



บทความ

เรื่อง

การแก้ไขอคติของมาตรฐานวัดการรับรู้ในการวัด
ตัวแปรนามธรรมโดยใช้แองเคอร์รีંગ วินเยตต์
(ตีพิมพ์ในวารสารวิชาการระดับประเทศ)

โดย

อัศรพงศ์ อันทอง และ ชยุตม์ ภิรมย์สมบัติ

กันยายน 2560

บทความ

เรื่อง

การแก้ไขข้อคิดของมาตรการรับรู้ในการวัด
ตัวแปรนามธรรมโดยใช้แองเคอร์รีง วินเยตต์
(ตีพิมพ์ในวารสารวิชาการระดับประเทศ)

โดย

อัครพงศ์ อันทอง

ชยุตม์ ภิรมย์สมบัติ

มหาวิทยาลัยแม่โจ้

จุฬาลงกรณ์มหาวิทยาลัย

ภายใต้ชุดโครงการวิจัย

ประสิทธิภาพและธรรมาภิบาลของนโยบายสาธารณะไทย

การวิจัยเพื่อเสริมสร้างน่านโยบายสาธารณะที่ดี

(The Efficiency and Government of Thai Public Policies)

(SRI59X0602)

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัย (สกว.)

(ความเห็นในรายงานนี้เป็นของผู้วิจัย สกว.ไม่จำเป็นต้องเห็นด้วยเสมอไป)

วารสารวิธีวิทยาการวิจัย
JOURNAL OF RESEARCH METHODOLOGY

ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

**The Department of Educational Research and Psychology,
Faculty of Education, Chulalongkorn University
ISSN 0857-2933**

Journal of Research Methodology

Senior Editors

Somwung Pitiyanuwat, Ph.D. (Prof.)	National Institute of Educational Testing Service
Teara Archawamety, Ph.D. (Prof.)	University of Nebraska at Kearney, USA

Editors-in-Chief

Suwimon Wongwanich, Ph.D. (Prof.)	Chulalongkorn University
Auyporn Ruengtrakul, Ph.D. (Assoc. Prof.)	Chulalongkorn University
Chayut Piromsombat, Ph.D.	Chulalongkorn University

Editorial Board

Arunee Onsawad, Ph.D. (Assoc. Prof.)	Naresuan University
Chatsiri Piyapimonsit, Ph.D. (Asst. Prof.)	Kasetsart University
Choosak Khamphalikit, Ph.D.	Srinakharinwirot University
Kamonwan Tangdhanakanond, Ph.D. (Assoc. Prof.)	Chulalongkorn University
Nonglak Wiratchai, Ph.D. (Prof.)	Chulalongkorn University
Pennee Narrot, Ph.D. (Assoc. Prof.)	Khon Kaen University
Piyapong Sumettikoon, Ph.D.	Chulalongkorn University
Sageemas Na Wichian, Ph.D. (Assoc. Prof.)	King Mongkut's University of Technology North Bangkok
Samphan Punprug, Ph.D. (Assoc. Prof.)	National Institute of Educational Testing Service
Samran Meejang, Ph.D. (Assoc. Prof.)	Naresuan University
Shotiga Pasiphol, Ph.D. (Assoc. Prof.)	Chulalongkorn University
Sirichai Karnjanawasee, Ph.D. (Prof.)	Chulalongkorn University
Siridej Sujiva, Ph.D. (Assoc. Prof.)	Chulalongkorn University
Somkit Promjouy, Ph.D. (Assoc. Prof.)	Sukhothai Thammathirat Open University
Sungworn Ngadkratoke, Ph.D. (Asst. Prof.)	Sukhothai Thammathirat Open University

Coordinators

Kanit Sriklaub, Ph.D.	Utumporn Chatiphuak
Surasak Kao-ian	

วารสารวิธีวิทยาการวิจัย
ปีที่ 29 ฉบับที่ 3 (กันยายน-ธันวาคม 2559)

สารบัญ

255

การวิเคราะห์องค์ประกอบเชิงยืนยันของดัชนีสุขภาพภาวะภาครัฐ:

การศึกษานำร่องโรงพยาบาลของรัฐ

ดวงเนตร ธรรมกุล ศิริพร ครูทภาศ จริยา ชื่นศิริมงคล

ยุพาวรรณ ทองตะนูนาม และ อัจฉรา ประเสริฐสิน

273

อิทธิพลของความสัมพันธ์ระหว่างมิติ และวิธีประมาณค่าความเที่ยงที่มีต่อค่าความเที่ยง

แบบพหุมิติ และประสิทธิภาพของการประมาณค่าความเที่ยง

วิภาวรรณ มัทขุนทด และ ญัฐภรณ์ หลาวทอง

295

ปัจจัยสำคัญสู่ความสำเร็จของการควบคุมคุณภาพทั้งองค์กร

และผลการดำเนินงานเมื่อนำมาปฏิบัติในมหาวิทยาลัย: การเสนอแบบจำลอง

อภิรักษ์ เอื้ออังกูร

315

การนำภูมิปัญญาวัฒนธรรมและทรัพยากรไทยมาสร้างประสบการณ์ในมุมการเรียนรู้

สำหรับเด็กเล็ก: การวิจัยระยะยาว

อัญมณี บุญเชื้อ และ สุรศักดิ์ เก้าเอี้ยน

343

การจัดการเรียนการสอนด้วยโปรแกรมสองภาษาและผลลัพธ์ทางการศึกษา:

นักเรียนที่ใช้ภาษาอังกฤษเป็นหลัก

Jenna Rycek and Tammi Ohmstede

359

โมเดลการวัดการร่วมมือรวมพลังของครู: การเปรียบเทียบโมเดลคู่แข่งชั้นสองโมเดล

เพียรกิจ นิมิตรดี ดวงกลม ไตรวิจิตรคุณ และ อวยพร เรืองตระกูล

375

การแก้ไขอคติของมาตรวัดการรับรู้ในการวัดตัวแปรนามธรรมโดยใช้แองเคอร์ริง วินเซนต์

อัศวพงศ์ อ้นทอง และ ชยุตม์ ภิรมย์สมบัติ

JOURNAL OF RESEARCH METHODOLOGY

Volume 29, Number 3 (November - December 2016)

CONTENT

255

**Confirmatory Factor Analysis of Happy-Public Organization Index:
A Pilot Study in Public Hospital**

*Doungnetre Thummakul, Siriporn Kruttakart, Jariya Cheunsirimongkol,
Yupawan Thongtanunam and Ujsara Prasertsin*

273

**Effect of Multidimensionality and Reliability Estimation Methods on
Multidimensional Reliability and Effect of Reliability Estimation**

Wiphaphan Mukkhuntod and Nuttaporn Lawthong

295

**Critical Success Factors of Total Quality Management Implementation and
Organizational performance in University Sector: A Proposed Model**

Apinan Aue-aungkul

315

**Organizing Thai Cultural Wisdom and Resources to Develop
Young Children's Experience in Learning Centers: Longitudinal Study**

Anyamanee Boonsue and Surasak Kao-ian

343

**Dual Language Program and Academic Outcomes:
Looking at the English Dominate Students**

Jenna Rycek and Tammi Ohmstede

359

A Measurement Model of Teacher Collaboration: Competitive Models

Phiankit Nimitdee, Duangkamol Traiwichitkhun and Auyporn Ruengtrakul

375

**Correcting Scale Perception Bias in Measuring Subjective Variables
Using Anchoring Vignettes**

Akarapong Untong and Chayut Piromsombat

Correcting Scale Perception Bias in Measuring Subjective Variables Using Anchoring Vignette

Akarapong Untong¹
Chayut Piromsombat²

ABSTRACT

This paper presents the correcting scale perception bias by using anchoring vignettes method that proposed by King et al. (2004). This method uses a standard set of data for adjusts the response scales of respondents upon two identifying assumptions: response consistency and vignette equivalence. The two methods for adjusting the response scales are the non-parametric method and parametric method. The non-parametric method is convenient and easy, while the parametric method is more efficient, and it can correct the differential item functioning (DIF-corrected) of the question that each respondent has different thresholds. The anchoring vignettes can be comparing the answers as ordinal scale, which assess by the respondents who are having different personal characteristic and culture as more correct and reliable. Therefore, it should be more applied anchoring vignettes method in Thailand; especially in the social research groups that are commonly use the questionnaires for measure subjective variables by rating scale.

Keywords: Scale Perception Bias, Subjective Variables, Anchoring Vignettes

¹ Corresponding author, School of Tourism, Maejo University, Chiang Mai 50200.
email: akarapong_un@hotmail.com

² Department of Educational Research and Psychology, Faculty of Education, Chulalongkorn University, Bangkok 10330. email: chayut.p@chula.ac.th

การแก้ไขอคติของมาตรวัดการรับรู้ในการวัดตัวแปรนามธรรม โดยใช้แองเคอร์ริง วินเยตต์

อัศรพงศ์ อ้นทอง¹
ชยุตม์ ภิรมย์สมบัติ²

บทคัดย่อ

บทความนี้เสนอการแก้ไขอคติของมาตรวัดการรับรู้ด้วยวิธีแองเคอร์ริง วินเยตต์ ที่เสนอโดย King et al. (2004) วิธีดังกล่าวใช้การเพิ่มชุดข้อมูลที่เป็นเกณฑ์มาตรฐานในการปรับค่าสเกลคำตอบของผู้ตอบภายใต้ข้อตกลงเบื้องต้นที่สำคัญ 2 ข้อ ได้แก่ ความสอดคล้องในการตอบและความเท่าเทียมกันของวินเยตต์ โดยมีวิธีปรับค่าสเกล 2 วิธี ได้แก่ วิธีที่ไม่อิงพารามิเตอร์ซึ่งมีความสะดวกและง่ายในการดำเนินการ และวิธีที่อิงพารามิเตอร์ซึ่งมีประสิทธิภาพมากกว่าและสามารถจัดการปัญหาอคติจากการทำหน้าที่ต่างกันของข้อคำถามที่เกิดจากผู้ตอบแต่ละคน มีขีดแบ่งในการให้ค่าสเกลคำตอบต่างกัน วิธีแองเคอร์ริง วินเยตต์ ช่วยให้สามารถนำคำตอบที่มีมาตรวัดแบบเรียงลำดับที่ประเมินจากผู้ตอบที่มีลักษณะเฉพาะส่วนบุคคลและ/วัฒนธรรมที่แตกต่างกันทางมาเปรียบเทียบได้อย่างถูกต้องและน่าเชื่อถือมากขึ้น ดังนั้นจึงควรมีการนำวิธีดังกล่าวมาประยุกต์ใช้มากขึ้นในประเทศไทย โดยเฉพาะในกลุ่มงานวิจัยด้านสังคมที่นิยมใช้แบบสอบถามในการวัดตัวแปรนามธรรมด้วยมาตรวัดประมาณค่า

