

# วิธีการเชิงปริมาณใหม่ๆ สำหรับศาสตร์ประเมิน

## Modern Quantitative Methods for Evaluation Science

รัตนะ บัวสนธิ<sup>1</sup>

### บทคัดย่อ

ในช่วงระยะเวลาประมาณ 10 กว่าปีที่ผ่านมา หรือตั้งแต่ปี ค.ศ. 2000 เป็นต้นมา วิธีการเชิงปริมาณที่ใช้สำหรับการวิจัย ได้มีการพัฒนาและนำเทคนิควิธีวิเคราะห์ทางสถิติใหม่ๆ มาใช้เป็นเครื่องมือในการวิเคราะห์ข้อมูลเพื่อหาคำตอบให้กับคำถามการวิจัยอย่างแพร่หลาย ซึ่งวิธีการดังกล่าวนี้สามารถที่จะนำมาใช้สำหรับศาสตร์การประเมิน โดยเฉพาะ การประเมินโครงการทางสังคม หรือทางการศึกษาได้เป็นอย่างดี วิธีการเชิงปริมาณใหม่ๆ นี้ แบ่งออกเป็น 3 กลุ่มใหญ่ๆ คือ 1) การสรุปเชิงสาเหตุหรือการออกแบบเชิงทดลอง 2) โมเดลพหุระดับ และ 3) สมการเชิงโครงสร้าง หรือโมเดลตัวแปรแฝง

**คำหลัก** 1) การสรุปอ้างอิงเชิงสาเหตุ/การออกแบบการทดลอง 2) โมเดลพหุระดับ 3) สมการเชิงโครงสร้าง หรือโมเดลตัวแปรแฝง

### บทนำ

ผลจากการพัฒนาการที่ก้าวหน้าของวิธีการเชิงปริมาณสำหรับงานวิจัย โดยเฉพาะเกี่ยวกับด้านวิธีการวิเคราะห์ทางสถิติก่อให้เกิดการเปิดพรมแดนความรู้ และการใช้วิธีการที่ทันสมัยเพื่อหาคำตอบในสิ่งที่สงสัยให้กับนักวิจัยได้อย่างกว้างขวางขึ้น ซึ่งผลดังกล่าวก็ย่อมส่งผลกระทบต่อ การเปิดพรมแดนเกี่ยวกับวิธีการประเมินโครงการหรือแผนงานให้กับนักประเมินที่นิยมใช้วิธีการประเมินเชิงปริมาณตามไปด้วยเช่นกัน

ในบทความนี้ มุ่งนำเสนอแนวคิดเกี่ยวกับวิธีการเชิงปริมาณใหม่ๆ สำหรับใช้ประเมินโครงการทั้ง 3 กลุ่ม โดยจะกล่าวถึงแนวคิดที่มาและตัวอย่างตามลำดับ ดังต่อไปนี้

<sup>1</sup> ศาสตราจารย์ประจำสาขาวิจัยและพัฒนาศาสตร์ศึกษา มหาวิทยาลัยนครสวรรค์

## 1. การสรุปอ้างอิงเชิงสาเหตุหรือการออกแบบการทดลอง: Causal Inference/ Experimental Design

วิธีการเชิงปริมาณในกลุ่มนี้ก็คือ การอาศัยหลักคิดเกี่ยวกับการออกแบบการทดลองในงานวิจัยเพื่อที่จะให้สามารถสรุปผลของการวิจัย (ซึ่งก็คือตัวแปรตาม) ว่าเกิดขึ้นเนื่องมาจากตัวแปรต้นที่เป็นสาเหตุอย่างแท้จริง โดยอาศัยหลักการควบคุมสุ่ม (randomized control) กลุ่มผู้เข้ารับเงื่อนไขการทดลองให้มีความเท่าเทียมกันก่อนที่จะได้รับเงื่อนไขการทดลอง (ซึ่งเป็นตัวแปรต้นหรือสาเหตุ) และเมื่อทำการทดลองหรือให้เงื่อนไขการทดลองกับกลุ่มผู้เข้ารับการทดลอง (subjects) เสร็จสิ้นแล้วการสรุปผลก็จะมีลักษณะการสรุปอ้างอิงความสัมพันธ์เชิงสาเหตุ (causal relationship) ระหว่างตัวแปรต้นและตัวแปรตาม การออกแบบการทดลองสามารถออกแบบได้หลายแบบ ซึ่งแต่ละแบบแผนการทดลองก็จะให้ความตรงภายใน (internal validity) และความตรงภายนอก (external validity) ของผลการวิจัยแตกต่างกัน ดังที่ Campbell and Stanley (1963) ได้วิเคราะห์ วิจัย เสนอไว้ในหนังสือที่ทั้งสองคนเขียนไว้ในปีดังกล่าว

