The participants was 6,432 undergraduate students from two universities (59.81% Srinakharinwirot University, 40.19% Kasetsart University)

The results revealed that the first order factors model, the second order 2-factors model and 4-factors models, and the third order 1-factor models based on second order 2-factors and 4-factors models were fitted well with the empirical data. Factor loadings and R^2 of each indicator in each factor showed the convergent validity of each indicator with its factor but only hours per week studied outside class was not.

ความสำคัญและความเป็นมาของปัญหา

ภารกิจหลักของสถาบันอุดมศึกษาไทย ประกอบด้วย การจัดการเรียนการสอน การวิจัย การบริการ วิชาการแก่สังคมและการทำนุบำรุงศิลปวัฒนธรรม ภารกิจด้านการจัดการเรียนการสอน เป็นภารกิจหลักที่สำคัญ ประการหนึ่งและมีผลกระทบต่อสังคมไทยในวงกว้าง เนื่องจากการสอนในระดับอุดมศึกษามีจุดมุ่งหมายเพื่อ สร้างคนดี อยู่ในสังคมได้อย่างปกติสุข และเตรียมความพร้อม ในการประกอบอาชีพ การจัดการเรียนการสอนเพื่อ ผลิตบุคลากรเพื่อรับใช้สังคมจึงต้องกระทำอย่างเป็น กระบวนการ มีการวางแผนที่ดี มีการดำเนินการที่มี ประสิทธิภาพ และมีการประเมินเพื่อปรับปรุงแก้ไขจุดที่ บกพร่องเกี่ยวกับกระบวนการจัดการเรียนการสอน

การประเมินเป็นกระบวนการตัดสินคุณค่าจาก ข้อมูลที่รวบรวมมาอย่างเป็นระบบ โดยข้อมูลที่ใช้ในการ ตัดสินคุณค่าสามารถเป็นได้ทั้งข้อมูลที่ได้จากการวัดและ ข้อมูลที่ไม่ใช่การวัด (ศิริชัย กาญจนวาสี, 2543: 6) แหล่งข้อมูลในการประเมินการสอนควรมาจากหลายแหล่ง และมีวิธีการเก็บรวบรวมที่หลากหลาย เพื่อเป็นการยืนยัน

ความเที่ยงตรงของการประเมิน (Stake, 1989:14) แหล่งข้อมูลหนึ่งที่ได้รับคามนิยมในการประเมินการสอนของอาจารย์ คือ การประเมินการสอนโดยผู้เรียนเนื่องจากผู้เรียนเป็นลูกค้าโดยตรงของการสอนได้รับประสบการณ์การเรียนรู้โดยตรงจากอาจารย์ (Hativa, 2000: 336-337) นอกจากนี้การประเมินโดยผู้เรียนยังเป็นวิธีการที่สะดวกและจำเป็นต่อการพัฒนาคุณภาพการเรียนการสอนอีกด้วย (ศิริชัย กาญจนวาสี, 2543: 11)

วิธีการเก็บรวบรวมข้อมูลเพื่อการประเมินการสอนของอาจารย์โดยผู้เรียนที่นิยมใช้กันมากคือผู้เรียนให้ระดับคุณภาพการสอนของอาจารย์ วิธีการดังกล่าวคุณภาพของการวัดอยู่ที่เนื้อหาที่ใช้วัดว่าครอบคลุมพฤติกรรมการสอนที่พึงประสงค์ของอาจารย์มากน้อยเพียงใด (Marsh, 1987) ในทัศนะของมาร์ชเห็นว่าเนื้อหาการประเมินการสอนควรมีหลายมิติ เช่น อาจารย์ที่มีการเตรียมการสอนที่ดีอาจจะไม่ใช่อาจารย์ที่สามารถกระตุ้นผู้เรียนให้มีความสนใจในการเรียนก็ได้ การมีองค์ประกอบหลายมิติ จะให้สารสนเทศที่ตอบสนองวัตถุประสงค์ของการประเมินการสอนกับผู้ที่เกี่ยวข้องได้หลากหลายมากกว่า เช่น ผลการประเมินมิติต่าง ๆ ด้านกระบวนการจัดการเรียนการสอนสามารถนำผลการประเมินมาปรับปรุงพฤติกรรมการสอนของอาจารย์ ผลการประเมินสิ่งที่ผู้เรียนได้เรียนรู้จากรายวิชาสามารถนำผลการประเมินมาใช้ในการตัดสินใจเลือกเรียนรายวิชาของผู้เรียน ใช้ข้อมูลในการพิจารณาเลื่อนชั้นแก่อาจารย์ได้อย่างมีประสิทธิภาพ และจากผลการวิจัยเกี่ยวกับความเที่ยงตรงของแบบประเมินการสอนโดยผู้เรียนที่คูลิก (Kulik, 2001) ได้สังเคราะห์ความสัมพันธ์ระหว่างวิธีการประเมินการสอนโดยผู้เรียนกับวิธีการสอนโดยแหล่งข้อมูลอื่นๆ จากงานวิจัยที่ผ่านมาพบว่า ผลการประเมินการสอนโดยผู้เรียนมีความสัมพันธ์เชิงบวกกับเกณฑ์หลายๆ เกณฑ์ ได้แก่ ผลการเรียนและข้อสังเกตของผู้เรียน การประเมินจากผู้สังเกตที่ได้รับการฝึกมาอย่างดี การประเมินโดยศิษย์เก่า เป็นต้น ดังนั้นกล่าวได้ว่าสารสนเทศจากการประเมินการสอนโดยผู้เรียนที่มีคุณภาพและมีหลากหลายมิติเป็นตัวแทนที่ดีในการประเมินการสอน

ในปี ค.ศ. 1982 มาร์ช (Marsh, 1982) ได้ทำการพัฒนาแบบประเมินการสอนโดยผู้เรียน (Students' Evaluations of Educational Quality (SEEQ)) ลักษณะของแบบวัด เป็นมาตราประเมิน (Rating scale) 9 ระดับ ตั้งแต่ไม่

เห็นด้วยมากที่สุด (ให้ 1 คะแนน) ถึง เห็นด้วยมากที่สุด (ให้ 9 คะแนน) การทำแบบประเมินเป็นการให้นักศึกษาแต่ละคนเป็นผู้ให้คะแนนระดับคุณภาพการจัดการเรียนการสอนผู้สอนเป็นรายบุคคล ตามข้อคำถามที่กำหนดขึ้นซึ่งมี 35 ข้อ แบ่งเป็น 9 องค์ประกอบ มาร์ชได้ศึกษาเพื่อพัฒนาคุณภาพแบบประเมิน SEEQ โดยได้แสดงหลักฐานความเที่ยงตรงของแบบประเมิน SEEQ ไว้อย่างหลากหลาย ได้แก่ การศึกษาความเที่ยงตรงข้ามกลุ่ม (Marsh, & Hocevar, 1984) ความเที่ยงตรงเชิงโครงสร้างโดยการใช้การวิเคราะห์องค์ประกอบเชิงสำรวจ (Marsh, 1982) การวิเคราะห์องค์ประกอบเชิงยืนยัน (Marsh, & Hocevar, 1984) การวิเคราะห์คุณลักษณะหลากหลายวิธีหลาย (Marsh, 1982) และการศึกษาความเที่ยงตรงเชิงโครงสร้างกับผู้เรียนที่ใช้ภาษาอื่นที่ไม่ใช่ภาษาอังกฤษ เช่น ภาษาสเปน (Marsh, Touron, & Wheeler, 1985; citing Marsh, 1986) และภาษาจีน (Marsh, Hau, Chung, & Siu, 1998) เป็นต้น โดยผลการวิจัยทุกชิ้นยืนยันความเที่ยงตรงขององค์ประกอบของแบบประเมินการสอนทั้ง 9 องค์ประกอบเป็นอย่างดี แบบประเมินการสอน SEEQ จึงมีความน่าเชื่อถือในการให้สารสนเทศการประเมินการสอนแก่ผู้ที่มีส่วนเกี่ยวข้องกับการประเมินการสอนได้เป็นอย่างดี ซึ่งชี้ให้เห็นถึงคุณภาพและความเป็นสากลของแบบประเมิน SEEQ