คำสำคัญ: อคติของมาตรวัดการรับรู้ ตัวแปรนามธรรม แองเคอร์ริง วินเยตต์

¹ ผู้รับผิดชอบบทความหลัก คณะพัฒนาการท่องเที่ยว มหาวิทยาลัยแม่โจ้ เชียงใหม่ 50200

อีเมล: akarapong_un@hotmail.com

² ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร 10330

อีเมล: chayut.p@chula.ac.th

บทนำ

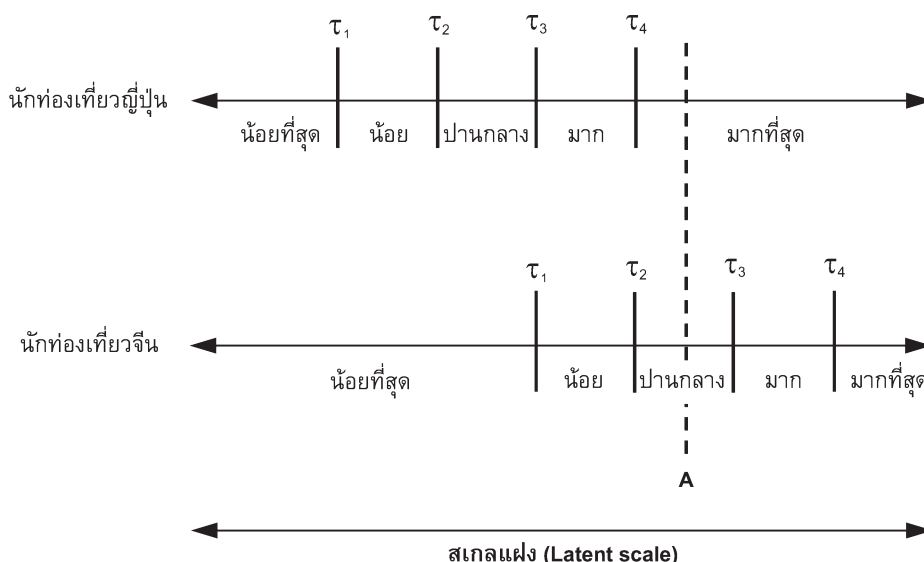
การศึกษาด้านสังคมศาสตร์มักต้องเผชิญกับตัวแปรทางทฤษฎีและ/แนวคิดที่มีลักษณะเป็นนามธรรมสูง โดยตัวแปรทางสังคมมักเป็นตัวแปรคุณภาพที่เป็นนามธรรม อย่างเช่น พฤติกรรมของบุคคล ทศนคติ ความคิดเห็น ความพึงพอใจ เป็นต้น และนิยมให้ผู้ตอบเป็นผู้ประเมินหรือให้คะแนน (Respondent-assessment) ซึ่งมาตรวัดนามธรรม (Subjective scales) ที่ผู้ตอบแต่ละคนให้คะแนน (หรือประเมิน) ขึ้นอยู่กับความเป็นจริง (Objective reality) ของสถานการณ์ อย่างเช่น ผลกระทบของการท่องเที่ยวที่มีต่อรายได้ของตนเอง ความพึงพอใจต่อบริการที่ได้รับจากสปา เป็นต้น และมีชุดของขีดแบ่งที่เป็นนามธรรม (Subjective thresholds) ที่มีลักษณะเฉพาะ อย่างเช่น ดีเยี่ยม ดีมาก ดี หรือ เห็นด้วยอย่างมาก เห็นด้วยมาก เห็นด้วย เป็นต้น (van Soest, Delaney, Harmon, Kapteyn and Smith, 2011) ซึ่งแตกต่างกันในแต่ละบุคคล (Lindeboom and van Doorslaer, 2004) ตามลักษณะเฉพาะที่แตกต่างกันของบุคคลและ/กลุ่มทางเศรษฐกิจและสังคม (Socio-economic) โดยในการวัด/ประเมินนิยมใช้มาตรวัดประมาณค่า (Rating scale) อย่าง Likert scale ในการวัด/ประเมินตัวแปรทางสังคมที่เป็นตัวแปรนามธรรม

ตัวแปรทางสังคมที่เป็นนามธรรมไม่อาจวัดด้วยคำถามเพียงข้อเดียว ควรวัดด้วยชุดคำถามที่สะท้อนและ/บ่งชี้ถึงตัวแปรที่ต้องการวัด ผู้ตอบแต่ละคนรับรู้และ/เข้าใจข้อคำถามและให้น้ำหนักคะแนนแตกต่างกัน นอกจากนี้ผู้ตอบแต่ละคนยังมีจุดตัด (Cut-points) หรือขีดแบ่ง (Thresholds) ที่แตกต่างกันตามลักษณะเฉพาะที่แตกต่างกันของบุคคล (อาจกล่าวว่าเป็นสโตร์ส่วนบุคคล) หรือวัฒนธรรม (Chevalier and Fielding, 2011) อย่างเช่น นักท่องเที่ยวญี่ปุ่นมักให้คะแนนความพึงพอใจสูงกว่านักท่องเที่ยวจีน เป็นต้น ดังนั้นจึงไม่สามารถเปรียบเทียบคำตอบที่ประเมินจากบุคคลที่มีลักษณะเฉพาะและ/วัฒนธรรมที่แตกต่างกันได้ เนื่องจากคำตอบที่ได้จากผู้ตอบที่มีลักษณะเฉพาะส่วนบุคคลและ/วัฒนธรรมที่แตกต่างกันอาจมีอคติจากการทำหน้าที่ต่างกันของข้อคำถาม (Differential Item Functioning: DIF) (Holland and Wainer, 1993)

ภาพ 1 เป็นตัวอย่างที่แสดงให้เห็นว่า ผู้ตอบที่มาจากวัฒนธรรมที่แตกต่างกัน (แทนด้วยประเทศ) อาจมีขีดแบ่งในการพิจารณาให้ค่าสเกลที่แตกต่างกัน โดยในตัวอย่างเป็นการประเมินความพึงพอใจของนักท่องเที่ยวที่มีต่อการให้บริการของพนักงานต้อนรับของโรงแรม ด้วยการสอบถามว่า “ท่านพึงพอใจต่อการให้บริการของพนักงานต้อนรับของโรงแรมในระดับใด” โดยให้นักท่องเที่ยวเป็นผู้ประเมินด้วยตนเอง สมมติว่านักท่องเที่ยวที่ตอบคำถามดังกล่าวมาจากประเทศญี่ปุ่น และสาธารณรัฐประชาชนจีน ซึ่งมีวัฒนธรรมที่แตกต่างกัน ดังนั้นนักท่องเที่ยวทั้งสองคนที่มาจากวัฒนธรรมที่แตกต่างกันอาจมีขีดแบ่งนามธรรมของการให้ค่าสเกลแตกต่างกัน (เรียกว่า Subjective thresholds) ดังแสดงในรูปที่ 1 จึงไม่สามารถนำค่าคะแนนความพึงพอใจที่ประเมินด้วยนักท่องเที่ยวทั้งสองคนมาเปรียบเทียบกันได้

อย่างเช่น ระดับความพึงพอใจ “มากที่สุด” ของนักท่องเที่ยวญี่ปุ่น มีระดับเท่ากับ “ปานกลาง” เมื่อเทียบกับนักท่องเที่ยวจีน (เปรียบเทียบผ่านเส้นตรง A) เป็นต้น

นอกจากนี้การวัดตัวแปรทางสังคมศาสตร์ที่เป็นตัวแปรนามธรรมด้วยมาตรวัดประมาณค่า มักเผชิญกับปัญหาที่สำคัญ คือ 1) การออกแบบคำถามที่ใช้ในการวัด ซึ่งต้องมีความชัดเจนและเป็นรูปธรรม และ 2) ผู้ตอบแต่ละคนมีความเข้าใจข้อคำถามเดียวกันแตกต่างกัน (King, Murray, Salomon and Tandon, 2004) ส่วนใหญ่แก้ไขปัญหาดังกล่าวด้วยการปรับปรุงการใช้ภาษาให้มีความชัดเจนและเป็นรูปธรรมมากขึ้น เพื่อเพิ่มความสามารถในการเปรียบเทียบความแตกต่างระหว่างบุคคล (มณีการ์ ชูทอง และชยุตม์ ภิรมณ์สมบัติ, 2558) รวมทั้งการแก้ไขปัญหการทำหน้าที่ต่างกันของข้อคำถาม (DIF) ซึ่งต้องตัดข้อคำถามที่มีค่า DIF สูงทิ้ง จึงทำให้สูญเสียข้อคำถามบางข้อ หากชุดคำถามที่ใช้วัด/บ่งชี้ตัวแปรนามธรรมเผชิญกับ DIF (มณีการ์ ชูทอง, 2557)



ที่มา: ดัดแปลงมาจากรูปที่ 1 ของ Wilgenburg (2010)

ภาพ 1 ความแตกต่างของขีดแบ่ง (Thresholds) ของคำตอบที่มีลักษณะเชิงกลุ่ม (Response category)

การวัดตัวแปรนามธรรมที่ให้ผู้ตอบประเมินด้วยตนเองมักเผชิญกับ Response bias ที่มีความอคติจากการที่ผู้ตอบตอบคำถามบนพื้นฐานของปัจจัยอย่างอื่นมากกว่าเนื้อหาของคำถามที่ใช้ในการสอบถาม อย่างเช่น ผู้ตอบตอบคำถามบนพื้นฐานที่ต้องการแสดงให้เห็นว่าตนเองเป็นคนดี

ของสังคมโดยไม่พิจารณาหรือคำนึงถึงเนื้อหาของคำถามที่ใช้สอบถาม เป็นต้น ซึ่งอาจกล่าวได้ว่า ผู้ตอบคำถามไม่ตรงกับความคิดที่แท้จริงของตนเอง แต่ถ้าหากผู้ตอบยังคงมีรูปแบบหรือลักษณะของการตอบที่สอดคล้องกันในทุกช่วงเวลาและ/สถานการณ์ อย่างเช่น ผู้ตอบนิยมเลือกที่จะตอบค่ากลางอยู่เสมอไม่ว่าจะตอบคำถามในประเด็นใด ก็จะมีลักษณะเช่นนี้ว่า Response style (Paulhus, 1991) หรือ Scale perception bias (Araña and Leon, 2013) ซึ่งขึ้นอยู่กับความแตกต่างตามลักษณะเฉพาะส่วนบุคคล สังคมและเศรษฐกิจ (Socio-economic) และวัฒนธรรม โดยผลของ Scale perception bias ทำให้ค่าส่วนเบี่ยงเบนมาตรฐานที่คำนวณได้ไม่แท้จริง และเกิดการบิดเบือนของความสัมพันธ์ระหว่างตัวแปรที่ใช้ในการศึกษา (Dolnicar and Grun, 2009) ทำให้ผลลัพธ์ที่ได้จากการวิเคราะห์ด้วยสถิติหรือเศรษฐมิติที่อยู่บนพื้นฐานของการวิเคราะห์สหสัมพันธ์ (Correlation analysis) เกิดอคติ (Araña and Leon, 2013) รวมทั้งยังทำให้เกิดความเข้าใจผิดในผลลัพธ์ที่ได้จากการเปรียบเทียบ (King et al., 2004)