จากหลักคิดของวิธีการเชิงปริมาณในงานวิจัยเกี่ยวกับการออกแบบการทดลองที่กล่าวมานั้น ก็จะสามารถนำมาประยุกต์ใช้กับการประเมินโครงการได้เช่นเดียวกัน ซึ่งในการประเมินโครงการนั้น จะถือว่าโครงการ (program) หรือสิ่งแทรกแซง (intervention) ก็คือ ตัวแปรต้นหรือเงื่อนไขที่เป็นสาเหตุ (บางครั้งอาจเรียกว่า ตัวแปรสาเหตุ หรือ treatment variable) ส่วนผลที่เกิดขึ้นกับกลุ่มเป้าหมายที่ได้รับบริการจากโครงการ (target or client group) ซึ่งมักเรียกว่า ประสิทธิภาพ (effectiveness) ของโครงการ ก็คือ ตัวแปรผลหรือตัวแปรตามนั่นเอง ด้วยเหตุนี้เอง ความน่าเชื่อถือของการสรุปอ้างอิงเชิงสาเหตุของโครงการกับผลที่เกิดขึ้น จึงต้องอาศัยหลักการออกแบบการทดลองจากงานวิจัยที่เป็นไปตามหลัก “Max-Min-Con” (Kerlinger, 1986, pp. 279–291) เพื่อให้เกิดความตรงภายในและความตรงภายนอกของการประเมินโครงการให้มากที่สุดนั่นเอง

แต่อย่างไรก็ดี บางครั้งในการนำโครงการไปให้บริการหรือใช้กับกลุ่มผู้รับบริการนั้น ไม่สามารถจะอาศัยหลักการควบคุมสุ่ม ซึ่งเป็นหลักสำคัญอย่างหนึ่งของหลัก “Max-Min-Con” เพราะกลุ่มผู้รับบริการได้กำหนดไว้อย่างเจาะจงแล้ว ยกตัวอย่างเช่น การนำโครงการป้องกันไข้เลือดออกไปใช้กับกลุ่มเป้าหมายที่อาศัยอยู่ตามแนวชายแดนของประเทศซึ่งเป็นกลุ่มเสี่ยง หรือโครงการส่งเสริมการรู้หนังสือภาษาไทยสำหรับผู้อพยพชาวโรฮิงญา เป็นต้น กรณีเช่นนี้ Shadish, Cook and Campbell (2002) ได้เสนอให้ใช้แบบแผนเชิงกึ่งทดลอง (quasi-experimental designs)

ที่มีความแกร่งในการสรุปอ้างอิงเชิงสาเหตุของโครงการไว้สองแบบแผน คือ การถดถอยแบบไม่ต่อเนื่อง (regression discontinuity: RD) การรบกวนแบบอนุกรมเวลา (interrupted time series: ITS) นอกเหนือจากสามคนดังกล่าวนี้ Morgan and Winship (2007) ก็ยังได้นำเสนอเกี่ยวกับแบบแผนทางเลือกอื่นๆ รวมทั้งการใช้วิธีการทางสถิติสำหรับการสรุปอ้างอิงเชิงสาเหตุเกี่ยวกับโครงการหรือสิ่งแทรกแซงทางสังคม โดยอาศัยพื้นฐานทางคณิตศาสตร์สำหรับการประมาณค่าเชิงสาเหตุที่แตกต่างกันในแต่ละแบบแผนทางเลือก (design alternatives) นั้นๆ นอกจากนี้ Morgan and Winship ยังเสนอให้ใช้แผนภูมิเป็นตัวแทนแสดงให้เห็นถึงเครือข่ายเชิงสาเหตุของตัวแปรต่างๆ ที่เกิดขึ้นอย่างเกี่ยวข้องเชื่อมโยงกันของแต่ละแบบแผนการทดลองใช้โครงการที่นำไปใช้ในบริบทที่แตกต่างกัน ตลอดจนการใช้วิธีการทางสถิติสำหรับปรับประมาณค่าความสัมพันธ์เชิงสาเหตุของโครงการกับผลที่เกิดขึ้นจากการใช้โครงการอีกด้วย

ตัวอย่างวิธีการเชิงปริมาณใหม่ๆ ที่ใช้ในการสรุปอ้างอิงเชิงสาเหตุหรือการออกแบบการทดลอง ได้แก่ งานของ Clair, Cook and Hallberg (2014, pp. 311–327) ที่ทำการศึกษาเปรียบเทียบความตรงภายในและความแม่นยำทางสถิติระหว่างแบบแผนการทดลองแบบการรบกวนอนุกรมเวลากับแบบแผนการทดลองแบบสุ่ม ซึ่งผู้วิจัยต้องการทดสอบยืนยันให้เห็นว่าในกรณีการประเมินโครงการโดยอาศัยหลักคิดของงานวิจัยเชิงทดลองเกี่ยวกับการออกแบบการทดลองใช้โครงการที่ไม่สามารถจะใช้แบบแผนการทดลองแบบสุ่มกับผู้รับบริการจากโครงการได้ด้วยเหตุผลข้อจำกัดตามที่กล่าวมานั้น นักประเมินก็จะสามารถใช้แบบแผนการทดลองแบบการรบกวนอนุกรมเวลาเพื่อสรุปอ้างอิงเชิงสาเหตุระหว่างโครงการหรือสิ่งแทรกแซงกับผลที่เกิดขึ้นได้อย่างน่าเชื่อถือไม่น้อยไปกว่าการใช้แบบแผนการทดลองแบบสุ่ม ผู้วิจัยทั้งสามใช้เทคนิควิธีการเปรียบเทียบการศึกษาภายใน (within-study comparison: WSC) ระหว่างแบบแผนการทดลองแบบการรบกวนอนุกรมเวลากับแบบแผนการทดลองแบบสุ่ม โดยใช้ข้อมูลจากการประเมินโครงการ Indiana's Diagnostic Assessment Intervention ที่ออกแบบการประเมินแบบสุ่ม และใช้ผลปฏิบัติทางการเรียนของนักเรียนตั้งแต่เกรด 3–8 ที่ได้จากการทดสอบของรัฐเป็นตัวบ่งชี้ประสิทธิผลของโครงการ จากข้อมูลการประเมินดังกล่าวนี้ นักวิจัยได้นำมาใช้ออกแบบการประเมินแบบรบกวนอนุกรมเวลาหลายแบบ เช่น เพิ่มจำนวนครั้งการสอบก่อน (pretest) รวมทั้งมีการจับคู่และไม่จับคู่ระหว่างโรงเรียนด้วย หลังจากนั้นก็ประมาณค่าความแม่นยำของคะแนนผลการเรียนของนักเรียน โดยอาศัยสมการทางคณิตศาสตร์สำหรับประมาณค่าคะแนนเปรียบเทียบกันระหว่างค่าที่ได้จากสมการตามแบบแผนการประเมินแบบสุ่มและแบบแผนการประเมินแบบการรบกวนอนุกรมเวลาว่าจะได้ผลเป็นประการใด ตามวัตถุประสงค์