จากการตรวจสอบเอกสารซึ่งพบว่าแบบประเมิน SEEQ มีเนื้อหาในการประเมินการสอนที่หลากหลายและมีการศึกษาถึงคุณภาพของแบบประเมินที่หลากหลายในประเทศ ดังนั้นผู้วิจัยจึงมีความสนใจศึกษาคุณภาพแบบประเมิน SEEQ ฉบับภาษาไทย โดยมีคำถามการวิจัยว่า เมื่อนำแบบประเมิน SEEQ มาใช้ในบริบทของสังคมไทยแล้ว องค์ประกอบการประเมินการสอนทั้ง 9 องค์ประกอบยังคงมีโครงสร้างองค์ประกอบเดิมอยู่ในเกณฑ์ที่ยอมรับได้หรือไม่ อย่างไรก็ตามการใช้ประโยชน์จากการประเมินการสอนควรให้สารสนเทศในหลาย ๆ มุมมอง และเหมาะสมกับผู้ต้องการใช้ข้อมูล สำหรับผู้ใช้ข้อมูลการประเมินการสอนเพื่อการตัดสินใจ เช่นฝ่ายบุคลากรของสถาบันอุดมศึกษา การใช้ข้อมูลสำหรับการประกัน

คุณภาพการศึกษา สารสนเทศที่ต้องการไม่จำเป็นต้องมีหลากหลายมิติเหมือนกับความต้องการใช้ข้อมูลของอาจารย์ผู้สอน ดังนั้นในงานวิจัยครั้งนี้จึงเสนอทางเลือกของการนำเสนอผลการประเมินการสอนที่มีจำนวนสารสนเทศที่น้อยกว่าจากฐานข้อมูลองค์ประกอบ 9 องค์ประกอบจำนวน เป็นโมเดลองค์ประกอบอันดับสองจำนวน 2 โมเดล ได้แก่ (1) โมเดลตามแนวคิดการประเมินการสอนของอุทุมพร จามรมาน (2530) ที่เสนอว่าสารสนเทศในการประเมินการสอนควรครอบคลุมเนื้อหาใน 2 มิติ คือ พฤติกรรมการสอน (process) และผลลัพธ์ที่ผู้เรียนได้จากการสอน (product) และ (2) โมเดลที่ปรับปรุงจากผลการวิจัยของมาร์ช (Marsh, 1991) ซึ่งมาร์ชพบว่าแบบประเมินการสอน SEEQ สามารถนำมาอธิบายใหม่เป็นโมเดล 4 องค์ประกอบ ได้แก่ ผู้นำเสนอ (presenter) ความสัมพันธ์ (rapport) เนื้อหา (material) และภาระงาน (workload) อย่างไรก็ตาม อนุ เจริญวงศ์ระยับ (2549) พบว่าการกำหนดคุณลักษณะเฉพาะโมเดล 4 องค์ประกอบของมาร์ชถึงแม้ว่าจะมีความสอดคล้องกับข้อมูลเชิงประจักษ์กับบริบทสังคมไทย แต่ผลการประมาณค่าพารามิเตอร์มีค่าไม่เหมาะสมเนื่องจากโมเดลของมาร์ชให้องค์ประกอบอันดับหนึ่งด้านการมอบหมายงาน เป็นตัวบ่งชี้ในองค์ประกอบอันดับสองถึง 2 องค์ประกอบ คือ เนื้อหาและภาระงาน ดังนั้นในการวิจัยครั้งนี้ผู้วิจัยจึงกำหนดคุณลักษณะเฉพาะใหม่ (Respecification) โดยให้องค์ประกอบอันดับหนึ่งการมอบหมายงาน (Assignments: AS) เป็นตัวบ่งชี้ขององค์ประกอบอันดับสอง ภาระงาน (Workload) เพียงองค์ประกอบเดียว และให้องค์ประกอบอันดับสองเนื้อหา (Material) มีตัวบ่งชี้เพียงตัวเดียวคือการทดสอบ/การให้เกรด (Examination/Grading: EG) หลังจากนั้นจะทำการตรวจสอบว่าโมเดลใดในสองโมเดลจะมีความสอดคล้องกับ

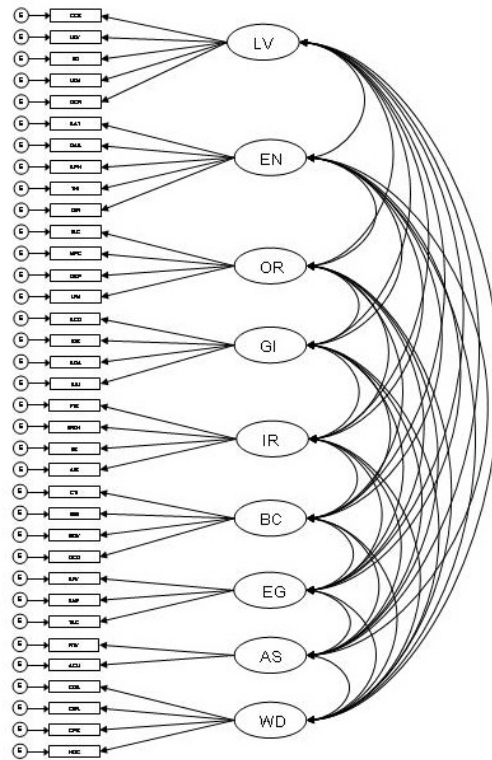
ข้อมูลเชิงประจักษ์มากกว่ากัน เมื่อได้โมเดลองค์ประกอบอันดับสองที่กลมกลืนกับข้อมูลเชิงประจักษ์แล้วจะทำการทดสอบว่าองค์ประกอบอันดับสองสามารถนำมารวมกันเป็นแนวคิดการประเมินการสอนในภาพรวมเป็นโมเดลองค์ประกอบอันดับสามโมเดล 1 องค์ประกอบได้หรือไม่

วัตถุประสงค์ของการวิจัย

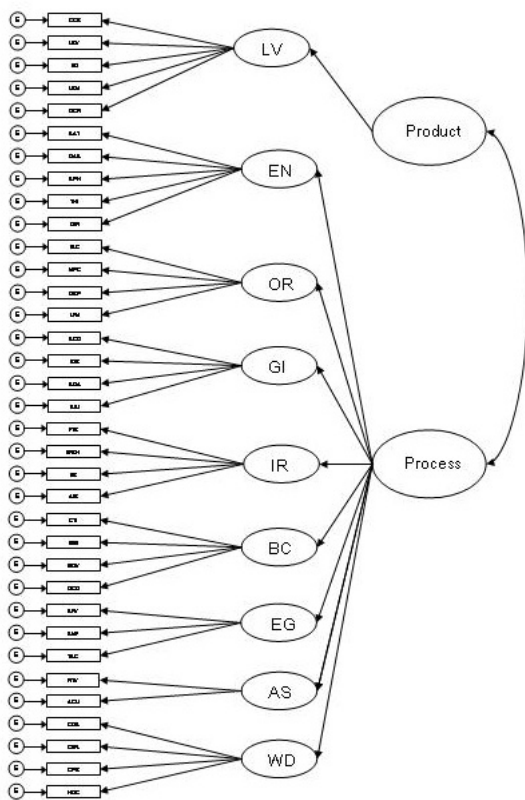
วัตถุประสงค์หลักของการวิจัยครั้งนี้เพื่อศึกษาความเที่ยงตรงของโครงสร้างองค์ประกอบของแบบประเมินการสอน SEEQ ฉบับภาษาไทยจำนวน 9 องค์ประกอบ และเสนอองค์ประกอบทางเลือกที่ได้จากพื้นฐานว่าโครงสร้างองค์ประกอบ 9 องค์ประกอบเป็นองค์ประกอบอันดับสอง และองค์ประกอบอันดับสาม โดยมีวัตถุประสงค์ย่อยดังนี้