โดยทั่วไปแบ่ง Response style ออกเป็น 4 รูปแบบ (Araña and Leon, 2013) คือ 1) ผู้ตอบนิยมเลือกที่จะตอบค่าสูงสุดหรือต่ำสุด (Extremes) (เรียกว่า Extreme Response Style: ERS) 2) ผู้ตอบเลือกที่จะตอบในเชิงบวกและหลีกเลี่ยงที่จะตอบในเชิงลบ (เรียกว่า Acquiescence Response Style: ARS) 3) ผู้ตอบนิยมเลือกที่จะตอบค่ากลางของสเกล (เรียกว่า Midpoint Responding: MPR) และ 4) ผู้ตอบเลือกที่จะตอบโดยพยายามให้ตนเองดูดีในสังคม (เรียกว่า Socially Desirable Response Style: SRS)

จากข้างต้นจะเห็นได้ว่า Response style/Scale perception bias เป็นอคติที่เกิดจากการที่ผู้ตอบมีมาตรฐานในการเลือกตอบที่แตกต่างกันระหว่างบุคคล (Interpersonal) หรือข้ามวัฒนธรรม (Cross-cultural) จึงไม่สามารถนำคำตอบมาเปรียบเทียบกันได้ ที่ผ่านมามีการเสนอแนวทางแก้ไขปัญหาดังกล่าวหลากหลายแนวทาง อย่างเช่น เสนอให้ระมัดระวังในการใช้คำในการตั้งคำถาม การประชุมกลุ่มย่อย (Focus groups) การซักถามถึงความรู้ความเข้าใจ (Cognitive debriefing) และเทคนิคอื่นๆ ซึ่งส่วนใหญ่มุ่งเน้นและให้ความสำคัญกับการปรับปรุงคำถามที่ใช้ในการสอบถาม (King and Wand, 2007; Wand, King and Lau, 2011) อย่างไรก็ตาม การดำเนินการต่างๆ เหล่านี้ยังไม่สามารถขจัดปัญหาที่อาจเกิดขึ้นจากความเข้าใจที่แตกต่างกันระหว่างผู้ตอบแต่ละกลุ่ม และการใช้ชุดคำตอบเชิงกลุ่มที่มีลักษณะแบบเรียงลำดับ (Ordinal response categories) อย่างเช่น เห็นด้วยมากที่สุด เห็นด้วยมาก เห็นด้วย เห็นด้วยน้อย เห็นด้วยน้อยที่สุด เป็นต้น (King and Wand, 2007) ซึ่งผู้ตอบแต่ละคนอาจมีจุดตัดหรือขีดแบ่ง (Cut-points/Thresholds) ในการพิจารณาที่จะให้ค่าสเกลคำตอบที่มีลักษณะเชิงกลุ่ม (Response category) แตกต่างกันระหว่างผู้ตอบแต่ละคน (Araña and Leon, 2013)

ดังนั้น King, et al. (2004) จึงเสนอวิธีที่เรียกว่า “แอ่งเคอร์ริง วินเยตต์ (Anchoring vignettes: AV)” โดยการเพิ่มเติมชุดข้อมูลเข้าไปจัดการกับปัญหาที่เกิดจากการมีมาตรฐานการตอบที่แตกต่างกันระหว่างบุคคล (Wand, King and Lau, 2011) กล่าวคือ แอ่งเคอร์ริง วินเยตต์ เป็นวิธีที่เพิ่มเติมชุดข้อมูล (หรือที่เรียกว่า วินเยตต์) เพื่อใช้ในการแก้ไข Scale perception bias ที่เป็นผลมาจากการที่ผู้ตอบมีมาตรฐานในการเลือกตอบที่แตกต่างกันตามความแตกต่างระหว่างบุคคล หรือวัฒนธรรม โดยชุดข้อมูลที่เพิ่มเติมเข้าไปเปรียบเสมือนสมอที่ชั้ยึด (Anchors) หรือเกณฑ์ (หรือหุมด) มาตรฐาน (Standard) ที่ใช้รับการตอบที่แตกต่างกันระหว่างบุคคลให้สามารถเปรียบเทียบระหว่างกันได้ ซึ่งต่อมาถูกนำมาประยุกต์ใช้อย่างกว้างขวางในหลายสาขาวิชา เช่น สาธารณสุข รัฐศาสตร์ เศรษฐศาสตร์ การศึกษา การท่องเที่ยว เป็นต้น โดยเฉพาะในกลุ่มงานวิจัยที่ใช้แบบสอบถามที่ให้ผู้ตอบเป็นผู้ประเมิน นอกจากนี้ยังมีการพิสูจน์ให้เห็นในเชิงประจักษ์ว่า วิธีแอ่งเคอร์ริง วินเยตต์ ยังสามารถช่วยบรรเทาปัญหา Sample selection bias ได้ (Paccagnella, 2011; 2016)

แอ่งเคอร์ริง วินเยตต์ (Anchoring Vignettes)

วิธีแอ่งเคอร์ริง วินเยตต์ นำเสนอครั้งแรก โดย King et al. (2002) (Rabe-Hesketh and Skrondal, 2002) ก่อนที่ได้รับการตีพิมพ์เผยแพร่อย่างเป็นทางการในวารสารวิชาการในปี ค.ศ. 2004 (King et al., 2004) และมีการขยายความเพิ่มเติมเพื่อให้เข้าใจง่ายขึ้นโดย King and Wand (2007) ในปี ค.ศ. 2007 วิธีนี้เป็นวิธีที่มีประสิทธิภาพในการตรวจสอบความแตกต่างของสเกลคำตอบที่ผู้ตอบประเมินด้วยตนเองภายใต้ความแตกต่างของกลุ่มทางเศรษฐกิจและสังคม หรือวัฒนธรรม ซึ่งช่วยให้การวิเคราะห์คำตอบที่มีลักษณะแบบเรียงลำดับ (Ordinal) ที่ได้จากการให้ผู้ตอบเป็นผู้ประเมิน มีการคำนึงถึงความแตกต่างระหว่างบุคคลที่ประเมิน (Paccagnella, 2016)

แอ่งเคอร์ริง วินเยตต์ ถูกนำมาใช้อย่างกว้างขวางในการศึกษาวิจัยที่ใช้แบบสอบถามให้ผู้ตอบเป็นผู้ประเมิน อย่างเช่น การประเมินตนเองด้านความมีประสิทธิภาพทางการเมือง (King et al., 2004) ด้านสุขภาพ (Bago d’Uva, van Doorslaer, Lindeboom and O’Donnell, 2008) ความพึงพอใจต่องานที่ทำ (Kristensen and Johansson, 2008) ความพึงพอใจในชีวิต (Angelini, Cavapozzi, Corazzine and Paccagnella, 2014) ความพึงพอใจต่อการท่องเที่ยว (Araña and Leon, 2013) เป็นต้น การศึกษาส่วนใหญ่แสดงให้เห็นว่า วิธีนี้สามารถใช้แก้ไขปัญหาคำถามที่ต่างกันของข้อคำถาม (Differential Item Functioning: DIF)¹ อันเป็นผลมาจากความแตกต่างระหว่างบุคคล

¹ DIF เป็นปัญหาที่เกิดจากการที่คำตอบที่ประเมินโดยผู้ตอบได้รับอิทธิพลจากความแตกต่างของคุณลักษณะเฉพาะส่วนบุคคล อย่างเช่น ความแตกต่างทางด้านวัฒนธรรมของนักท่องเที่ยวมีผลทำให้การประเมินความพึงพอใจต่อการให้บริการแตกต่างกัน หรือความแตกต่างทางด้านฐานะทางเศรษฐกิจของนักท่องเที่ยวนำมาซึ่งการประเมินคุณค่าที่ได้รับจากการท่องเที่ยวแตกต่างกัน เป็นต้น

ที่มีพื้นฐานทางด้านเศรษฐกิจและสังคม รวมทั้งวัฒนธรรมและประสบการณ์ที่แตกต่างกัน อย่างไรก็ตาม แองเคอร์ริง วินเยตต์ ถูกออกแบบมาให้ใช้แก้ไขปัญหาเฉพาะกรณีของ Response-category DIF ไม่ใช่สำหรับทุกรูปแบบของ DIF (King and Wand, 2007)

แนวคิดของแองเคอร์ริง วินเยตต์ มีพื้นฐานมาจากเปรียบเทียบสเกลคำตอบที่ผู้ตอบเป็นผู้ประเมิน กับเกณฑ์มาตรฐาน (Standard) ที่เป็นสมอ (Anchor) ที่ได้จากคำตอบของชุดคำถามวินเยตต์ (Vignettes' questions) ซึ่งอยู่ในการสำรวจเดียวกัน จากพื้นฐานแนวคิดดังกล่าว การประยุกต์ใช้วิธีแองเคอร์ริง วินเยตต์ ประกอบด้วยขั้นตอนที่สำคัญ 2 ขั้นตอน คือ 1) การออกแบบชุดคำถามวินเยตต์เพื่อใช้เป็นเกณฑ์มาตรฐาน (Standard) สำหรับเป็นตัววัดในการเปรียบเทียบกับสเกลคำตอบที่ได้จากการประเมินด้วยผู้ตอบ และ 2) เทคนิคการปรับสเกลคำตอบที่ผู้ตอบเป็นผู้ประเมินโดยใช้สเกลคำตอบมาตรฐานเป็นจุดยึด (Anchor) ซึ่งมีทั้งวิธีที่ไม่อิงพารามิเตอร์ (Non-parametric) และวิธีที่อิงพารามิเตอร์ (Parametric) (Wand, King and Lau, 2011) โดยจะขอลงถึงในแต่ละส่วนพอสังเขปดังนี้

1) การออกแบบชุดคำถามวินเยตต์

วินเยตต์ (Vignette) มีรากศัพท์มาจากภาษาฝรั่งเศส หมายถึง รูปภาพ หรือข้อความสั้น ๆ ที่พรรณนาให้เห็นภาพ ซึ่งในการออกแบบวินเยตต์เพื่อใช้เป็นสมอสำหรับยึดคำตอบที่ผู้ตอบที่แตกต่างกันประเมิน ควรมีการให้ข้อมูลเท่าที่จำเป็นสำหรับผู้ตอบได้คิดวิเคราะห์ แต่ไม่มากเกินไปจนทำให้ผู้ตอบไม่อาจเข้าใจสถานการณ์ที่สร้างขึ้น ดังนั้นวินเยตต์จึงควรเป็นข้อความสั้น ๆ ที่ใช้ภาษาที่มีความเป็นรูปธรรมและชัดเจนโดยพรรณนาให้ผู้ตอบเห็นภาพสถานการณ์ที่แสดงคุณลักษณะของตัวแปรที่ต้องการวัดในระดับที่แตกต่างกันอย่างชัดเจนภายในชุดของวินเยตต์แต่ละชุด เพื่อให้ผู้ตอบสามารถจำแนกและเรียงลำดับวินเยตต์ในแต่ละชุดได้สอดคล้องกับที่ผู้วิจัยและ/ผู้เชี่ยวชาญกำหนด