การศึกษาเปรียบเทียบที่นักวิจัยตั้งไว้ทั้งสิ้น 4 ข้อ (Clair, Cook and Hallberg, 2014, p. 314) ผลการศึกษาเปรียบเทียบ พบว่า แบบแผนการทดลอง (หรือการประเมิน) แบบการรวบรวมอนุกรมเวลา ให้ผลการประมาณค่า ประสิทธิภาพของโครงการ (คะแนนผลการเรียนของนักเรียน) ใกล้เคียงกับแบบแผนการทดลองแบบสุ่ม จากงานวิจัยนี้ สรุปได้ว่า การใช้แบบแผนการประเมินโครงการแบบการรวบรวมอนุกรมเวลาสามารถใช้ประเมินโครงการและสรุปผลการอ้างอิงเชิงสาเหตุของโครงการได้แม่นยำไม่ต่างไปจากแบบแผนการทดลองแบบสุ่ม

## 2. โมเดลพหุระดับ: Multilevel Modeling

วิธีการเชิงปริมาณที่เรียกว่า โมเดลพหุระดับ นั้น เป็นวิธีการสถิติที่ใช้วิเคราะห์ข้อมูล โดยให้ความสำคัญและคำนึงถึงความแตกต่างของระดับข้อมูลที่ลดหลั่นกันลงมาตามระดับเล็กหรือระดับจุลภาค (Micro Level) และระดับใหญ่หรือระดับมหภาค (Macro Level) นั่นคือ แยกวิเคราะห์ตามระดับข้อมูลที่ลดหลั่นกัน ซึ่งเป็นวิธีการทางสถิติที่แตกต่างไปจากเดิม ซึ่งมักนำข้อมูลที่แม้จะอยู่ในระดับต่างกันมารวมวิเคราะห์ไว้ด้วยกัน ยกตัวอย่างเช่น ในการศึกษาหาคำตอบว่า ปัจจัยใดบ้างที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียน นักวิจัยได้ทำการทบทวนวรรณกรรมแล้วพบว่า ปัจจัยที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียน ได้แก่ ฐานะเศรษฐกิจของครอบครัว ลำดับที่เกิด สภาพแวดล้อมของครอบครัว เพศ เจตคติ วิธีการสอนของครู ขนาดห้องเรียน สภาพแวดล้อมของโรงเรียน การได้รับงบประมาณสนับสนุนการจัดการศึกษาของโรงเรียน และความสัมพันธ์ของชุมชนกับโรงเรียน จากตัวอย่างนี้ จะเห็นว่าสามารถจัดระดับของข้อมูลตามลักษณะปัจจัยที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนได้เป็น 3 ระดับลดหลั่นกัน กล่าวคือ ปัจจัยในระดับตัวนักเรียนแต่ละคน (Micro Level) ก็คือ ฐานะเศรษฐกิจของครอบครัว ลำดับที่เกิด สภาพแวดล้อมทางครอบครัว เพศ และเจตคติของนักเรียน ในขณะที่ปัจจัยในระดับชั้นเรียน (Macro Level) ที่ร่วมส่งผลต่อค่าเฉลี่ยผลสัมฤทธิ์ทางการเรียนของนักเรียนในห้องเรียนหนึ่งๆ ได้แก่ วิธีการสอนของครู และขนาดห้องเรียน ในทำนองเดียวกันกับปัจจัยในระดับโรงเรียน (Macro Level) ที่ร่วมส่งผลต่อค่าเฉลี่ยผลสัมฤทธิ์ทางการเรียนของนักเรียนในโรงเรียนหนึ่ง ได้แก่ สภาพแวดล้อมของโรงเรียน การได้รับงบประมาณสนับสนุน และความสัมพันธ์ของชุมชนกับโรงเรียน เป็นต้น วิธีการเชิงปริมาณที่เรียกว่า โมเดลพหุระดับนี้ ผู้ที่นำเสนอแนวคิดและเขียนอธิบายไว้ในตำราซึ่งเป็นที่รู้จักกันอย่างแพร่หลาย ได้แก่ Raudenbush

and Bryk (2002) Goldstein (2011) Luke (2004) Hox (2010) และ Snijder and Bosker (2012) และศิริชัย กาญจนวาสี (2542)