1. เพื่อแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างองค์ประกอบอันดับหนึ่ง โมเดล 9 องค์ประกอบ
2. เพื่อแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างองค์ประกอบอันดับสอง โมเดล 2 องค์ประกอบ ได้แก่ กระบวนการ และผลลัพธ์ และโมเดล 4 องค์ประกอบ ได้แก่ ผู้นำเสนอ ความสัมพันธ์ เนื้อหา และภาระงาน
3. เพื่อแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างองค์ประกอบอันดับสาม โมเดล 1 องค์ประกอบ ภาพรวมของการประเมินการสอน
4. เพื่อแสดงเปรียบเทียบความสอดคล้องกับข้อมูลเชิงประจักษ์ของโมเดลองค์ประกอบอันดับสองระหว่างโมเดล 2 องค์ประกอบ ได้แก่ กระบวนการ และผลลัพธ์ และโมเดล 4 องค์ประกอบ ได้แก่ ผู้นำเสนอ ความสัมพันธ์ เนื้อหา และภาระงาน

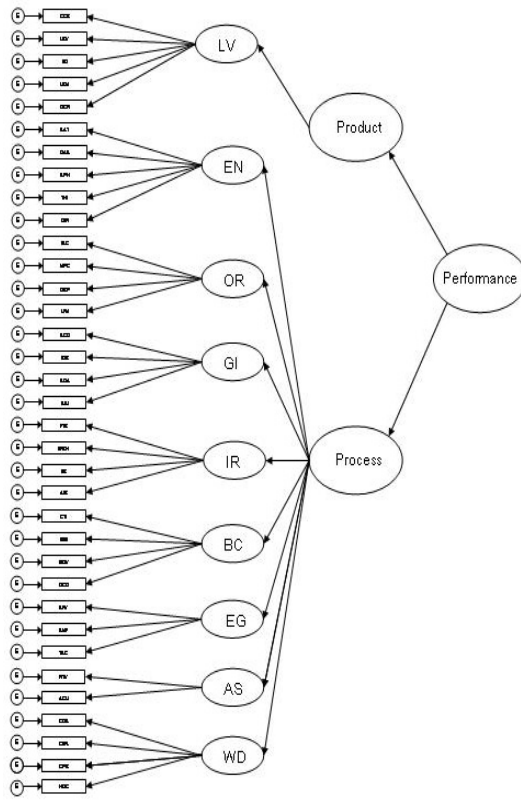
โมเดลสมมติฐานการวิจัย



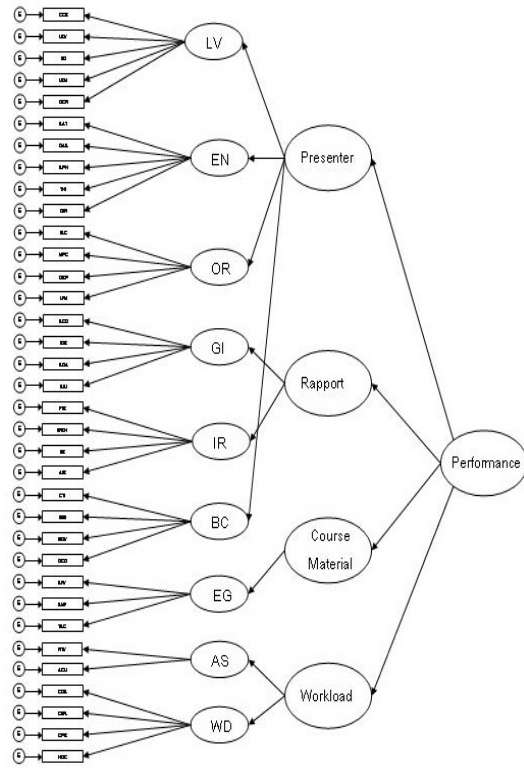
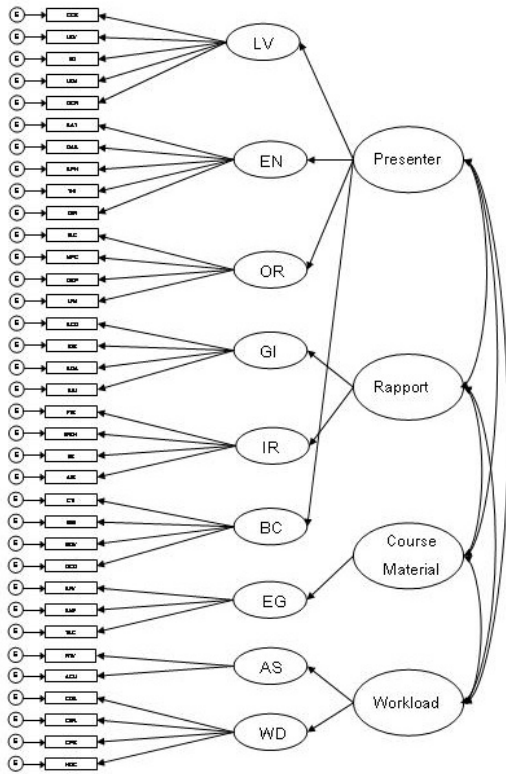
โมเดลองค์ประกอบอันดับหนึ่ง



โมเดลองค์ประกอบอันดับสอง: 2 องค์ประกอบ



โมเดลองค์ประกอบอันดับสาม: ฐาน 2 องค์ประกอบ



ภาพประกอบ 1 แสดงโมเดลสมมติฐานของการวิจัย

ระเบียบวิธีวิจัย

ประชากรในการวิจัย เป็นผู้เรียนระดับปริญญาตรี ภาคการศึกษาที่ 1 และ 2 ปีการศึกษา 2547 มหาวิทยาลัยเกษตรศาสตร์ และมหาวิทยาลัยศรีนครินทรวิโรฒ

กลุ่มตัวอย่างที่ใช้ในการวิจัย เป็นผู้เรียนระดับปริญญาตรี ภาคการศึกษาที่ 1 และ 2 ปีการศึกษา 2547 มหาวิทยาลัยเกษตรศาสตร์และมหาวิทยาลัยศรีนครินทรวิโรฒ จำนวน 6,432 คน (59.81% มหาวิทยาลัยศรีนครินทรวิโรฒ, 40.19% มหาวิทยาลัยเกษตรศาสตร์)

เครื่องมือที่ใช้ในการวิจัย เป็นแบบประเมินระดับคุณภาพการจัดการเรียนการสอน 9 ระดับ ตั้งแต่ไม่เห็นด้วยที่สุด (1) ถึง เห็นด้วยมากที่สุด (9) หากข้อความใดผู้เรียนไม่มีข้อมูล ให้ใส่ (0) จำนวนข้อความทั้งหมด 35 ข้อความ แบ่งเป็น 9 องค์ประกอบได้แก่ (1) การเรียนรู้/คุณค่า (Learning/Value: LV) (2) ความกระตือรือร้น (Enthusiasm: EN) (3) การเตรียมการสอน (Organization: OR) (4) ปฏิสัมพันธ์ภายในกลุ่ม (Group Interaction: GI) (5) ความเอาใจใส่ต่อผู้เรียน

(Individual Rapport: IR) (6) ความลุ่มลึกในเนื้อหาการสอน (Breadth of Coverage: BC) (7) การทดสอบ/การให้เกรด (Examination/Grading: EG) (8) การมอบหมายงาน (Assignments: AS) และ (9) ปริมาณงาน/ความยาก (Workload/Difficulty: WD)