ก่อนการออกแบบวินเยตต์ผู้วิจัยต้องกำหนดกรอบแนวคิด ทำความเข้าใจ และระบุให้ชัดเจนว่า ในการศึกษาต้องการวัดและ/ประเมินอะไร (กำหนดตัวแปรที่ต้องการวัดและ/ประเมินให้ชัดเจน) ใช้มาตรวัดแบบใดในการวัด (หากต้องการใช้วิธีแองเคอร์ริง วินเยตต์ ในการปรับค่าสเกลของคำตอบควรใช้มาตรวัดประมาณค่า) เพื่อให้สถานการณ์ในวินเยตต์วัดและ/ประเมินภายใต้แนวคิดเดียวกับสิ่งที่ต้องการศึกษา นอกจากนี้หลังจากการออกแบบวินเยตต์เสร็จแล้วควรให้ผู้เชี่ยวชาญพิจารณาความตรงเชิงเนื้อหาและความสอดคล้องของการเรียงลำดับวินเยตต์ ก่อนนำไปทดสอบกับชุดตัวอย่างที่สนใจศึกษา (มณีการ์ ชูทอง, 2557) ประมาณ 30-50 ตัวอย่าง เพื่อตรวจสอบว่าการเรียงลำดับวินเยตต์ของชุดตัวอย่างกับที่กำหนดโดยผู้วิจัยและผู้เชี่ยวชาญสอดคล้องกันหรือไม่ หากไม่สอดคล้องหรือมีการเรียงลำดับที่ขัดแย้งกัน หรือชุดตัวอย่างเข้าใจสถานการณ์วินเยตต์แตกต่าง

จากผู้วิจัยและผู้เชี่ยวชาญ ผู้วิจัยต้องปรับปรุงเนื้อหาของวินเยตต์และนำไปทดสอบอีกครั้ง จนสัดส่วนของผู้ที่เรียงลำดับวินเยตต์ถูกต้องมีประมาณร้อยละ 60-80 ของผู้ตอบทั้งหมด (ไม่มีกฎเกณฑ์ตายตัว ขึ้นอยู่ดุลยพินิจของผู้วิจัย)

สำหรับออกแบบวินเยตต์ไม่มีหลักเกณฑ์หรือแนวทางที่แน่นอนขึ้นอยู่กับประสบการณ์ของผู้ทำวิจัย ในที่นี้จะขอนำเสนอแนวทางในการออกแบบวินเยตต์แบบกว้างๆ พอสังเขปดังนี้

- วินเยตต์ควรเป็นการอธิบายสถานการณ์ที่เป็นรูปธรรม ใช้ภาษาที่เข้าใจง่ายและชัดเจน โดยคนที่มีความสามารถต่างกันสามารถเข้าใจได้ตรงกัน ซึ่งสามารถออกแบบโดยการยกตัวอย่างเป็นสถานการณ์ หรือพฤติกรรมของบุคคลก็ได้ อย่างเช่น ในกรณีตัวอย่างที่มีวัตถุประสงค์ของการวิจัย เพื่อประเมินทัศนคติของคนท้องถิ่นต่อผลกระทบที่เกิดจากการพัฒนาท่องเที่ยวในชุมชน สามารถออกแบบวินเยตต์ในลักษณะของการยกตัวอย่างสถานการณ์ 3 สถานการณ์ ที่มีระดับของขนาดผลกระทบที่เกิดจากการพัฒนาการท่องเที่ยวในชุมชนที่แตกต่างกัน 3 ระดับ ($V_1 < V_2 < V_3$) ดังนี้

การมาเที่ยวของนักท่องเที่ยว**สร้างประโยชน์**ต่อชุมชนต่อไปในระดับใด

V_1 ชุมชน ก: มีนักท่องเที่ยวพลุกพล่าน ชาวบ้านขายของได้มากและมีงานทำมากขึ้น แต่รำคาญที่นักท่องเที่ยวส่งเสียงดังและทิ้งขยะเกลื่อนกลาดไปทั่ว

น้อย	มาก	มากที่สุด	น้อยที่สุด
------	-----	-----------	------------

V_2 ชุมชน ข: มีนักท่องเที่ยวเข้ามามาก ชาวบ้านขายของรายได้ดี แม้มีขยะเพิ่มขึ้นมาก แต่คนในชุมชนช่วยกันเก็บขยะ ทำให้ยังสามารถใช้ชีวิตได้ตามปกติ

น้อยที่สุด	น้อย	มาก	มากที่สุด
------------	------	-----	-----------

V_3 ชุมชน ค: ชุมชนมีความครึกครื้นจากนักท่องเที่ยว ชาวบ้านกลับมาทำงานภายในชุมชนมากขึ้น แต่ต้องนำรายได้ส่วนหนึ่งจากการท่องเที่ยวมาจัดเก็บขยะ

น้อยที่สุด	น้อย	มาก	มากที่สุด
------------	------	-----	-----------

หรือในกรณีเดียวกัน แต่ออกแบบวินเยตต์ในลักษณะของการยกตัวอย่างบุคคล 3 สถานการณ์ ที่มีระดับของการได้รับผลกระทบจากการพัฒนาการท่องเที่ยวในชุมชนที่แตกต่างกัน 3 ระดับ ($V_2 < V_3 < V_1$) ดังนี้

บุคคลต่อไปนี้**ได้รับผลกระทบหรือประโยชน์**จากการท่องเที่ยวที่เกิดขึ้นภายในชุมชนในระดับใด

V₁ นาย ก เปิดร้านขายอาหารในหมู่บ้าน ซึ่งมีรายได้เพิ่มขึ้นจากการขายอาหารให้นักท่องเที่ยว และทุกๆ วันเสาร์ ภริยายังไปตั้งแผงขายของที่ระลึกที่ถนนคนเดินซึ่งจัดขึ้นภายในชุมชน

น้อยที่สุด	น้อย	มาก	มากที่สุด
------------	------	-----	-----------

V₂ นาย ข เป็นชาวไร่ชาวนา ซึ่งออกไปทำไร่และสวนนอกหมู่บ้านทุกวัน ทุกๆ วันเสาร์จะเห็นนักท่องเที่ยวเดินพลุกพล่านในหมู่บ้าน เนื่องจากมีถนนคนเดินภายในชุมชน

น้อยที่สุด	น้อย	มาก	มากที่สุด
------------	------	-----	-----------

V₃ นาย ค เป็นผู้สูงอายุที่มีความสามารถด้านดนตรีพื้นบ้าน ทุกๆ วันเสาร์จะไปแสดงดนตรีที่ถนนคนเดินซึ่งจัดภายในชุมชน และมีรายได้เล็กน้อยจากการบริจาคของนักท่องเที่ยวที่เข้ามาดู

น้อยที่สุด	น้อย	มาก	มากที่สุด
------------	------	-----	-----------

- สถานการณ์ในแต่ละวินยเขตต์ไม่ควรเป็นสถานการณ์สุดโต่ง (Outlier) ที่ดีที่สุดหรือแย่ที่สุด ซึ่งผู้ตอบสามารถตอบได้ทันทีโดยไม่ต้องพิจารณาหรือหยุดคิด หรือเป็นสถานการณ์ที่ทำให้ผู้ตอบทั้งหมดสามารถให้ค่าสเกลสูงสุดหรือต่ำสุด (Wand, 2013) (ทั้งนี้จะขอขยายความเพิ่มเติมเพื่อให้เห็นถึงปัญหาในประเด็นนี้ในหัวข้อของกระบวนการประมาณค่า) ในแต่ละวินยเขตต์ควรประกอบด้วยคุณลักษณะ (Attribute) ของสิ่งที่ต้องการวัด/ประเมินมากกว่า 1 คุณลักษณะ และในแต่ละชุดของวินยเขตต์ (ควรมีวินยเขตต์มากกว่า 1 วินยเขตต์) ควรมีระดับที่แตกต่างกัน ซึ่งผู้ตอบโดยเฉลี่ย (หรือส่วนใหญ่) สามารถเรียงลำดับไปในทิศทางเดียวกัน

- วินยเขตต์แต่ละชุดควรวัด/ประเมินภายใต้แนวคิดในมิติเดียว (Single unidimensional concept) และควรเป็นเรื่องเดียว (หรือแนวคิดเดียว) กับคำถามที่ให้ผู้ตอบเป็นผู้ประเมิน นอกจากนี้ควรมีมาตรวัดในระดับเดียวกับคำถามที่ให้ผู้ตอบเป็นผู้ประเมิน (เป็นเงื่อนไขจำเป็น: Necessary condition) อย่างเช่น ในกรณีของการศึกษาที่ต้องการประเมินความพึงพอใจของนักท่องเที่ยวต่อบริการที่ได้รับ ผู้วิจัยกำหนดสเกลของมาตรวัดไว้ 4 ระดับ คือ พึงพอใจมากที่สุด พึงพอใจมาก พึงพอใจน้อย และพึงพอใจน้อยที่สุด ดังนั้นการออกแบบวินยเขตต์ก็ควรเป็นสถานการณ์ที่ประเมินความพึงพอใจต่อบริการโดยกำหนดสเกลของมาตรวัดไว้ 4 ระดับ เช่นเดียวกัน เป็นต้น

- ควรระมัดระวังการให้รายละเอียดที่มากเกินไปในการออกแบบสถานการณ์สมมติที่ใช้เป็นวินยเขตต์ อย่างเช่น การระบุเพศ อายุ อาชีพ หน่วยงาน เป็นต้น เนื่องจากอาจทำให้เกิดอคติ

ในการตอบ ผู้ตอบอาจตีความโดยใช้การอ้างอิงจากประสบการณ์หรือข้อมูลอื่น ๆ ที่ไม่ใช่ข้อมูลจากสถานการณ์สมมติที่กำหนดขึ้น หรือผู้ตอบอาจสับสนและเพิ่มเติมมิติของการวิเคราะห์เกินกว่ามิติที่ต้องการวัด/ประเมินในสถานการณ์สมมติที่กำหนดขึ้น