สำหรับการประเมินโครงการต่างๆ นั้น ก็จะต้องอยู่บ่อยๆ ว่า นักประเมินได้มีการประเมินโครงการที่มีลักษณะข้อมูลเป็นไปในทำนองพหุระดับ หรือต่างระดับลดหลั่นกัน เช่น โครงการโรงเรียนสีขาว โครงการอาหารกลางวัน หรือโครงการเศรษฐกิจพอเพียง เหล่านี้เป็นต้น ในกรณีโครงการโรงเรียนสีขาวที่อาจประเมินประสิทธิผลของโครงการจากความรู้ความเข้าใจ พฤติกรรมของนักเรียนเกี่ยวกับยาเสพติด และการหลีกเลี่ยงยาเสพติด จะพบว่า มีปัจจัยในระดับต่างกันที่ส่งผลต่อประสิทธิผลของโครงการนี้ อาทิ ในระดับนักเรียน ได้แก่ ลักษณะของครอบครัว จำนวนสมาชิกในครอบครัว ลักษณะการอบรมเลี้ยงดู การคบเพื่อน ส่วนปัจจัยในระดับโรงเรียน และห้องเรียน ได้แก่ การควบคุมชั้นเรียนของครู บรรยากาศในห้องเรียน ขนาดห้องเรียน ที่ตั้ง/สภาพแวดล้อมของโรงเรียน ความสัมพันธ์ของผู้ปกครองกับโรงเรียน เหล่านี้เป็นต้น ดังนั้น การใช้วิธีการเชิงปริมาณโมเดลพหุระดับ สำหรับประเมินผลโครงการ จึงสามารถนำมาใช้และมีประโยชน์ต่อการอธิบายประสิทธิผลของโครงการได้เป็นอย่างดี

ตัวอย่างงานประเมินโครงการที่ใช้วิธีการเชิงปริมาณแบบใหม่ด้วยโมเดลพหุระดับ ได้แก่ งานของ Gebreselassie, Stephenson, Maplex, Johnson and Tucker (2014, pp. 214–229) ที่ทำการประเมินโครงการ การให้บริการความรู้ความเข้าใจทางสุขภาพจิตของชุมชนแก่เด็กและครอบครัว โดยศึกษาประเมินความคงทนของความรู้ความเข้าใจทางสุขภาพจิตของเด็กและครอบครัวที่ได้รับบริการจากโครงการดังกล่าว ซึ่งในการประเมินครั้งนี้ ผู้ประเมินได้จำแนกข้อมูลแบ่งเป็นระดับตามปัจจัยต่างๆ ได้แก่

2.1 ระดับที่ 1 ได้แก่ ปัจจัยช่วงเวลากการศึกษาติดตาม

2.2 ระดับที่ 2 ได้แก่ ปัจจัยส่วนบุคคลและครอบครัว ประกอบด้วย เพศ อายุ เชื้อชาติ การใช้ชีวิต (คนเดียว หรือ อยู่ร่วมกับคนอื่น) สถานที่ให้การบำบัด (คลินิก หรือ แหล่งอื่นๆ) คะแนนตรวจสอบพฤติกรรมเด็ก (CBCL: Child Behavior Checklist) และรายได้ครัวเรือน

2.3 ระดับที่ 3 ได้แก่ ปัจจัยคุณลักษณะต่างๆ ของชุมชน ประกอบด้วย ประสพการณ์ที่มีมาก่อนของพื้นที่หรือชุมชนในการได้รับการทางสุขภาพจิต จำนวนค่าใช้จ่ายให้กับผู้ให้สัมภาษณ์ภายหลังจากการสัมภาษณ์ข้อมูลพื้นฐานเสร็จแล้ว และจำนวนค่าใช้จ่ายที่จ่ายให้กับผู้ให้สัมภาษณ์ในแต่ละช่วงเวลากการศึกษาติดตาม

ผลจากการประเมินพบว่า แต่ละปัจจัยที่ต่างระดับกันมีผลต่อการคงทนของความรู้ความเข้าใจทางสุขภาพจิตของเด็กและครอบครัวในชุมชนที่ได้รับการทางสุขภาพจิตต่างกันด้วย