ขั้นตอนการพัฒนาเครื่องมือ ผู้วิจัยทำการแปลแบบประเมิน SEEQ ฉบับภาษาอังกฤษเป็นภาษาไทย และทำการตรวจสอบความถูกต้องของการแปลโดยให้ผู้เชี่ยวชาญด้านภาษาอังกฤษจำนวน 2 ท่าน ต่อจากนั้นให้ผู้เชี่ยวชาญอีก 2 ท่านแปลแบบประเมินการสอน SEEQ ที่ผ่านการปรับปรุงแล้วเป็นภาษาอังกฤษอีกครั้ง แล้วนำข้อความภาษาอังกฤษฉบับดั้งเดิมและฉบับแปลมาเปรียบเทียบกัน ซึ่งพบว่าใจความของทั้งสองฉบับมีความใกล้เคียงกัน และยังคงความหมายเดิมไว้อย่างครบถ้วน ต่อจากนั้นผู้วิจัยนำแบบประเมินการสอน SEEQ มาทดลองใช้กับนิสิตจำนวน 10 คน เพื่อตรวจสอบความเข้าใจของข้อความ ผลการทดลองพบว่านิสิตทุกคนมีความเข้าใจข้อคำถามและสามารถประเมินการสอนของอาจารย์

ได้อย่างถูกต้องแต่มีการปรับภาษาที่ค่อนข้างยืดเยื้อเล็กน้อย และนำแบบประเมินการสอน SEEQ ที่ผ่านการปรับปรุงมาใช้ ในการตอบสมมติฐานการวิจัย

การเก็บรวบรวมข้อมูล ผู้วิจัยทำการติดต่อกับ คณะ/ภาควิชาที่เป็นกลุ่มตัวอย่าง โดยให้เจ้าหน้าที่ของคณะ/ ภาควิชา เป็นผู้เก็บรวบรวมข้อมูล โดยที่เจ้าหน้าที่ทำการแจก แบบประเมินให้กับนักศึกษาตอนท้ายภาคการศึกษาเป็น รายบุคคล นักศึกษาแต่ละคนทำการประเมินคุณภาพการ จัดการเรียนการสอนของอาจารย์แต่ละท่านในแต่ละรายวิชา ตามข้อคำถามที่กำหนดไว้จำนวน 35 ข้อ

การวิเคราะห์ข้อมูล ผลการวิเคราะห์ข้อมูลที่ หลักฐานแสดงความเที่ยงตรงเชิงโครงสร้างองค์ประกอบของ โมเดลสมมติฐานในการวิจัยครั้งนี้ประกอบด้วยหลักฐาน 3 กลุ่ม คือ

(1) การแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้าง ในภาพรวม ใช้การประเมินความสอดคล้องกลมกลืน ระหว่างโมเดลสมมติฐานกับข้อมูลเชิงประจักษ์ใช้การวิเคราะห์ องค์ประกอบเชิงยืนยัน ใช้วิธีการประมาณค่าแบบ Weighted Least Square ข้อมูลนำเข้าใช้ Polychoric Correlation และ Asymtotic Covariance Matrix ตัชนีและเกณฑ์ที่ใช้ในการ ประเมินความสอดคล้องกลมกลืนระหว่างโมเดลสมมติฐานกับ ข้อมูลเชิงประจักษ์ใช้ค่า Chi-square ที่ไม่มีนัยสำคัญทางสถิติ ที่ระดับ .05, GFI สูงกว่า .95, AGFI สูงกว่า .95, RMSEA พร้อมกับ ช่วงความเชื่อมั่น ต่ำกว่า .06, CFI สูงกว่า .95, NNFI สูงกว่า .95

(2) การแสดงหลักฐานความเที่ยงตรงเชิงผู้เข้า ใช้การประเมินค่าพารามิเตอร์ คือพิจารณาจากค่า สัมประสิทธิ์องค์ประกอบในรูปคะแนนดิบมีนัยสำคัญทาง สถิติที่ระดับ .05 ขนาดสัมประสิทธิ์องค์ประกอบในรูป คะแนนมาตรฐาน ของตัวตัวบ่งชี้ในแต่ละองค์ประกอบมี ค่ามากกว่า .40 แต่ไม่เกิน 1.0 และความคงที่ของค่า องค์ประกอบร่วม พิจารณาจากค่า R^2 เข้าใกล้ 1

(3) เปรียบเทียบความสอดคล้องกับข้อมูลเชิง ประจักษ์ระหว่างโมเดลคู่แข่ง ได้แก่ โมเดลองค์ประกอบ อันดับสอง 2 องค์ประกอบ กับโมเดลองค์ประกอบอันดับ สอง 4 องค์ประกอบ โดยการทดสอบความแตกต่างของ ความสอดคล้องกับข้อมูลเชิงประจักษ์ระหว่างโมเดล ใช้ค่า ความแตกต่างไคสแควร์ ($\Delta\chi^2$) และความแตกต่างของ ค่าองศาอิสระ (Δdf) ในการทดสอบความแตกต่าง โดย ค่าความแตกต่างไคสแควร์มีนัยสำคัญทางสถิติที่ระดับ .05 หมายถึง โมเดลที่เป็นคู่แข่งชั้นมีการสอดคล้องกับข้อมูล เชิงประจักษ์ไม่เท่าเทียมกัน

สรุปผลการวิจัย

ผลการวิจัยสามารถสรุปได้ดังนี้

1. ผลการแสดงหลักฐานความเที่ยงตรงเชิง โครงสร้างในภาพรวม เป็นการตรวจสอบโมเดลสมมติฐาน ว่ามีความสอดคล้องกับข้อมูลเชิงประจักษ์หรือไม่ ผลการวิจัยมีดังนี้

ตาราง 1 ผลการประเมินความสอดคล้องของโมเดลสมมติฐานกับข้อมูลเชิงประจักษ์

โมเดลสมมติฐาน	Chi-square	RMSEA	GFI	AGFI	CFI	NNFI
		(ช่วงความเชื่อมั่น)				
องค์ประกอบอันดับหนึ่ง	5,188.42 (df=524)	.037 (.036 ถึง .038)	.99	.98	.97	.97
องค์ประกอบอันดับสอง: 2 องค์ประกอบ	6,457.87 (df=560)	.040 (.040 ถึง .041)	.98	.98	.97	.96
องค์ประกอบอันดับสอง : 4 องค์ประกอบ	5,597.46 (df=544)	.038 (.037 ถึง .039)	.99	.98	.97	.97
องค์ประกอบอันดับสาม: ฐาน 2 องค์ประกอบ	6,457,87 (df=560)	.040 (.040 ถึง .041)	.98	.98	.97	.96
องค์ประกอบอันดับสาม: ฐาน 4 องค์ประกอบ	5,872.66 df=554)	.039 (.038 ถึง .040)	.99	.98	.97	.97

ผลการประเมินโมเดลสมมติฐานกับข้อมูลเชิงประจักษ์พบว่า โมเดลสมมติฐานที่เสนอทั้ง 5 โมเดลนอกจากค่าโคสแควร์ที่มีนัยสำคัญทางสถิติที่ระดับ .05 แล้ว ดัชนีความกลมกลืนอื่นๆ บ่งชี้ว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ทั้งหมด กล่าวคือ ค่า GFI, AGFI, CFI, และ NNFI ในทุกโมเดลมีค่ามากกว่า .95 และ ค่า RMSEA รวมทั้งช่วงความเชื่อมั่นต่ำกว่า .06 ทุกโมเดล และพบว่าองค์ประกอบอันดับสองโมเดล 4 องค์ประกอบที่ปรับจากแนวคิดของมาร์ชผลการประมาณค่าพารามิเตอร์มีความเหมาะสมทุกค่า แสดงถึงหลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบประเมิน ทั้งใน (1) โมเดลอันดับหนึ่ง 9 องค์ประกอบของแบบประเมินการสอนของมาร์ชมีความกลมกลืนกับกลุ่มตัวอย่างผู้เรียนไทย สามารถนำแบบประเมินการสอนซึ่งมีสารสนเทศเกี่ยวกับการสอน 9 มิติ มาใช้ในการประเมินการสอนในบริบทของสังคมไทยได้อย่างน่าเชื่อถือ (2) องค์ประกอบทั้ง 9 องค์ประกอบสามารถนำมารวมกันเป็นองค์ประกอบใหม่ตามแนวทางการประเมินการสอนได้ 2 องค์ประกอบ คือ องค์ประกอบด้านกระบวนการจัดการเรียนการสอน และผลลัพธ์ที่ผู้เรียนได้รับจากการสอนได้อย่างน่าเชื่อถือ (3) องค์ประกอบทั้ง 9 องค์ประกอบสามารถรวมกันเป็นองค์ประกอบใหม่เป็น 4 องค์ประกอบซึ่งปรับปรุงจากแนวคิดของมาร์ชได้อย่างน่าเชื่อถือเช่นกัน และ (4) องค์ประกอบการประเมินการสอนทั้งโมเดลสององค์ประกอบและโมเดลสี่องค์ประกอบสามารถนำมารวมกันเป็นองค์ประกอบรวมคือการประเมินการสอน (Performance Appraisals) ได้อย่างน่าเชื่อถือ