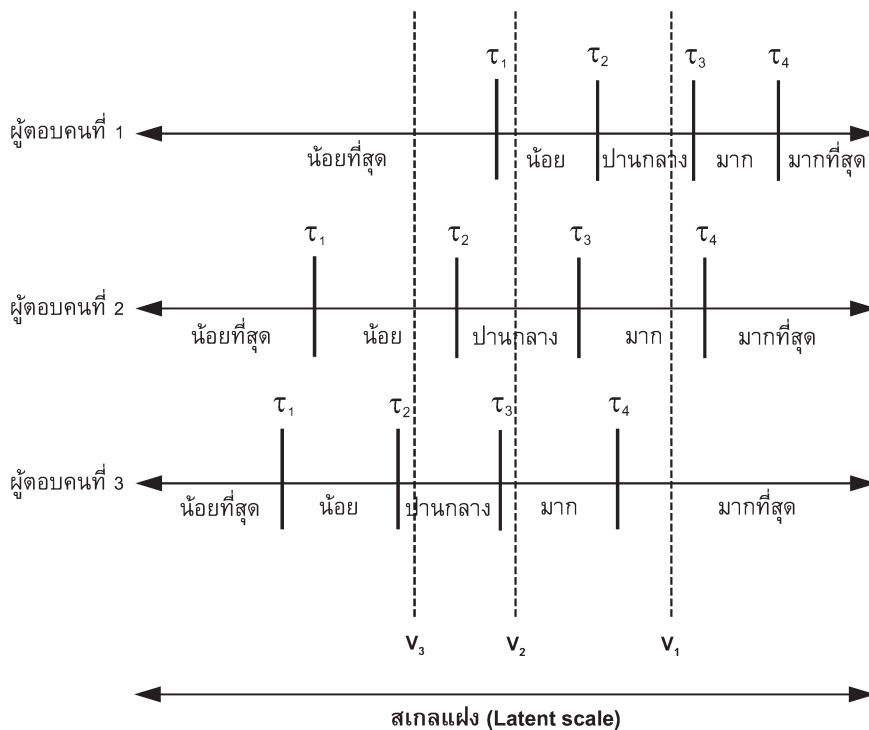
งานศึกษาที่ผ่านมายังไม่มีข้อสรุปว่า ในแต่ละวินเยตต์ควรมีคุณลักษณะที่ต้องการวัด/ประเมินผลสมอยู่ที่คุณลักษณะ (ในความเห็นของผู้เขียนแต่ละวินเยตต์ควรมีคุณลักษณะอย่างน้อย 3 คุณลักษณะ) ควรมีวินเยตต์อย่างน้อยกี่สถานการณ์ ในแต่ละมิติที่ต้องการวัด (ในความเห็นของผู้เขียนควรมีวินเยตต์อย่างน้อย 3 สถานการณ์) และควรใช้ชุดวินเยตต์เดียวกันทั้งแบบสอบถามหรือแยกตามมิติที่ต้องการวัดในแต่ละด้าน ควรสอบถามคำถามวินเยตต์ก่อนหรือหลังชุดคำถามที่ต้องการให้ผู้ตอบเป็นผู้ประเมิน (King et al. (2004) แนะนำให้สอบถามข้อคำถามที่ต้องการให้ผู้ตอบเป็นผู้ประเมินก่อนให้ผู้ตอบตอบคำถามวินเยตต์ อย่างไรก็ตาม การศึกษาผ่านมานิยมใช้การประชุมกลุ่มย่อย และการซักถามถึงความรู้ความเข้าใจ (Cognitive debriefing) กับผู้ตอบในขั้นตอนของการทดสอบวินเยตต์ รวมทั้งการนำวินเยตต์ไปทดสอบกับชุดตัวอย่าง เพื่อนำข้อมูลที่ได้มาปรับปรุงแก้ไขวินเยตต์ให้มีความตรงเชิงเนื้อหาและความสอดคล้องในการเรียงลำดับวินเยตต์ รวมทั้งการพิจารณาดัชนีวินเยตต์หรือคุณลักษณะบางรายการที่ไม่เกี่ยวข้องออกจากวินเยตต์ก่อนนำไปใช้จริง

ทั้งนี้เพื่อให้สามารถนำวินเยตต์มาใช้เป็นเกณฑ์มาตรฐาน (หรือจุดยึด) ในการปรับค่าสเกลคำตอบที่ผู้ตอบเป็นผู้ประเมิน จึงมีข้อตกลงเบื้องต้นในการวัด (Measurement assumptions) ที่สำคัญ 2 ข้อ ดังนี้ (King et al., 2004)

1. ความสอดคล้องในการตอบ (Response consistency) กล่าวคือ ผู้ตอบแต่ละคนมีระดับสเกลของการตอบเหมือนกันไม่ว่าจะเป็นการตอบในชุดคำถามวินเยตต์ หรือชุดคำถามที่ผู้ตอบเป็นผู้ประเมิน จากรูปที่ 2 ความสอดคล้องในการตอบ คือ ซีดแบ่งที่ 1 (τ_1) ถึง ซีดแบ่งที่ 4 (τ_4) ของผู้ตอบแต่ละคน ซึ่งเป็นซีดแบ่งที่เป็นนามธรรมภายในแต่ละบุคคล (Subjective thresholds) ซึ่งต้องอยู่ในตำแหน่งเดียวกันเมื่อผู้ตอบแต่ละคนตอบวินเยตต์และตอบคำถามที่ผู้ตอบเป็นผู้ประเมิน หมายความว่า ผู้ตอบแต่ละคนตอบชุดคำถามวินเยตต์และชุดคำถามที่ผู้ตอบเป็นผู้ประเมินภายใต้ซีดแบ่งที่เป็นนามธรรมเดียวกัน อย่างเช่นในกรณีของรูปที่ 2 หากไม่มีการละเมิดข้อตกลงความสอดคล้องในการตอบ ผู้ตอบคนที่ 1 ใช้ซีดแบ่งเดียวกันในการตอบวินเยตต์และคำถามที่ผู้ตอบเป็นผู้ประเมิน แต่ถ้าข้อตกลงความสอดคล้องในการตอบถูกละเมิด ผู้ตอบคนที่ 1 ใช้ซีดแบ่งของตนเองในการตอบวินเยตต์ แต่ใช้ซีดแบ่งของผู้ตอบคนที่ 2 หรือ 3 ในการตอบคำถามที่ผู้ตอบเป็นผู้ประเมิน เป็นต้น

2. ความเท่าเทียมกันของวินเยตต์ (Vignette equivalence) กล่าวคือ ในแต่ละวินเยตต์ ผู้ตอบรับรู้และประเมินระดับของตัวแปรนามมาตรวัดแบบมิติเดียว (Unidimensional scale) ที่เป็นไป

ในทางเดียวกัน แม้ว่าวินเยตต์แต่ละสถานการณ์จะเป็นอิสระจากกัน แต่ต้องมีการเรียงลำดับวินเยตต์เหมือนกัน โดยความแตกต่างที่เกิดขึ้นจากการให้ค่าสเกลวินเยตต์เป็นความคลาดเคลื่อนจากการวัดแบบสุ่ม (Random measurement error) จากรูปที่ 2 ความเท่าเทียมกันของวินเยตต์ แสดงโดยเส้นตรง (เส้น V_1-V_3) ที่ลากผ่านสเกลแฝง (Latent scale) ของผู้ตอบทั้งสามคน ภายใต้ข้อตกลงที่ว่าผู้ตอบแต่ละคนเรียงลำดับวินเยตต์เหมือนกันในลักษณะ $V_1 > V_2 > V_3$ หมายความว่า ผู้ตอบทั้งสามคนให้ค่าสเกลของวินเยตต์ที่มีการเรียงลำดับวินเยตต์เหมือนกัน ($V_1 > V_2 > V_3$) หากข้อตกลงนี้ถูกละเมิด เส้น V_1-V_3 ที่ลากผ่านผู้ตอบทั้งสามคนไม่เป็นเส้นตรงตามที่แสดงในรูปที่ 2 แสดงว่าผู้ตอบบางคนให้ค่าสเกลที่ทำให้วินเยตต์เรียงลำดับที่ผิดจากคนอื่น ๆ (Grol-Prokopczyk, 2014) อย่างเช่น จากรูปที่ 2 หากผู้ตอบคนที่ 2 ให้ค่าสเกลวินเยตต์ที่มีการเรียงลำดับในลักษณะ $V_2 > V_1 > V_3$ ก็จะทำให้เส้น V_1 และ V_2 ไม่เป็นเส้นตรงเหมือนกับที่แสดงในภาพ 2



หมายเหตุ: V_1-V_3 คือ วินเยตต์ที่ 1 ถึง 3 โดยที่ $V_1 > V_2 > V_3$

ที่มา: ดัดแปลงมาจากรูปที่ 1 ของ Grol-Prokopczyk (2014)

ภาพ 2 การใช้แองเคอร์เรียง วินเยตต์เปรียบเทียบคำตอบที่แตกต่างกันระหว่างบุคคล

ความสอดคล้องในการตอบมักถูกประเมินโดยอ้อมบนพื้นฐานของความน่าเชื่อถือของ Vignette-adjusted (Grol-Prokopczyk, 2014) โดยการเปรียบเทียบจุดตัดหรือขีดแบ่ง (Cut-points/ Thresholds) ที่ได้จากการให้ค่า (Rating) ของผู้ตอบต่อคำถามที่เป็นวินยัตต์และต่อคำถามที่ผู้ตอบเป็นผู้ประเมินกับผลของการวัดในเชิงรูปธรรม (Objective measures) อย่างเช่น ในงานศึกษาของ Bago d'Uva, Lindeboom, O'Donnell and Doorslaer (2011) Grol-Prokopczyk, Freese and Hauser (2011) van Soest et al., (2011) เป็นต้น ขณะทำงานศึกษา Kapteyn, Smith, van Soest and Yonkova (2011) ใช้วิธีที่ไม่อิงพารามิเตอร์ (Wilcoxon test และ Sign test) และวิธีที่อิงพารามิเตอร์ (แบบจำลอง CHOPIT) ตรวจสอบความสอดคล้องในการตอบแล้วพบว่า วิธีทั้งสองให้ผลการทดสอบแตกต่างกัน และให้ข้อเสนอว่า วินยัตต์ควรมีรายละเอียดที่ชัดเจนรัดกุม และสมบูรณ์เพื่อลดการตีความที่แตกต่างกันระหว่างผู้ตอบ นอกจากนี้ในงานศึกษาของ Angelini et al. (2014) ได้วิเคราะห์สหสัมพันธ์ (Correlation) ระหว่างค่าคะแนนที่ผู้ตอบเป็นผู้ประเมินกับวินยัตต์ แล้วพบว่า ผลลัพธ์ที่ได้สนับสนุนข้อตกลงเบื้องต้นในเรื่องของความสอดคล้องในการตอบ เช่นเดียวกับกับ Crane, Rissel, Greaves and Gebel (2016) ก็ได้ประยุกต์ใช้ Spearman's rank correlation ตรวจสอบข้อตกลงเบื้องต้นของความสอดคล้องในการตอบ