### 3. สมการเชิงโครงสร้าง หรือโมเดลตัวแปรแฝง: Structural Equation/Latent Variable Modeling

วิธีการเชิงปริมาณที่เรียกว่า สมการเชิงโครงสร้างหรือโมเดลตัวแปรแฝง เป็นวิธีการวิเคราะห์ทางสถิติที่มุ่งอธิบายความสัมพันธ์ในเชิงสาเหตุ (causal relationship) ระหว่างตัวแปรแฝง (Latent variable) ที่เป็นตัวแปรสาเหตุ หรือตัวแปรอิสระ (cause or independent variable) โดยที่ตัวแปรแฝงที่เป็นตัวแปรสาเหตุนั้น อาจมีอิทธิพลหรือส่งผลต่อตัวแปรแฝงที่เป็นตัวแปรผล หรือตัวแปรตามทั้งที่เป็นการส่งผลทางตรง และ/หรือทางอ้อม (direct and/or indirect effect) ทั้งนี้ โดยพิจารณาจากการลากเส้นลูกศรจากตัวแปรแฝงสาเหตุตัวหนึ่งๆ ไปยังตัวแปรแฝงผลตัวหนึ่งๆ ความสัมพันธ์เชิงสาเหตุทั้งที่เป็นความสัมพันธ์ทางตรงและหรือทางอ้อมนี้ เรียกว่า สมการเชิงโครงสร้าง (structural equation) ซึ่งอาศัยการวิเคราะห์การถดถอย (regression) ในการประมาณค่าสัมประสิทธิ์ความสัมพันธ์ของเส้นทาง (path coefficient)

ในขณะที่ตัวแปรแฝงแต่ละตัว (ทั้งตัวแปรสาเหตุและตัวแปรแฝงผล) ต่างก็ประกอบด้วย ตัวแปรสังเกตได้ (observe variable) มากกว่า 1 ตัวขึ้นไป ซึ่งความสัมพันธ์ของตัวแปรสังเกตได้ในตัวแปรแฝงหนึ่งๆ นั้น เรียกว่า โมเดลการวัด (measurement model) และอาศัยการวิเคราะห์องค์ประกอบ (factor analysis) ในการประมาณค่าความสัมพันธ์ของตัวแปรสังเกตได้ภายใต้ตัวแปรแฝงหนึ่ง หรือเรียกอีกอย่างหนึ่งว่า ค่าน้ำหนักองค์ประกอบ (factor loading) นั่นเอง

ในการศึกษาหาคำตอบของปัญหาการวิจัยหรือการประเมินนั้น นักวิจัยหรือนักประเมินต้องมีการทบทวนวรรณกรรมในส่วนทฤษฎีแนวคิดที่มีอยู่ และงานวิจัย หรืองานประเมินที่มีผู้ดำเนินการวิจัยไว้ก่อนหน้าแล้วเกี่ยวกับลักษณะและความสัมพันธ์ของตัวแปรแฝงต่างๆ รวมทั้งความสัมพันธ์เกี่ยวข้องระหว่างตัวแปรสังเกตได้กับตัวแปรแฝงหนึ่งๆ หลังจากนั้นก็นำเสนอโมเดลเชิงสมมติฐานหรือเชิงทฤษฎี (hypothetical or theoretical model) ที่แสดงความสัมพันธ์ระหว่างตัวแปรสังเกตได้กับตัวแปรแฝงหนึ่งๆ (โมเดลการวัด) และที่แสดงความสัมพันธ์ระหว่างตัวแปรแฝงแต่ละตัว (โมเดลเชิงโครงสร้าง)

ภายหลังจากนำเสนอโมเดลเชิงสมมติฐานหรือเชิงทฤษฎีแล้วนักวิจัยหรือนักประเมินก็ต้องทำการเก็บรวบรวมข้อมูลเชิงประจักษ์เกี่ยวกับตัวแปรสังเกตได้แต่ละตัว ต่อจากนั้นก็นำมาวิเคราะห์ด้วยโปรแกรมการวิเคราะห์ที่เรียกว่า ลิสเรล (LISREL: **L**inear **S**tructural **R**elationship) เพื่อประมาณค่าพารามิเตอร์ต่างๆ แล้วพิจารณาค่าดัชนีต่างๆ ที่จะบ่งบอกว่าข้อมูลเชิงประจักษ์

หรือโมเดลเชิงประจักษ์ (empirical data or model) มีความสอดคล้อง (fit) กับโมเดลเชิงสมมุติฐานหรือไม่ ซึ่งค่าดัชนีที่บ่งบอกว่าโมเดลสอดคล้องกันหรือไม่ก็มีหลายค่า เช่น สัดส่วนของค่าไควร์สแควร์กับองศาอิสระ ( $\chi^2/df$ ) ยิ่งน้อยยิ่งดี โดยปกติจะใช้เกณฑ์  $\chi^2/df \leq 2$ ) หรือค่า p (probability มากกว่า 0.05) หรือค่า GFI, AGFI มีค่าใกล้ 1.00 (GFI หรือ AGFI คือ **G**oodness of **F**it Index และ **A**just **G**oodness of **F**it Index) เหล่านี้เป็นต้น ทั้งนี้ ผู้ที่นำเสนอแนวคิดและเขียนตำราที่เป็นที่รู้จักแพร่หลาย ได้แก่ Kline (2011) Shumaker and Lomax (2010) Raykov and Marcoulides (2006) Kaplan (2009) Mulaik (2009) และ Loehlin (2004) และนงลักษณ์ วิรัชชัย (2554) ตำราเหล่านี้ มีทั้งที่เป็นตำราเสนอแนวคิดเบื้องต้นและตำราที่นำเสนอแนวคิดวิธีวิเคราะห์อย่างลึกซึ้งซับซ้อน