ผลการทดสอบความเที่ยงตรงเชิงโครงสร้างในระดับภาพรวมทั้งโมเดล พบว่าทุกโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ในระดับดี แสดงว่าโมเดลที่กำหนดขึ้นมีความเที่ยงตรงเชิงโครงสร้างในภาพรวม ในลำดับต่อไป เป็นการพิจารณาความเที่ยงตรงเชิงโครงสร้าง ของข้อคำถามรายข้อ โดยพิจารณาถึงความเที่ยงตรงเชิงผู้เข้าของข้อคำถามเป็นรายข้อว่ามีความแปรปรวนร่วมกันในการอธิบายความแปรปรวนขององค์ประกอบมากน้อยเพียงใด ซึ่งพิจารณาจากค่าพารามิเตอร์ต่างๆ ดังนี้

2. ผลการประเมินค่าพารามิเตอร์ พบว่าค่าพารามิเตอร์ทุกดัชนีอยู่ในเกณฑ์ที่แสดงว่าข้อคำถามเกือบ

ทุกข้อมีความเที่ยงตรงเชิงผู้เข้ากับองค์ประกอบของตนเอง ได้แก่ ค่าสัมประสิทธิ์ขององค์ประกอบในรูปคะแนนดิบทุกตัวมีนัยสำคัญทางสถิติที่ระดับ .05 ค่าสัมประสิทธิ์องค์ประกอบในรูปคะแนนมาตรฐานเกือบทุกค่า สูงกว่า .40 และค่า R^2 เข้าใกล้ 1 โดยรายละเอียดค่าสัมประสิทธิ์ขององค์ประกอบในรูปคะแนนมาตรฐานและค่า R^2 ในแต่ละโมเดลมีดังนี้

2.1 โมเดลองค์ประกอบอันดับหนึ่ง มีจำนวน 9 องค์ประกอบ มีค่าพารามิเตอร์ดังนี้ (1) องค์ประกอบการเรียนรู้/คุณค่ามีค่าสัมประสิทธิ์องค์ประกอบในรูปคะแนนมาตรฐานระหว่าง .84 ถึง .94 ค่า R^2 ระหว่าง .71 ถึง .88 (2) ความกระตือรือร้น .88 ถึง .97 ค่า R^2 .77 ถึง .95 (3) การเตรียมการสอน .92 ถึง .96 ค่า R^2 .84 ถึง .91 (4) ปฏิสัมพันธ์ภายในกลุ่ม .94 ถึง .96 ค่า R^2 .89 ถึง .93 (5) ความเอาใจใส่ต่อผู้เรียน .89 ถึง .97 ค่า R^2 .80 ถึง .93 (6) ความลุ่มลึกในเนื้อหาการสอน .94 ถึง .95 ค่า R^2 .89 ถึง .90 (7) การทดสอบ/การให้เกรด .95 ถึง .95 ค่า R^2 .90 ถึง .91 (8) การมอบหมายงาน .94 ถึง .95 ค่า R^2 .89 ถึง .90 และ (9) ปริมาณงาน/ความยาก .19 ถึง .76 ค่า R^2 .04 ถึง .58

2.2 โมเดลองค์ประกอบอันดับสอง: 2 องค์ประกอบ คือ (1) กระบวนการ ซึ่งประกอบด้วยตัวบ่งชี้ องค์ประกอบอันดับหนึ่งจำนวน 8 องค์ประกอบ คือ ความกระตือรือร้น การเตรียมการสอน ปฏิสัมพันธ์ภายในกลุ่ม ความเอาใจใส่ต่อผู้เรียน ความลุ่มลึกในเนื้อหาการสอน การทดสอบ/การให้เกรด การมอบหมายงาน และปริมาณงาน/ความยาก พบว่าค่าสัมประสิทธิ์องค์ประกอบอันดับสองในรูปคะแนนมาตรฐานอยู่ระหว่าง .43 ถึง 1.0 ค่า R^2 ระหว่าง .18 ถึง .99 และ (2) ผลลัพธ์ประกอบด้วยตัวบ่งชี้องค์ประกอบอันดับหนึ่งจำนวน 1 องค์ประกอบ คือ การเรียนรู้/คุณค่า เนื่องจากองค์ประกอบผลลัพธ์มีตัวบ่งชี้เพียง 1 ตัว จึงกำหนดค่าสัมประสิทธิ์องค์ประกอบอันดับหนึ่งเป็น 1 ค่าสัมประสิทธิ์องค์ประกอบเฉพาะของข้อความเป็น 0 และค่า R^2 เป็น 1 เช่นกัน.3 โมเดลองค์ประกอบอันดับสอง: 4 องค์ประกอบที่ปรับจากแนวคิด

ของมาร์ช คือ (1) ผู้นำเสนอ ซึ่งประกอบด้วยตัวบ่งชี้ องค์ประกอบอันดับหนึ่งจำนวน 3 องค์ประกอบ คือ การเรียนรู้/คุณค่า ความกระตือรือร้น การเตรียมการสอน และ ความลุ่มลึกในเนื้อหาการสอน พบว่าค่าสัมประสิทธิ์องค์ประกอบอันดับสองในรูปคะแนนมาตรฐานอยู่ระหว่าง .98 ถึง .99 ค่า R^2 ระหว่าง .96 ถึง .99 (2) ความสัมพันธ์ ซึ่งประกอบด้วยตัวบ่งชี้ องค์ประกอบอันดับหนึ่งจำนวน 2 องค์ประกอบ คือ ปฏิสัมพันธ์ภายในกลุ่ม และ ความเอาใจใส่ พบว่าค่าสัมประสิทธิ์องค์ประกอบอันดับสองในรูปคะแนนมาตรฐานเท่ากับ .96 และ .98 ตามลำดับ ค่า R^2 เท่ากับ .92 และ .97 ตามลำดับ (3) เนื้อหางาน ซึ่งประกอบด้วยตัวบ่งชี้องค์ประกอบอันดับหนึ่งเพียงองค์ประกอบเดียวคือ การทดสอบ/การให้เกรด จึงกำหนดค่าสัมประสิทธิ์องค์ประกอบอันดับหนึ่งและสัมประสิทธิ์องค์ประกอบอันดับสองเป็น 1 ค่าสัมประสิทธิ์องค์ประกอบเฉพาะของข้อความเป็น 0 และค่า R^2 เป็น 1 เช่นกัน และ (4) ภาระงาน ซึ่งประกอบด้วยตัวบ่งชี้องค์ประกอบอันดับหนึ่งจำนวน 2 องค์ประกอบ คือ การมอบหมายงาน และปริมาณงาน/ความยาก พบว่าค่าสัมประสิทธิ์องค์ประกอบอันดับสองในรูปคะแนนมาตรฐานเท่ากับ .99 และ .39 ตามลำดับ ค่า R^2 เท่ากับ .98 และ .15 ตามลำดับ