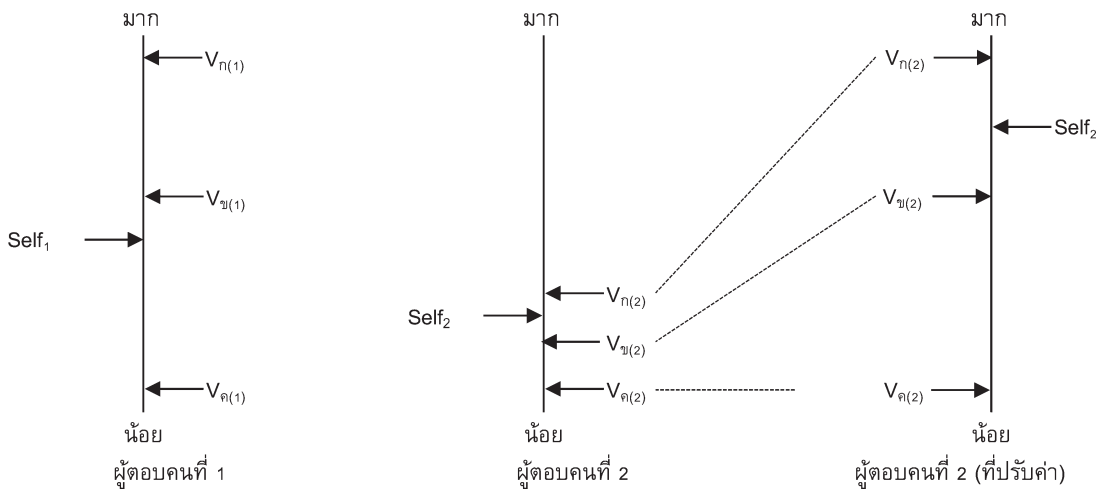
สำหรับการตรวจสอบข้อตกลงเบื้องต้นเกี่ยวกับความเท่าเทียมกันของวินยัตต์ ในเบื้องต้นพิจารณาจากการเรียงลำดับวินยัตต์ที่ถูกต้องตามการรับรู้โดยทั่วไปของบุคคลที่ผ่านการพิจารณาโดยผู้เชี่ยวชาญ กล่าวคือ ผู้ตอบโดยเฉลี่ย (หรือส่วนใหญ่) เรียงลำดับวินยัตต์ถูกต้องสอดคล้องกับการรับรู้โดยทั่วไป หรือบรรทัดฐาน (Norm) ของสังคมที่ผ่านการพิจารณาจากผู้เชี่ยวชาญ ซึ่งยังไม่มีกฎหรือหลักเกณฑ์ตายตัวในการกำหนดสัดส่วนของผู้ตอบที่เรียงลำดับวินยัตต์ถูกต้องตามที่กำหนดโดยผู้เชี่ยวชาญ อย่างไรก็ตาม ในงานศึกษาของ Araña and Leon (2013) ใช้สัดส่วนที่เรียงลำดับวินยัตต์ถูกมากกว่าร้อยละ 70 ของผู้ตอบทั้งหมด เป็นเกณฑ์พิจารณาสำหรับตรวจสอบข้อตกลงเบื้องต้นเกี่ยวกับความเท่ากันของวินยัตต์ (มีข้อสังเกตว่า เกณฑ์ดังกล่าวตกอยู่ในช่วงความเชื่อมั่นทางสถิติมากกว่าร้อยละ 95 ภายใต้การกระจายแบบปกติมาตรฐานของชุดตัวอย่าง) นอกจากนี้ Araña and Leon (2013) ยังประยุกต์ใช้ Spearman's rank correlation ตรวจสอบความสัมพันธ์ระหว่างลำดับของวินยัตต์ของผู้ตอบแต่ละคนกับชุดตัวอย่างทั้งหมด

นอกจากนี้การศึกษาที่ผ่านมายังไม่มีบทสรุปที่กำหนดเป็นกฎหรือหลักเกณฑ์ตายตัว สำหรับการออกแบบวินยัตต์เพื่อไม่ให้ละเมิดข้อตกลงเบื้องต้นทั้งสอง มีเพียงการเสนอให้ใช้แนวทางเชิงคุณภาพในการขั้นตอนของการทดสอบวินยัตต์ อย่างเช่น การประชุมกลุ่มย่อย การสัมภาษณ์ผู้ตอบแบบตัวต่อตัว เป็นต้น (Araña and Leon, 2013) อย่างไรก็ตาม การศึกษาในอดีตที่ประเมินและทดสอบความน่าเชื่อถือของวินยัตต์ แสดงให้เห็นว่า ค่าคะแนนที่ประเมินตนเองที่ปรับด้วยวินยัตต์

มีความใกล้เคียงกับค่าที่ได้จากการวัดในเชิงรูปธรรมมากกว่าค่าคะแนนที่ไม่ได้ปรับด้วยวินยเกตต์ (Grol-Prokopczyk, 2014)

2) กระบวนการประมาณค่า

จากข้อตกลงเบื้องต้นทั้งสองข้อตกลง สามารถนำวินยเกตต์มาปรับสเกลคำตอบที่แตกต่างกันระหว่างบุคคล ดังตัวอย่างที่แสดงในภาพ 3 โดยสมมติให้มีผู้ตอบจำนวน 2 คน แต่ละคนต้องตอบคำถามเกี่ยวกับการประเมินการรับรู้ผลกระทบจากการท่องเที่ยวที่ผู้ตอบเป็นผู้ประเมินจำนวน 1 ข้อ (Self) ที่เป็นมาตรวัดประมาณค่า (Rating scale) และวินยเกตต์จำนวน 3 สถานการณ์ (V_n , V_u และ V_c) โดยผู้ตอบทั้งสองคนเรียงลำดับวินยเกตต์เหมือนกัน ($V_n > V_u > V_c$) หากเปรียบเทียบคำตอบก่อนการปรับสเกลของผู้ตอบทั้งสองคน พบว่า ผู้ตอบคนที่ 1 (ภาพทางซ้าย) ให้ระดับคะแนนการรับรู้ผลกระทบจากการท่องเที่ยวสูงกว่าผู้ตอบคนที่ 2 (ภาพตรงกลาง) แต่ถ้าพิจารณาเปรียบเทียบกับวินยเกตต์ พบว่า ระดับคะแนนของผู้ตอบคนที่ 1 มีค่าต่ำกว่า V_u ส่วนระดับคะแนนของผู้ตอบคนที่ 2 มีค่าสูงกว่า V_u ดังนั้นเมื่อปรับระดับสเกลของผู้ตอบคนที่ 2 ให้เท่ากับผู้ตอบคนที่ 1 ดังภาพที่แสดงทางด้านขวามือ พบว่า ผู้ตอบคนที่ 2 มีระดับคะแนนการประเมินการรับรู้ผลกระทบจากการท่องเที่ยวในระดับคะแนนที่สูงกว่าผู้ตอบคนที่ 1 ดังนั้นจะเห็นได้ว่า การปรับสเกลด้วยวินยเกตต์ช่วยให้สามารถเปรียบเทียบระดับค่าคะแนนการตอบระหว่างผู้ตอบที่แตกต่างกันได้



หมายเหตุ: v_n-v_c คือ วินยเกตต์ โดย $v_n > v_u > v_c$

ที่มา: ดัดแปลงมาจากรูปที่ 1 ของ King et al. (2004)

ภาพ 3 การปรับสเกลคำตอบที่ผู้ตอบเป็นผู้ประเมินระหว่างบุคคลด้วยวินยเกตต์

สำหรับการปรับค่าสเกลคำตอบด้วยการใช้วินเยตต์เป็นจุดยึดอาศัยแนวความคิดการปรับตำแหน่ง (Recode) ของคะแนนลำดับขั้นที่ผู้ตอบเป็นผู้ประเมินเทียบกับ (Relative) ชุดของวินเยตต์ภายใต้ข้อตกลงที่ว่าผู้ตอบแต่ละคนมีการเรียงลำดับวินเยตต์เหมือนกัน (หรือไม่มีการละเมิดข้อตกลงความเท่าเทียมกันของวินเยตต์) โดยมีเงื่อนไขในการปรับตำแหน่งดังนี้ ถ้าคำตอบที่ผู้ตอบเป็นผู้ประเมินอยู่ต่ำกว่า V_d ให้ค่าสเกลในตัวแปรใหม่ (สมมติว่าเป็น C) มีค่าเป็น 1 ($Self < V_d; C = 1$) ถ้ามีค่าเท่ากับ V_d ให้มีค่าเป็น 2 ($Self = V_d; C = 2$) ถ้ามีค่าอยู่ระหว่าง V_d กับ V_n ให้มีค่าเป็น 3 ($V_d < Self < V_n; C = 3$) ถ้ามีค่าเท่ากับ V_n ให้มีค่าเป็น 4 ($Self = V_n; C = 4$) ถ้ามีค่าอยู่ระหว่าง V_n กับ V_n ให้มีค่าเป็น 5 ($V_n < Self < V_n; C = 5$) ถ้ามีค่าเท่ากับ V_n ให้มีค่าเป็น 6 ($Self = V_n; C = 6$) และถ้ามีค่ามากกว่า V_n ให้มีค่าเป็น 7 ($Self > V_n; C = 7$) (King et al., 2004) ภายใต้เงื่อนไขข้างต้น เมื่อปรับค่าสเกลของคำตอบที่ผู้ตอบแต่ละคนเป็นผู้ประเมิน จะเห็นได้ว่า ผู้ตอบคนที่ 1 มีระดับค่าสเกลใหม่เท่ากับ 3 ($C = 3$) ส่วนผู้ตอบคนที่ 2 มีระดับค่าสเกลใหม่เท่ากับ 5 ($C = 5$) แสดงว่า ผู้ตอบคนที่ 2 มีระดับคะแนนการประเมินการรับรู้ผลกระทบจากการท่องเที่ยวในระดับคะแนนสูงกว่าผู้ตอบคนที่ 1

จากแนวความคิดการปรับค่าสเกลคำตอบที่นำเสนอข้างต้น ขอยกยความเพิ่มเติมในประเด็นของการออกแบบวินเยตต์ ซึ่งถ้าหากนักวิจัยกำหนดสถานการณ์วินเยตต์แบบสุดโต่งที่ทำให้ผู้ตอบให้ค่าสเกลวินเยตต์สูงสุดหรือต่ำสุดเหมือนกันทั้งหมด ก็จะมีผลต่อช่วงของค่าสเกลใหม่ อย่างเช่น ในกรณีตัวอย่างของรูปที่ 3 หากสมมติว่า ค่าสเกลที่ใช้ในการประเมินมี 4 ระดับ (1-4) และผู้วิจัยสร้างสถานการณ์วินเยตต์ที่ผู้ตอบทั้งสองคนให้ค่าสเกลของวินเยตต์ที่อยู่ระดับสูงเท่ากับ 4 ($V_n = 4$) เหมือนกัน ค่าสเกลใหม่ (C) ที่เป็นไปได้จะไม่มีค่า 7 หรือไม่มีการที่ $Self > V_n$ ในทางตรงกันข้าม หากสถานการณ์วินเยตต์ที่สร้างขึ้นทำให้ผู้ตอบทั้งสองคนให้ค่าสเกลที่อยู่ในระดับต่ำเท่ากับ 1 ($V_d = 1$) เหมือนกัน ค่าสเกลใหม่ (C) ที่เป็นไปได้จะไม่มีค่า 1 หรือไม่มีการที่ $Self < V_d$ เป็นต้น

ตัวแปรที่มีการปรับค่าใหม่ (ตัวแปร C ในกรณีตัวอย่างในรูปที่ 3) เป็นตัวแปรที่มีมาตรวัดแบบเรียงลำดับ (Ordinal scale variables) และไม่มีอิทธิพลจากความแตกต่างของคุณลักษณะเฉพาะส่วนบุคคล (DIF-free) สามารถใช้อธิบายและวิเคราะห์เปรียบเทียบระหว่างบุคคลที่แตกต่างกันได้ เช่นเดียวกับตัวแปรที่มีมาตรวัดแบบเรียงลำดับอื่นๆ อย่างเช่น การเปรียบเทียบโดยฮิสโตแกรม (Histograms) การวิเคราะห์ด้วยตารางจร (Contingency table) หรือตารางไขว้ (Cross-tabulation) การวิเคราะห์ด้วยแบบจำลอง Ordered probit เป็นต้น

จากแนวคิดพื้นฐานข้างต้นที่ผ่านมาได้มีการเสนอวิธีปรับค่าสเกลคำตอบที่ผู้ตอบเป็นผู้ประเมินเทียบกับวินเยตต์ 2 วิธี คือ วิธีที่ไม่อิงพารามิเตอร์ (Non-parametric) และวิธีที่อิงพารามิเตอร์ (Parametric) (King, et al., 2004; King and Wand, 2007; Wand, King and Lau, 2011) ดังมีรายละเอียดในแต่ละวิธีพอสังเขปดังนี้