ยกตัวอย่างสมมติของการใช้สมการเชิงโครงสร้าง/โมเดลตัวแปรแฝงในการประเมินโครงการ ได้แก่ นักประเมินต้องการประเมินประสิทธิผลของโครงการการส่งเสริมการรู้เท่าทันสื่อ (media literacy) ของครูสังกัดสำนักงานเขตพื้นที่การศึกษาในจังหวัดภาคเหนือตอนล่าง จำนวน 8 จังหวัด ได้แก่ จังหวัดพิษณุโลก สุโขทัย ตาก กำแพงเพชร อุตรดิตถ์ เพชรบูรณ์ พิษณุตร และนครสวรรค์ ซึ่งประสิทธิผลการรู้เท่าทันสื่อ นั้น พิจารณาหรือวัดได้จาก ความรู้ความเข้าใจในเนื้อหาสาระของสื่อ การเข้าถึงสื่อ และการสร้างองค์ความรู้จากสื่อ ทั้งนี้ นักประเมินได้ทำการศึกษาทฤษฎีแนวคิดและงานวิจัยหรืองานประเมินเกี่ยวกับการรู้เท่าทันสื่อของกลุ่มบุคคลต่างๆ ทั้งจากงานวิจัยในประเทศและต่างประเทศก็พบว่า มีตัวแปรแฝงและตัวแปรสังเกตได้ในแต่ละตัวแปรแฝงที่ส่งผลต่อการรู้เท่าทันสื่อทั้งเป็นการส่งผลทางตรงและทางอ้อม ดังนี้

ตัวแปรแฝง	ตัวแปรสังเกตได้	การส่งผลต่อการรู้เท่าทันสื่อ		
		ทางตรง	ทางอ้อม	ทั้งสองทาง
1. ฐานะทางเศรษฐกิจสังคม	1.1 รายได้/เดือน 1.2 ค่าใช้จ่ายสำหรับสื่อ 1.3 การมีสิ่งอำนวยความสะดวกในชีวิตประจำวัน			✓
2. พื้นภูมิหลัง	2.1 เพศ 2.2 อายุ 2.3 สถานภาพสมรส 2.4 คุณวุฒิและสาขาการศึกษาที่จบสูงสุด			✓

ตัวแปรแฝง	ตัวแปรสังเกตได้	การส่งผลต่อการรู้เท่าทันสื่อ		
		ทางตรง	ทางอ้อม	ทั้งสองทาง
3. การใช้เวลาว่าง	3.1 ท่องเที่ยว 3.2 พักผ่อนอยู่กับบ้าน/ ครอบครัว 3.3 ทำงานอดิเรก 3.4 ชมภาพยนตร์/ดนตรี 3.5 มากกว่า 1 กรณี	✓		
4. บุคลิกภาพ	4.1 รูปร่าง 4.2 นิสิต 4.3 อารมณ์		✓	