2.4 โมเดลองค์ประกอบอันดับสามองค์ประกอบรวมการประเมินการสอน มีจำนวน 2 โมเดล ได้แก่ (1) โมเดลองค์ประกอบอันดับสาม ซึ่งมาจากฐานองค์ประกอบอันดับสองจำนวน 2 องค์ประกอบ คือ กระบวนการ และผลลัพธ์ พบว่ากระบวนการค่าสัมประสิทธิ์องค์ประกอบอันดับสามในรูปคะแนนมาตรฐานเท่ากับ .97 ค่า R^2 เท่ากับ .94 และผลลัพธ์กำหนดให้ค่าสัมประสิทธิ์องค์ประกอบเป็น 1 และค่า R^2 เป็น 1 และ (2) โมเดลองค์ประกอบอันดับสาม ซึ่งมาจากฐานองค์ประกอบอันดับสองจำนวน 4 องค์ประกอบ คือ ผู้นำเสนอ ความสัมพันธ์ เนื้อหา และ ภาระงาน พบว่า ผู้นำเสนอ ความสัมพันธ์ และภาระงาน ค่าสัมประสิทธิ์องค์ประกอบอันดับสามในรูปคะแนนมาตรฐานเท่ากับ 1.0, .99, และ .99 ตามลำดับ ค่า R^2 เท่ากับ 1.0, .98, และ .99 ตามลำดับ สำหรับเนื้อหาสัมประสิทธิ์องค์ประกอบอันดับสามในรูปคะแนนมาตรฐานกำหนดให้เป็น 1 และค่า R^2 เป็น 1

3. การเปรียบเทียบความสอดคล้องกับข้อมูลเชิงประจักษ์ระหว่างโมเดลอันดับสอง: / องค์ประกอบ และ โมเดลอันดับสอง: 4 องค์ประกอบ เพื่อใช้เป็นเกณฑ์ในการตัดสินใจว่าโมเดลทางเลือกที่ผู้วิจัยเสนอขึ้น (โมเดล 2 องค์ประกอบ) เมื่อเปรียบเทียบกับโมเดลที่มาร์ช (Marsh, 1991) โมเดลใดมีความเที่ยงตรงเชิงโครงสร้างในภาพรวมในบริบทของสังคมไทยมากกว่ากัน ผลการวิจัยพบว่า ค่าไคสแควร์ในโมเดลองค์ประกอบอันดับสอง: 4 องค์ประกอบมีค่าน้อยกว่า โดยความแตกต่างต่างของค่าไคสแควร์เท่ากับ 860.41 และความแตกต่างของค่าองศาอิสระเท่ากับ 16 ความแตกต่างดังกล่าวมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่าองค์ประกอบอันดับสองโมเดล 4 องค์ประกอบที่ปรับจากโมเดลของมาร์ชมีความกลมกลืนกับข้อมูลเชิงประจักษ์ดีกว่าองค์ประกอบอันดับสองโมเดล 2 องค์ประกอบในบริบทของสังคมไทย อย่างไรก็ตามทั้งสองโมเดลต่างก็มีความกลมกลืนกับข้อมูลเชิงประจักษ์ในระดับดีเช่นเดียวกัน

อภิปรายผลการวิจัย

1. จากผลการวิจัยที่โมเดลองค์ประกอบอันดับหนึ่ง 9 องค์ประกอบมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ แสดงว่าลักษณะความสัมพันธ์ระหว่างตัวบ่งชี้สามารถรวมกันอธิบายในมิติใหม่ได้ 9 มิติ ได้อย่างน่าเชื่อถือ ผลการวิจัยดังกล่าวเป็นการยืนยันถึงความเป็นสากลของแบบประเมินการสอน SEEQ ของมาร์ช ที่สามารถนำมาใช้ในบริบทต่างๆ ได้อย่างกว้างขวางรวมทั้งในประเทศไทยด้วย อย่างไรก็ตามเพื่อเป็นการแสดงคุณภาพของแบบประเมินการสอน SEEQ ในรายละเอียดว่าตัวบ่งชี้แต่ละตัวมีคุณภาพในการอธิบายองค์ประกอบแต่ละตัวหรือไม่ ผู้วิจัยจึงทำการแสดงหลักฐานความเที่ยงตรงเชิงผู้เข้า โดยดัชนีที่บ่งบอกถึงคุณภาพของตัวบ่งชี้ที่ใช้อธิบายองค์ประกอบ พบว่าเป็นตัวบ่งชี้ที่ดีในเกือบทุกองค์ประกอบยกเว้นตัวบ่งชี้ในองค์ประกอบปริมาณงาน/ความยากพบว่าสัมประสิทธิ์องค์ประกอบคะแนนมาตรฐานค่อนข้างต่ำกว่าสัมประสิทธิ์ของตัวบ่งชี้ในองค์ประกอบอื่นๆ แต่ผ่านเกณฑ์ขั้นต่ำที่ค่าสัมประสิทธิ์องค์ประกอบควรสูงกว่า .40 ยกเว้นข้อคำถามสุดท้ายที่

ถามถึงจำนวนชั่วโมงต่อสัปดาห์นอกชั้นเรียน แสดงว่าตัวบ่งชี้ที่มีความเที่ยงตรงเชิงผู้เข้าตำมากสมควรมีการศึกษาเพิ่มเติมในอนาคตว่าในบริบทสังคมไทยตัวบ่งชี้ที่เหมาะสมในการใช้วัดในองค์ประกอบปริมาณงาน/ความยากหรือไม่ หรือสมควรมีการปรับปรุงมาตรฐานในการวัดหรือไม่ เนื่องจากตามต้นฉบับภาษาอังกฤษ ข้อคำถามนี้เป็นข้อคำถามเดียวที่ไม่ได้กำหนดมาตรฐานวัดได้ เหมือนกับข้อคำถามอื่นๆ แต่เป็นการถามโดยตรงถึงจำนวนชั่วโมงที่ใช้ในการศึกษานอกชั้นเรียน วิธีการวัดดังกล่าวอาจจะเป็นสาเหตุหนึ่งที่ทำให้เกิดความแปรปรวนปลอม (spurious variance) ที่ทำให้สัมประสิทธิ์องค์ประกอบของตัวบ่งชี้มีค่าต่ำมากกับกลุ่มตัวอย่างผู้เรียนในประเทศไทย ซึ่งอาจเกิดมากจากความคุ้นเคยในการตอบแบบสอบถามที่ต่างวัฒนธรรม

2. จากการเปรียบเทียบความกลมกลืนระหว่างโมเดลองค์ประกอบอันดับสองระหว่างโมเดล 2 องค์ประกอบตามแนวคิดของอุทุมพร จามรมาน (2530) และ โมเดล 4 องค์ประกอบ ที่ปรับมาจากแนวคิดของมาร์ช (Marsh, 1991) พบว่าโมเดล 4 องค์ประกอบมีความกลมกลืนกับข้อมูลเชิงประจักษ์มากกว่า ความแตกต่างดังกล่าวอาจเกิดขึ้นมากจากการกำหนดคุณลักษณะเฉพาะของโมเดลของทั้งสองโมเดล กล่าวคือในองค์ประกอบอันดับสองในทั้งสองโมเดล มีตัวบ่งชี้ที่เป็นองค์ประกอบอันดับหนึ่งเพียงองค์ประกอบเดียว โดยในโมเดล 2 องค์ประกอบองค์ประกอบผลลัพธ์มีตัวบ่งชี้คือองค์ประกอบอันดับหนึ่งการเรียนรู้/คุณค่า ซึ่งมีข้อความจำนวน 4 ข้อ แต่ในโมเดล 4 องค์ประกอบ องค์ประกอบเนื้อหางานมีตัวบ่งชี้คือองค์ประกอบอันดับหนึ่งการทดสอบ/การให้เกรด ซึ่งมีข้อความจำนวน 3 ข้อความ การกำหนดค่าพารามิเตอร์ (Constrained) ที่ไม่ตรงกับลักษณะความแปรปรวนแปรปรวนร่วมของข้อมูลจะส่งผลให้ค่าไคสแควร์เพิ่มขึ้น ในที่นี้จำนวนการกำหนดพารามิเตอร์ของโมเดล 2 องค์ประกอบมีมากกว่าโมเดล 4 องค์ประกอบ ทำให้ค่าไคสแควร์สูงกว่า