ก. วิธีที่ไม่อิงพารามิเตอร์ (Non-parametric)

การปรับค่าสเกลด้วยวิธีที่ไม่อิงพารามิเตอร์เป็นวิธีที่ง่ายในการนำมาใช้ ไม่อิงกับข้อตกลงเบื้องต้นหรือการอนุมาน ทางสถิติ ไม่มีการใช้แบบจำลองหรือวิธีทางสถิติปรับค่าสเกลคำตอบ อาศัยเพียงข้อตกลงเบื้องต้นที่สำคัญ 2 ข้อ ดังที่ได้กล่าวมาแล้ว คือ ความสอดคล้องในการตอบ และความเท่าเทียมกันของวินเยตต์ วิธีนี้มีจุดอ่อนที่สำคัญ 2 ข้อ (King et al., 2004) คือ 1) ผู้ตอบแต่ละคนต้องให้ค่าสเกลคำตอบ (หรือตอบคำถาม) ทั้งที่เป็นคำถามที่ผู้ตอบเป็นผู้ประเมินและวินเยตต์ ดังนั้นหากในการศึกษามีชุดตัวอย่างและสิ่งที่ต้องการวัดหลายมิติก็จะส่งผลให้มีต้นทุนในการจัดการสูงขึ้นตามไปด้วย และ 2) ไม่มีการอนุมานหรือการทดสอบในทางสถิติ รวมทั้งไม่มีความคลาดเคลื่อนจากการวัด

จากแนวคิดการปรับตำแหน่งของคะแนนลำดับขั้นที่ผู้ตอบเป็นผู้ประเมินเทียบกับชุดของวินเยตต์ที่นำเสนอข้างต้น เมื่อกำหนดให้ y_i เป็นค่าสเกลคำตอบที่ประเมินโดยผู้ตอบคนที่ i ($i = 1, \dots, n$) และ v_{ij} คือ ค่าสเกลคำตอบที่ผู้ตอบคนที่ i ตอบในวินเยตต์ที่ j ($j = 1, \dots, J$) ภายใต้ข้อตกลงที่ว่า ผู้ตอบแต่ละคนมีการเรียงลำดับวินเยตต์เหมือนกัน คือ $v_{i,j-1} < v_{ij}$ สำหรับทุกๆ i และ j

เมื่อกำหนดให้ C_i คือ ตัวแปรที่เป็นชุดของค่าสเกลคำตอบที่ผู้ตอบเป็นผู้ประเมินซึ่งได้จากการปรับด้วยการเทียบกับค่าสเกลคำตอบของชุดวินเยตต์ของผู้ตอบคนที่ i ดังนั้นสามารถแสดงสมการคณิตศาสตร์ของตัวแปร C_i ที่ขจัดปัญหา DIF (DIF-corrected) ของค่าสเกลที่ผู้ตอบเป็นผู้ประเมินเทียบกับค่าสเกลของชุดวินเยตต์ ได้ดังนี้

$$C_i = \begin{cases} 1 & \text{if } y_i < v_{i,1} \\ 2 & \text{if } y_i = v_{i,1} \\ 3 & \text{if } v_{i,1} < y_i < v_{i,2} \\ \vdots & \vdots \\ 2J+1 & \text{if } v_i > v_{i,J} \end{cases} \quad (1)$$

ทั้งนี้เพื่อให้เข้าใจการปรับสเกลตามแนวทางของวิธีนี้ จึงขอยกตัวอย่าง โดยกำหนดให้มีคำถามที่ให้ผู้ตอบเป็นผู้ประเมินเกี่ยวกับทัศนคติของผู้ตอบต่อผลกระทบของการท่องเที่ยว จำนวน 3 ข้อ ($y_1 - y_3$) และวินเยตต์ 3 ระดับ คือ ระดับต่ำ (v_1) ระดับปานกลาง (v_2) และระดับสูง (v_3) โดยคำถามที่ให้ผู้ตอบเป็นผู้ประเมินและวินเยตต์ใช้มาตรวัดประมาณค่า 4 ระดับเหมือนกัน คือ ไม่เห็นด้วยมาก (แทนค่าด้วย 1) ไม่เห็นด้วย (แทนค่าด้วย 2) เห็นด้วย (แทนค่าด้วย 3) และ

เห็นด้วยมาก (แทนค่าด้วย 4) จากแนวคิดที่เสนอในสมการที่ (1) สามารถกำหนดเป็นเกณฑ์การปรับสเกลการตอบของผู้ตอบแต่ละคนได้ 7 ระดับ ดังแสดงในตาราง 1

ตาราง 1 เกณฑ์การปรับสเกลคำตอบของผู้ตอบแต่ละคนด้วยการเปรียบเทียบกับวินเยตต์

สเกลคำตอบในแต่ละคำถาม ที่ให้ผู้ตอบเป็นผู้ประเมิน (y)	ไม่เห็นด้วยมาก	ไม่เห็นด้วย	เห็นด้วย	เห็นด้วยมาก			
	1	2	3	4			
สเกลคำตอบที่ปรับโดยเทียบกับ วินเยตต์ (v ₁ , v ₂ , v ₃)	y < v ₁ 1	y = v ₁ 2	v ₁ < y < v ₂ 3	y = v ₂ 4	v ₂ < y < v ₃ 5	y = v ₃ 6	v ₃ < y 7

ที่มา: ดัดแปลงมาจากตาราง 17.1 ของ OECD (2014)

จากเกณฑ์การปรับสเกลการตอบที่แสดงในตารางที่ 1 เพื่อให้เห็นภาพของการเปรียบเทียบที่ชัดเจน ในที่นี้จึงสมมติให้ผู้ตอบจำนวน 3 คน ตอบคำถาม y_1 – y_3 ในระดับสเกลคำตอบที่เหมือนกัน หรือผู้ตอบแต่ละคนมีระดับทัศนคติต่อผลกระทบของการท่องเที่ยวเหมือนกัน (ชุดที่ 1) แต่มีค่าสเกลการตอบวินเยตต์แต่ละข้อ (v_1 – v_3) แตกต่างกัน (ชุดที่ 2) ดังแสดงในตารางที่ 2 ส่วนชุดที่ 3 เป็นสเกลคำตอบที่ปรับโดยเทียบกับวินเยตต์ (c_1 – c_3) ซึ่งจะเห็นได้ว่า เมื่อปรับค่าสเกลคำตอบของผู้ตอบด้วย แองเคอร์รีંગ วินเยตต์ ที่เป็นวิธีที่ไม่อิงพารามิเตอร์ พบว่า ผู้ตอบแต่ละคนมีทัศนคติต่อผลกระทบของการท่องเที่ยวในระดับที่แตกต่างกัน (ดูรายละเอียดในตาราง 2)

ตาราง 2 ตัวอย่างการปรับสเกลคำตอบด้วย แองเคอร์รีંગ วินเยตต์ ที่เป็นวิธีที่ไม่อิงพารามิเตอร์

	ผู้ตอบคนที่ 1				ผู้ตอบคนที่ 2				ผู้ตอบคนที่ 3					
y_1	1	2	3	(4)	1	2	3	(4)	1	2	3	(4)		
y_2	1	(2)	3	4	1	(2)	3	4	1	(2)	3	4		
y_3	1	2	(3)	4	1	2	(3)	4	1	2	(3)	4		
v_1	(1)	2	3	4	(1)	2	3	4	1	2	(3)	4		
v_2	1	(2)	3	4	1	2	(3)	4	1	2	(3)	4		
v_3	1	2	(3)	4	1	2	3	(4)	1	2	3	(4)		
c_1	1	2	3	4	5	6	(7)	1	2	3	4	5	(6)	7
c_2	1	2	3	(4)	5	6	7	1	2	(3)	4	5	6	7
c_3	1	2	3	4	5	(6)	7	1	2	3	(4)	5	6	7

หมายเหตุ: y_1 – y_3 คือ คำตอบที่ผู้ตอบเป็นผู้ประเมิน, v_1 – v_3 คือ วินเยตต์ โดย $v_1 < v_2 < v_3$ และ c_1 – c_3 คือ สเกลคำตอบที่ปรับโดยเทียบกับวินเยตต์

ที่มา: ดัดแปลงมาจากรูปที่ 17.1 ของ OECD (2014)

ทั้งนี้เพื่อความเข้าใจถึงวิธีปรับค่าสเกลคำตอบด้วยวิธีที่ไม่อิงพารามิเตอร์ที่แสดงในตารางที่ 2 ในที่นี้จึงขอยกตัวอย่างในกรณีของการปรับค่าสเกลคำตอบใหม่ของผู้ตอบคนที่ 1 ซึ่งให้ค่าสเกลคำตอบวินเยตต์ระดับต่ำ (v_1) เท่ากับ 1 วินเยตต์ระดับปานกลาง (v_2) เท่ากับ 2 และวินเยตต์ระดับสูง (v_3) เท่ากับ 3 ตามลำดับ โดยผู้ตอบคนที่ 1 ให้ค่าสเกลคำตอบของคำถามในข้อที่ 1 เท่ากับ 4 ซึ่งมีค่าสเกลมากกว่าค่าสเกลของวินเยตต์ระดับสูง ($y_1 > v_3$) ดังนั้นจากเงื่อนไขในตารางที่ 1 ค่าสเกลใหม่ในข้อที่ 1 (c_1) จึงมีค่าเท่ากับ 7 ส่วนคำตอบที่ผู้ตอบประเมินในข้อที่ 2 มีค่าสเกลเท่ากับ 3 ซึ่งเท่ากับค่าสเกลของวินเยตต์ระดับสูง ($y_2 = v_3$) ดังนั้นค่าสเกลใหม่ในข้อที่ 2 (c_2) จึงมีค่าเท่ากับ 6 สำหรับคำตอบของคำถามในข้อที่ 3 มีค่าสเกลเท่ากับ 2 ซึ่งเท่ากับค่าสเกลของวินเยตต์ระดับปานกลาง ($y_3 = v_2$) ดังนั้นค่าสเกลใหม่ในข้อที่ 3 (c_3) จึงมีค่าเท่ากับ 4