เมื่อให้  $Y$  : การรู้เท่าทันสื่อ

$y_1$  : ความรู้ความเข้าใจในเนื้อหาสาระของสื่อ

$y_2$  : การเข้าถึงสื่อ

$y_3$  : การสร้างองค์ความรู้จากสื่อ

$A$  : พันภูมิลัษ

$a_1$  : เพศ

$a_2$  : อายุ

$a_3$  : สถานภาพสมรส

$a_4$  : คุณวุฒิและสาขาการศึกษาที่จบสูงสุด

$B$  : ฐานะทางเศรษฐกิจสังคม

$b_1$  : รายได้/เดือน

$b_2$  : ค่าใช้จ่ายสำหรับสื่อ

$b_3$  : การมีสิ่งอำนวยความสะดวก

$C$  : การใช้เวลาว่าง

$c_1$  : ท่องเที่ยว

$c_2$  : พักผ่อนอยู่กับบ้าน/ครอบครัว



$C_3$  : ทำงานอดิเรก

$C_4$  : ชมภาพยนตร์/ดนตรี

$C_5$  : มากกว่า 1 กรณี

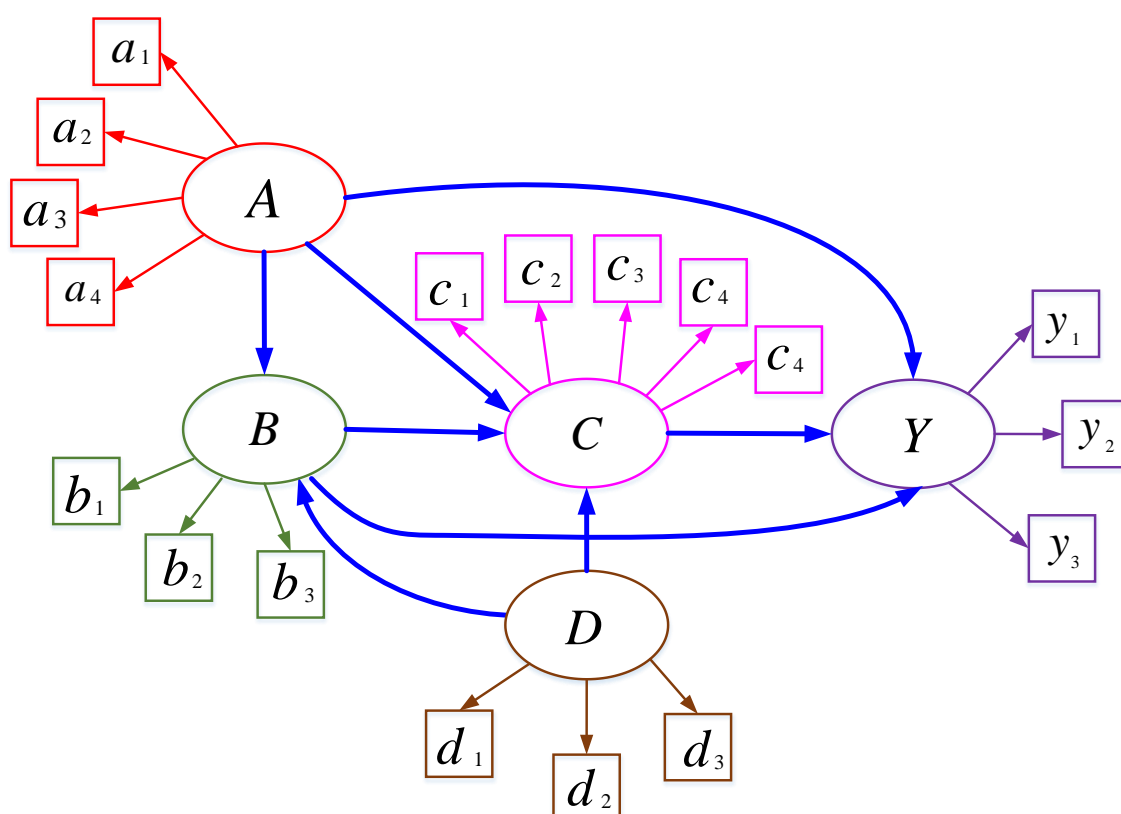
$D$  : บุคลิกภาพ

$d_1$  : รูปร่าง

$d_2$  : นิสัย

$d_3$  : อารมณ์

โมเดลเชิงสมมุติฐานที่แสดงโครงสร้างความสัมพันธ์ระหว่างตัวแปรแฝงกับตัวแปรแฝงและตัวแปรแฝงแต่ละตัวกับตัวแปรสังเกตได้ในตัวแปรแฝงหนึ่งๆ เขียนเป็นแผนภาพของโมเดลเชิงสมมุติฐานได้ดังนี้



จากโมเดลเชิงสมมุติฐานนี้ แสดงให้เห็นว่านักประเมินคิดว่า ตัวแปรแฝงพื้นฐาน (ซึ่งประกอบด้วย ตัวแปรสังเกตได้ เพศ อายุ สถานภาพสมรส และคุณวุฒิ/สาขาการศึกษาที่จบสูงสุด) ส่งผลทั้งทางตรงต่อประสิทธิผลการรู้เท่าทันสื่อของครู (ซึ่งวัดได้หรือประกอบด้วย ตัวแปรสังเกตได้ ความรู้ความเข้าใจในเนื้อหาสาระของสื่อ การเข้าถึงสื่อ และการสร้างองค์ความรู้จากสื่อ) และส่งผลทางอ้อมต่อประสิทธิผลการรู้เท่าทันสื่อผ่านตัวแปรแฝงการใช้เวลาว่างของครู (ซึ่งวัดได้จากตัวแปรสังเกตได้ การท่องเที่ยว พักผ่อนอยู่กับบ้าน/ครอบครัว การทำงานอดิเรก ชมภาพยนตร์/ดนตรี และกระทำมากกว่า 1 กรณี) ส่วนความสัมพันธ์ระหว่างตัวแปรแฝงอื่นๆ ก็อธิบายได้ในทำนองเดียวกันนี้ หลังจากเสนอโมเดลนี้แล้ว นักประเมินก็เก็บรวบรวมข้อมูลตามตัวแปรสังเกตได้จากครูในสังกัดเขตพื้นที่การศึกษาภาคเหนือตอนล่าง โดยสุ่มตัวอย่างครูมาจำนวนหนึ่ง (การใช้จำนวนตัวอย่างเท่าไรในการใช้วิธี LISREL นั้น สามารถศึกษาได้จากตำราที่กล่าวถึงผ่านมา) แล้วนำข้อมูลที่ได้อามาวิเคราะห์ทดสอบโมเดลเชิงสมมุติฐานดังกล่าวนี้ ถ้าผลการวิเคราะห์ทดสอบแสดงให้เห็นว่าข้อมูลเชิงประจักษ์ (หรือโมเดลเชิงประจักษ์) สอดคล้องกับโมเดลเชิงสมมุติฐาน (โดยพิจารณาค่าดัชนีตามเกณฑ์ต่างๆ ที่กล่าวมา) ก็แปลและสรุปผลการประเมินโดยชี้ให้เห็นว่าตัวแปรแฝงใดมีอิทธิพลหรือส่งผลต่อประสิทธิผลการรู้เท่าทันสื่อของครูมากที่สุด ทั้งที่เป็นการส่งผลรวม (ทางตรง+ทางอ้อม) และการส่งผลทางตรงหรือส่งผลทางอ้อมทำนองนี้ เป็นต้น