3. ถึงแม้จากผลการประเมินโมเดลจะพบว่าองค์ประกอบอันดับสองจำนวน 2 องค์ประกอบจะมีความสอดคล้องกับข้อมูลเชิงประจักษ์น้อยกว่าโมเดล 4 องค์ประกอบ แต่ทั้งสองโมเดลก็มีความกลมกลืนกับข้อมูลเชิงประจักษ์ในระดับดีเช่นเดียวกัน แสดงว่าลักษณะความสัมพันธ์ระหว่างตัว

แปรแฝงอันดับหนึ่งสามารถรวมกันอธิบายในมิติใหม่ได้ 2 มิติ และ 4 มิติได้อย่างมีความน่าเชื่อถือ ผลการวิจัยดังกล่าวเป็นการยืนยันเกี่ยวกับแนวคิดการประเมินการสอนของอุทุมพร จามรมาน (2530) และแนวคิดของมาร์ช (1991) นอกจากนี้พบว่าโมเดลองค์ประกอบอันดับสองทั้งสองโมเดลมีความเที่ยงตรงเชิงผู้เข้า โดยพบว่าค่าสัมประสิทธิ์องค์ประกอบส่วนใหญ่มีขนาดสัมประสิทธิ์สูงกว่า .40 แสดงว่าองค์ประกอบอันดับหนึ่งที่กำหนดให้เป็นตัวบ่งชี้ขององค์ประกอบอันดับสองสามารถนำมารวมคะแนนเป็นองค์ประกอบได้อย่างมีความหมายและน่าเชื่อถือ อย่างไรก็ตามองค์ประกอบอันดับหนึ่งปริมาณงาน/ความยากมีขนาดสัมประสิทธิ์คะแนนมาตรฐานที่ต่ำกว่าองค์ประกอบอันดับหนึ่งอื่นๆ ในการอธิบายความแปรปรวนขององค์ประกอบในทั้งสองโมเดล ซึ่งเมื่อพิจารณาแล้วสามารถอธิบายได้ว่า องค์ประกอบปริมาณงาน/ความยากเป็นองค์ประกอบที่วัดพฤติกรรมเกี่ยวกับรายวิชาโดยตรง เป็นการวัดการรับรู้ของผู้เรียนเกี่ยวกับเนื้อหา ปริมาณงานของรายวิชา ระดับการเรียนรู้ของผู้เรียนในรายวิชา ส่วนองค์ประกอบอื่นๆ จะเป็นการรับรู้ถึงบรรยากาศต่างๆ ในชั้นเรียน เช่น ความเอาใจใส่ผู้เรียน ความสัมพันธ์ระหว่างครูกับผู้เรียน เป็นต้น ลักษณะความสัมพันธ์ขององค์ประกอบปริมาณงาน/ความยาก กับองค์ประกอบอื่นๆ จึงไม่เป็นเชิงเส้นตรง (ซึ่งส่งผลให้ค่าองค์ประกอบร่วมกับตัวบ่งชี้อื่นๆ ในการอธิบายความแปรปรวนขององค์ประกอบอันดับสองด้านกระบวนการมีค่าต่ำ) เนื่องจากเมื่อผู้เรียนรับรู้ว่ามีภาระงานยิ่งมากจะมีความสนใจที่จะเรียนมากขึ้น และเมื่อรับรู้ว่ามีภาระงานน้อยผู้เรียนจะไม่ค่อยกระตือรือร้นที่จะเรียน อย่างไรก็ตามเมื่อผู้เรียนรับรู้ว่ามีภาระงานมากเกินไปความผู้เรียนจะรู้สึกเบื่อหน่ายในการเรียนเพิ่มขึ้นด้วย (Marsh, & Roche, 2000)

4. สำหรับองค์ประกอบอันดับสามซึ่งเป็นการรวมองค์ประกอบอันดับสองเป็นองค์ประกอบการประเมินการสอนในภาพรวมก็พบว่ามีความสอดคล้องกับข้อมูลเชิงประจักษ์เช่นกัน และค่าสัมประสิทธิ์องค์ประกอบของทั้งสององค์ประกอบมีค่าสูงมากในทั้งสองโมเดล (.97 ถึง 1.0)

แสดงว่าลักษณะความสัมพันธ์ระหว่างตัวแปรแฝงอันดับสองทั้งสองโมเดลสามารถรวมกันอธิบายในภาพรวมเป็นการประเมินการสอนได้อย่างมีความน่าเชื่อถือ

ข้อเสนอแนะ

การวิจัยครั้งนี้ ผู้วิจัยมีข้อเสนอแนะ 2 ส่วน คือ ข้อเสนอแนะเพื่อนำผลการวิจัยไปใช้ และข้อเสนอแนะสำหรับงานวิจัยครั้งต่อไปดังนี้

ข้อเสนอแนะเพื่อนำผลการวิจัยไปใช้

1. จากผลการวิจัยที่พบว่า แบบประเมิน SEEQ จำนวน 35 ข้อ เมื่อนำมารวมกันเป็นองค์ประกอบจำนวน 9 องค์ประกอบ มีความสอดคล้องกับข้อมูลเชิงประจักษ์ในบริบทของสังคมไทยเช่นเดียวกับผลการวิจัยที่ผ่านมา หลักฐานดังกล่าวแสดงว่าสารสนเทศการประเมินผลการสอนโดยผู้เรียนสามารถวัดพฤติกรรมการสอนของอาจารย์ได้เป็น 9 ด้าน สารสนเทศดังกล่าว อาจารย์สามารถนำคะแนนมารวมกันเป็น 9 ด้าน และเพื่อให้ทราบถึงระดับคุณภาพของการจัดการเรียนการสอนของตนเองตามการรับรู้ของผู้เรียน ความหลากหลายของมิติการประเมินพฤติกรรมการสอนทำให้ทราบว่า ผู้สอนมีจุดแข็งและ/หรือจุดอ่อนในมิติใดบ้าง ข้อมูลดังกล่าวสามารถนำไปใช้ในการพัฒนา/ปรับปรุงการจัดการเรียนการสอนของผู้สอนได้อย่างมีประสิทธิภาพ

2. จากการแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างองค์ประกอบที่ผู้วิจัยเสนอไว้ นอกจากเสนอโมเดลอันดับหนึ่งจำนวน 9 องค์ประกอบแล้ว ยังทำการเสนอโมเดลทางเลือกอื่นๆ ได้แก่ โมเดลองค์ประกอบอันดับสอง: 2 องค์ประกอบ และ 4 องค์ประกอบ ซึ่งพบว่าโมเดลดังกล่าวมีความสอดคล้องกับข้อมูลเชิงประจักษ์ในบริบทของสังคมไทย ทั้ง 2 โมเดล ดังนั้นสารสนเทศในภาพรวมที่มีข้อมูลน้อยกว่า 9 องค์ประกอบ สามารถนำไปใช้ประโยชน์ในหน่วยงานอื่นๆ ที่ต้องการข้อมูลเกี่ยวกับคุณภาพการจัดการเรียนการสอนของอาจารย์ เช่น การบริหารบุคคล สามารถนำผลการประเมินไปใช้เพื่อการประเมินผลการทำงานของอาจารย์ (Performance Appraisal) ความหลากหลายของข้อมูลที่มีหลายมิติ สามารถเป็นข้อมูลในการตัดสินใจที่หลากหลาย ผู้บริหารสามารถใช้ข้อมูลทุกมิติหรือเลือกใช้ผลการประเมินบางมิติ ขึ้นอยู่กับการให้ความสำคัญในเชิงบริหารว่าให้ความสำคัญกับการจัดการ

เรียนการสอนในมิติใด ผลการวิจัยครั้งนี้ได้เสนอทางเลือกในการนำเสนอข้อมูลเพื่อการพัฒนาการจัดการเรียนการสอน การบริหารงานบุคคล และการประกันคุณภาพการศึกษา ที่หน่วยงานภายในมหาวิทยาลัยสามารถนำไปประยุกต์ใช้ในการนำเสนอผลการประเมินได้อย่างหลากหลาย ทั้งมิติ 9 องค์ประกอบ มิติ 2 องค์ประกอบ มิติ 4 องค์ประกอบ และมิติในภาพรวมของการประเมินการสอน

ข้อเสนอแนะสำหรับงานวิจัยครั้งต่อไป

1. เนื่องจากรายงานวิจัยครั้งนี้ใช้หน่วยการวิเคราะห์เป็นผู้เรียนซึ่งพบว่าให้ผลการวิเคราะห์ที่ดี อย่างไรก็ตามธรรมชาติของการประเมินการสอนมีความแปรปรวนอย่างเป็นระบบที่แตกต่างกันระหว่างชั้นเรียน ทำให้การใช้ข้อมูลระดับผู้เรียนทำให้ผลการประเมินโมเดลและพารามิเตอร์มีค่าสูงเกินความเป็นจริง (Muthen, 1994) การวิจัยที่ผ่านมาจึงนิยมใช้ค่าเฉลี่ยชั้นเรียนในการวิจัยในประเด็นเกี่ยวกับแบบประเมินการสอน อย่างไรก็ตามการใช้ค่าเฉลี่ยชั้นเรียนมีข้อจำกัดอยู่ใน 3 ประเด็นได้แก่ (1) การใช้ค่าเฉลี่ยชั้นเรียนเป็นการลดสารสนเทศที่ได้จากข้อมูลระดับผู้เรียน (2) ค่าเฉลี่ยที่ได้ต้องเป็นตัวแทนที่ดีของชั้นเรียน ซึ่งหมายความว่าจำนวนผู้เรียนในชั้นเรียนต้องใหญ่พอ และความแตกต่างระหว่างจำนวนผู้เรียนในแต่ละชั้นเรียนต้องไม่มากนัก ซึ่งในทางปฏิบัติทำให้ชั้นเรียนบางชั้นเรียนต้องถูกตัดออกจากการวิเคราะห์ ซึ่งชั้นเรียนนั้นอาจจะมีลักษณะที่ผู้วิจัยต้องการศึกษาก็เป็นได้ และ (3) การใช้ค่าเฉลี่ยชั้นเรียนในการวิเคราะห์ข้อมูลต้องใช้กลุ่มตัวอย่างขนาดใหญ่มาก จำเป็นต้องใช้ทุนในการวิจัยสูง ดังนั้นการศึกษาต่อไปจึงควรมีวิธีการศึกษาที่สามารถใช้สารสนเทศของข้อมูลทั้งสองระดับในการศึกษาไปพร้อมๆ กัน สามารถช่วยลดขนาดของข้อมูลระดับชั้นเรียนลงมาได้ ซึ่งพัฒนาการล่าสุดของการวิเคราะห์องค์ประกอบเชิงยืนยันที่สามารถวิเคราะห์ข้อมูลได้ทั้งสองระดับคือการวิเคราะห์องค์ประกอบเชิงลดหลั่น (Hierarchical Factor Analysis) ดังนั้นการวิจัยครั้งต่อไปจึงควรมีวิจัยการวิเคราะห์ดังกล่าวมาใช้ เพื่อลดข้อจำกัดในการวิเคราะห์ที่กล่าวมาข้างต้น

2. ควรมีการศึกษาเพิ่มเติมเกี่ยวกับปัจจัยที่ส่งผลต่อการประเมินการสอนของนักศึกษา เพื่อให้ทราบถึง

สาเหตุที่ทำให้ให้นักศึกษาให้ผลการประเมินที่แตกต่างกัน เพื่อให้ได้ทราบว่ามีสาเหตุจากผลที่เกิดจากการสอนแล้วมีปัจจัยอะไรบ้าง ทั้งปัจจัยระดับชั้นเรียนและปัจจัยระดับผู้เรียน ที่ทำให้นักศึกษาเกิดการรับรู้ถึงคุณภาพของการสอนที่แตกต่างกัน

3. ควรมีการศึกษาความเที่ยงตรงแบบประเมินการสอน SEEQ โดยการให้แหล่งข้อมูลอื่นๆ ในการประเมินคุณภาพ

การจัดการเรียนการสอน เช่น ผลการประเมินตนเอง ผลการประเมินจากเพื่อนร่วมงาน และ ผลการประเมินจากผู้บังคับบัญชา เพื่อให้ได้ผลการประเมินที่รอบด้าน และทำการศึกษาความเที่ยงตรงเชิงผู้เข้าและความเที่ยงตรงเชิงจำแนกโดยนำผลการประเมินจากแหล่งข้อมูลหลายแหล่งมาศึกษาร่วมกัน

บรรณานุกรม

- ศิริชัย กาญจนวาสี. (2543). การประเมินการสอนระดับอุดมศึกษา. กรุงเทพฯ: ทบวงมหาวิทยาลัย. ถ่ายเอกสาร
- อนุ เจริญวงศ์ระยับ. (2549). การแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบประเมินการสอนโดยผู้เรียน
ของมาร์ช ด้วยการประยุกต์ใช้วิธีการวิเคราะห์กลุ่มพหุ. ปรินซิพนิพนธ์ กศ.ม.(การวัดผลการศึกษา). กรุงเทพฯ:
บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร
- อุทุมพร จามรมาน. (2530). การวัดและประเมินการเรียนการสอนระดับอุดมศึกษา. กรุงเทพฯ: โครงการตำรา
วิทยาศาสตร์อุตสาหกรรม.
- Hativa, N. (2000). *Teaching for Effective Learning in Higher Education*. Dordrecht: Kluwer.
- Kulik, J. A. (2001). *Student Rating: Validity, Utility, and Controversy*. In *The student Ratings Debate: Are They Valid? How Can We Best Use Them?* Theall, M. , P.C. Abrami, and L.A. Mets. (editors). pp. 9-26. San Francisco: Jossey-Bass.
- Marsh, H. W.; Hau, K. T. Chung, C.M.; & Siu, T. L. P. (1998). *Confirmatory Factor Analyses of Chinese Students' Evaluations of University Teaching*. *Structural Equation Modeling*. 5(2): pp.143-164.
- Marsh, H.W.; & Hocevar, D. (1985). *The Factorial Invariance of Student Evaluations of College Teaching*. *American Educational Research Journal*. 21(2): pp.341-366.
- Marsh, H. W.; Roche. L. A. (2000). *Effects of Grading Leniency and Low Workload on Students' Evaluations of Teaching: Popular Myth, Bias, Validity, or Innocent Bystanders?*. *Journal of Educational Psychology*. 92(1): pp.202-228.
- Marsh, H. W. (1982). *Validity of Students' Evaluations of College Teaching: A Multitrait-Multimethod Analysis*. *Journal of Educational Psychology*. 74(2): pp.264-279.
- Marsh, H. W. (1986). *Applicability Paradigm: Students' Evaluations of Teaching Effectiveness in Different Countries*. *Journal of Educational Psychology*. 78(6): pp.465-473.
- Marsh, H. W. (1987). *Students' Evaluations of University Teaching: Research Findings, Methodological Issues, and Directions of Future Research*. *International Journal Educational Research*. 11: pp.253-388.
- (1991). *Multidimensional Students' Evaluations of Teaching Effectiveness: A Test of Alternative Higher-Order Structures*. *Journal of Educational Psychology*. 83(2): 285-296.
- Muthen, Bengt O. (1994). *Multilevel Covariance Structure Analysis*. *Sociological Methods & Research*. 22(3): 376-398.
- Stake, R. E. 1989. *The Evaluation of Teaching*. In *Rethinking Appraisal and Assessment*. Simons, H. , J. Elliott. (editors). Bristol: Open University.