### สรุป

จากที่นำเสนอผ่านมาจะเห็นว่า การประเมินผลโครงการ หรือการเพิ่มความก้าวหน้าของศาสตร์การประเมินสามารถที่จะนำวิธีการเชิงปริมาณใหม่ๆ ที่เป็นวิธีการวิเคราะห์ทางสถิติมาใช้ในการศึกษาหาคำตอบสำหรับการประเมินได้อย่างลุ่มลึกทันสมัยวิธีการทางสถิติหรือวิธีการออกแบบการทดลองใช้โครงการใหม่ๆ จะช่วยขยายขอบเขตพรมแดนความรู้ของศาสตร์การประเมินได้ดียิ่งขึ้น ดังนั้น นักประเมินจึงจำเป็นที่จะต้องเสาะแสวงหาและนำวิธีการใหม่ๆ มาใช้ในการประเมินอยู่เสมอๆ การติดขัดกับวิธีการประเมินเดิมๆ แม้ว่าจะเป็นวิธีการที่ยอมรับนำมาใช้อย่างแพร่หลายนั้น ในด้านหนึ่งอาจทำให้ผลการประเมินน่าเชื่อถือได้ดี แต่ในอีกด้านหนึ่งก็เท่ากับปิดกั้นโอกาสการแสวงหาสิ่งใหม่ๆ มาเติมเต็มองค์ความรู้ของศาสตร์การประเมินไปโดยปริยาย

### เอกสารอ้างอิง

- นงลักษณ์ วิรัชชัย. (2542). **โมเดลริสเรลและสถิติวิเคราะห์สำหรับการวิจัยทางสังคมศาสตร์ และพฤติกรรมศาสตร์**. กรุงเทพมหานคร. โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2554). **การวิเคราะห์พหุระดับ**. กรุงเทพมหานคร. ภาควิชาวิจัยและจิตวิทยาการศึกษา คุรุศาสตร์จุฬาลงกรณ์มหาวิทยาลัย.
- Campbell, D and Stanley, J. (1963). **Experimental and quasi-experimental designs for research Chicago**. IL: Houghton-Mifflin.
- Clair, St. T., Cook, T.D. and Hallberg, K. (2014). **“Examining the Internal Validity and Statistical Precision of the Comparative Interrupted Time Series Design by Comparison With a Random ized Experiment”**. American Journal of Evaluation. 35(3), 311-327.
- Gebreselassie, T., Stephens, R. L., Maples, C. J., Johnson, S. F. and Tucker, A. L. (2014). **“Retention of Children and Their Families in the Longitudinal Outcome Study of the Comprehensive Community Mental Health Services for Children and Their Families Program: A Multilevel Analysis”**. American Journal of Evaluation. 35(2), 214-229.
- Goldstein, H. (2011). **Multilevel statistical models**. 4<sup>th</sup>, edition. West Sussex, England: Wiley.
- Hox, J. J. (2010). **Multilevel analysis. Techniques and applications**. NY.: Routledge.
- Kaplan, D. (2009). **Structural equation modeling: Foundations and extensions**. 2<sup>nd</sup>, edition. Thousand Oaks, CA.: Sage.
- Kerlinger, F. N. (1986). **Foundations of Behavioral Research**. 3<sup>rd</sup>, edition. NY.: Holt, Rinehart and Winston, Inc.
- Kline, R. B. (2011). **Principles and practice of structural equation modeling**. 3<sup>rd</sup>, edition. NY.: Guilford Press.

- Lochlin, J. C. (2004). **Latent variable models: An introduction to factor, path, and structural equation analysis**. 4<sup>th</sup>, edition. Mahwah, NJ.: Lawrence Erlbaum.
- Luke, D. A. (2004). **Multilevel Modeling**. Sage University Paper Series on Quantitative Applications in the Social Science (118.143). Thousand Oaks, CA.: Sage.
- Morgan, S. L. and Winship, C. (2007). **Counterfactuals and causal inference: Methods and principles for social research**. NY.: Cambridge University Press.
- Mulaik, S. A. (2009). **Linear causal Models and structural equations**. Boca Raton, FL.: CRC Press.
- Raudenbush, S. W. and Bryk, A. S. (2002). **Hierarchical Linear models: Applications and data analysis methods**. 2<sup>nd</sup>, edition. Thousand Oaks, CA.: Sage.
- Raykov, T. and Marcoulides, G. A. (2006). **A first course in structural equation modeling**. 2<sup>nd</sup>, edition. Mahwah, NJ.: Lawrence Erlbaum.
- Shadish, W. R., Cook, T. D. and Campbell, D. T. (2002). **Experimental and non-experimental design for generalized causal inference**. Belmont, CA.: Wadsworth.
- Shumaker, R. E. and Lomax, R. G. (2010). **A beginner & guide to structural equation modeling**. 3<sup>rd</sup>, edition, NY.: Routledge.
- Snijders, T. A. B. and Bosker, r> (2012). **Multilevel analysis: An introduction to basic and advanced multilevel modeling**. 2<sup>nd</sup>, edition. Thousand Oaks, CA.: Sage
- Zvoch, K. (2014). **“Modern quantitative methods for evaluation science: Recommendations for essential methodological text”**. American Journal of Evaluation. 35(3), 430–440.