อย่างไรก็ตาม มีโอกาสที่ค่าสเกลคำตอบที่ปรับโดยการเทียบกับวินเยตต์ (ตัวแปร C ซึ่งก็คือ $\{c_1, c_2, c_3\}$ ในกรณีตัวอย่างข้างต้น) อาจมีค่าสเกลเป็นช่วง (Interval values) ถ้าผู้ตอบให้ค่าสเกลวินเยตต์เท่ากัน (ดูตัวอย่างในตารางที่ 2 กรณีของผู้ตอบคนที่ 3 ซึ่งให้ค่า $v_1 = v_2$ ทำให้ค่าสเกลใหม่ของข้อที่ 3 (c_3) ตกอยู่ในช่วง 2-4 เนื่องจากค่าสเกลคำตอบที่ผู้ตอบเป็นผู้ประเมินของข้อที่ 3 ซึ่งมีค่าเท่ากับ 3 มีค่าเท่ากับ v_1 และ v_2 ($y_3 = v_1 = v_2$) ดังนั้นจากเงื่อนไขในตารางที่ 1 ค่าสเกลใหม่ของข้อคำถามที่ 3 จึงมีค่าตกอยู่ในช่วง 2-4) หรือมีการให้ค่าสเกลที่ทำให้การเรียงลำดับวินเยตต์ผิดจากที่กำหนด (อาจไม่สอดคล้องกับผู้ตอบโดยส่วนใหญ่ หรือโดยเฉลี่ย หรือที่กำหนดโดยนักวิจัยและ/ผู้เชี่ยวชาญ) อย่างเช่น หากผู้ตอบคนที่ 1 ให้ค่าสเกล v_3 (ระดับสูง) = 2 และให้ค่าสเกล v_2 (ระดับปานกลาง) = 4 หมายความว่า ผู้ตอบให้ค่าสเกลกับวินเยตต์ที่อยู่ในระดับสูง (v_3) ต่ำกว่าวินเยตต์ที่อยู่ในระดับปานกลาง (v_2) หรือ $v_3 < v_2$ เป็นต้น ทั้งสองกรณีส่งผลให้วิธีที่ไม่อิงพารามิเตอร์ไม่มีประสิทธิภาพในการปรับสเกลคำตอบ (King et al., 2004)

ทั้งนี้เพื่อให้เห็นภาพที่ชัดเจนมากขึ้นของทั้งสองกรณี ในที่นี้จึงขอยกตัวอย่างที่เสนอโดย King and Wand (2007) ที่แสดงให้เห็นถึงการประมาณค่าสเกลคำตอบที่ปรับโดยเทียบกับวินเยตต์ (แทนด้วย C) โดยใช้วิธีที่ไม่อิงพารามิเตอร์ปรับค่าสเกลคำตอบที่ผู้ตอบเป็นผู้ประเมิน (แทนด้วย y) ด้วยการเทียบกับวินเยตต์ 2 วินเยตต์ (แทนด้วย v_1 และ v_2) และกำหนดให้ $v_1 < v_2$ ดังแสดงในตาราง 3

ตาราง 3 กรณีตัวอย่างที่เป็นไปได้ของการปรับค่าสเกลคำตอบด้วยวินเยตต์ 2 วินเยตต์

ตัวอย่าง	คำตอบที่ได้ จากการสำรวจ	1	2	3	4	5	สเกลที่ปรับค่า (C)
		$y < v_1$	$y = v_1$	$v_1 < y < v_2$	$y = v_2$	$y > v_2$	
1	$y < v_1 < v_2$	1	0	0	0	0	1
2	$y = v_1 < v_2$	0	1	0	0	0	2
3	$v_1 < y < v_2$	0	0	1	0	0	3
4	$v_1 < y = v_2$	0	0	0	1	0	4
5	$v_1 < v_2 < y$	0	0	0	0	1	5
6	$y < v_1 = v_2$	1	0	0	0	0	1
7	$y = v_1 = v_2$	0	1	0	1	0	2,3,4
8	$v_1 = v_2 < y$	0	0	0	0	1	5
9	$y < v_2 < v_1$	1	0	0	0	0	1
10	$y = v_2 < v_1$	1	0	0	1	0	1,2,3,4
11	$v_2 < y < v_1$	1	0	0	0	1	1,2,3,4,5
12	$v_2 < y = v_1$	0	1	0	0	1	2,3,4,5
13	$v_2 < v_1 < y$	0	0	0	0	1	5

หมายเหตุ: y คือ คำถามที่ผู้ตอบเป็นผู้ประเมิน, v_1 และ v_2 คือ วินเยตต์ โดย $v_1 < v_2$ และ C คือ สเกลคำตอบที่ปรับโดยเทียบกับวินเยตต์

ที่มา: อ้างอิงมาจากตาราง 1 ของ King and Wand (2007)

จากตัวอย่างในตาราง 3 จะเห็นได้ว่า ตัวอย่างที่ 1-5 เป็นกรณีที่ผู้ตอบให้ลำดับของการเรียงวินเยตต์ถูกต้องตามที่กำหนดและไม่มีการให้ค่าสเกลของวินเยตต์เท่ากัน ($v_1 < v_2$) ผลลัพธ์ในคอลัมน์สเกลที่ปรับค่า (C) จึงเป็นค่าเดียว หรือเรียกว่า สเกลาร์ (Scalar) ตามเงื่อนไขที่กำหนดไว้ทั้งหมด ส่วนตัวอย่างที่ 6-8 เป็นกรณีที่ผู้ตอบให้ค่าสเกลของวินเยตต์เท่ากัน ($v_1 = v_2$) หากผู้วิจัยตรวจพบในขั้นตอนของการทดสอบวินเยตต์ว่า ผู้ตอบโดยเฉลี่ย (หรือส่วนใหญ่) ให้ค่าสเกลวินเยตต์ในลักษณะนี้ ผู้วิจัยควรปรับปรุงภาษาหรือสถานการณ์ในวินเยตต์เพื่อให้ผู้ตอบเข้าใจและเห็นถึงความแตกต่างระหว่างวินเยตต์ทั้งสอง ภายใต้กรณีนี้ค่า C ที่ประมาณได้ของตัวอย่างที่ 6 และ 8 ยังคงเป็นค่าเดียวตามเงื่อนไขที่กำหนด ขณะที่ค่า C ที่ประมาณได้ในกรณีตัวอย่างที่ 7 เป็นค่าช่วง หรือเรียกว่า Censored value ที่มีค่าในช่วง 2-4 สำหรับตัวอย่างที่ 9-13 เป็นกรณีที่ผู้ตอบให้ค่าสเกลของวินเยตต์ที่ทำให้ลำดับการเรียงวินเยตต์ไม่ถูกต้อง ($v_2 < v_1$) หากตรวจพบในขั้นตอนของการทดสอบวินเยตต์ว่า ผู้ตอบโดยเฉลี่ย (หรือส่วนใหญ่) ให้ค่าสเกลวินเยตต์ในลักษณะนี้ ผู้วิจัยอาจปรับปรุงลำดับของวินเยตต์ใหม่ ค่า C ที่ประมาณได้ภายใต้กรณีตัวอย่างที่ 9 และ 13 ยังคงเป็นค่าเดียว ส่วนตัวอย่างที่ 10-12 เป็นค่าช่วงที่มีค่าในช่วงที่แตกต่างกันตามตำแหน่งของ y

จากกรณีตัวอย่างจะเห็นได้ว่า การให้ค่าสเกลที่เท่ากันของวินเยตต์ และการให้ค่าสเกลที่ทำให้การเรียงลำดับวินเยตต์ผิดจากที่กำหนด ไม่ได้ทำให้เกิด Censoring effect ต่อการประมาณ

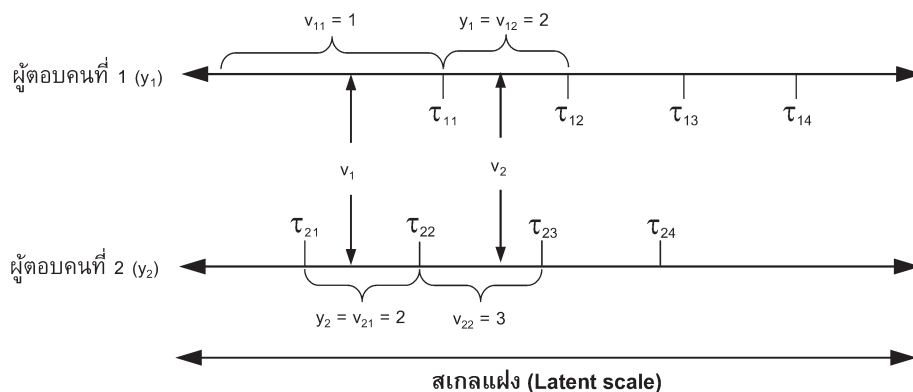
ค่า C ในบางกรณี อย่างเช่น กรณีตัวอย่างที่ 6 และ 8 (สำหรับกรณีที่มีการให้ค่าสเกลที่เท่ากันของ วินยเขตต์) และตัวอย่างที่ 9 และ 13 (สำหรับกรณีการให้ค่าสเกลที่ทำให้การเรียงลำดับวินยเขตต์ผิดจากที่กำหนด) ซึ่งในความเป็นจริงนั้น บางครั้งผู้ตอบอาจมีความยากลำบากที่จะจำแนกหรือเรียงลำดับ วินยเขตต์ที่มีความใกล้เคียงกัน อย่างเช่น หากสอบถามถึงความพึงพอใจต่อระยะเวลาที่ใช้ในการรอ การให้บริการระหว่าง 10 กับ 13 นาที ผู้ตอบอาจจะรู้สึกหรือให้ค่าคะแนนความพึงพอใจที่เท่ากัน เนื่องจากไม่สามารถจำแนกความแตกต่างของการรอการให้บริการระหว่าง 10 กับ 13 นาที ได้ เป็นต้น

ในงานศึกษาที่ผ่านมา มีการเสนอแนวทางการปรับแก้ค่า C ในกรณีทั้งสอง โดยพยายาม มุ่งแก้ไขเพื่อให้ค่า C ที่ประมาณได้มีค่าเป็นค่าเดียว (Uncensored) อย่างเช่น OECD (2014) ใช้แนวทางการปรับค่า C ในกรณีที่มีการให้ค่าสเกลที่เท่ากันของวินยเขตต์ด้วยการเทียบกับค่าสเกล ต่ำสุดของวินยเขตต์ที่เท่ากัน (Lower bound scores) อย่างเช่น หากผู้ตอบให้ค่าสเกล $v_1 = v_2$ ซึ่งไม่สอดคล้องกับที่กำหนดไว้ว่า $v_1 < v_2$ ก็จะประมาณค่า C โดยการเทียบคำตอบที่ผู้ตอบประเมิน กับค่า v_1 แต่ถ้าผู้ตอบให้สเกล $v_2 < v_1$ ก็จะ Re-classified ให้วินยเขตต์ทั้งสองเท่ากัน ($v_2 = v_1$) และประมาณค่า C โดยการเทียบคำตอบที่ผู้ตอบประเมินกับค่า v_2 เป็นต้น หรืออย่างเช่นในงานศึกษา ของ Crane et al. (2016) ได้ตัดชุดตัวอย่างที่มีการให้ค่าสเกลที่ทำให้การเรียงลำดับวินยเขตต์ผิดจากที่กำหนด หรือมีการละเมิดข้อตกลงความเท่าเทียมกันของวินยเขตต์ (Vignette equivalence) เป็นต้น อย่างไรก็ตาม การตัดชุดตัวอย่างออกจากการวิเคราะห์อาจเผชิญกับปัญหาอคติจากการเลือก (Selection bias) และทำให้สูญเสียข้อมูลที่สำคัญได้ (King and Wand, 2007)

Wand, King and Lau (2011) ได้สรุปวิธีการปรับแก้ค่า C ที่เป็นค่าช่วง (Interval value หรือ Censored) ให้เป็นค่าเดียว (Scalar) ภายใต้ข้อตกลงเบื้องต้นที่แตกต่างกัน 4 วิธี ได้แก่ 1) Omitting interval values: เป็นวิธีที่ง่ายที่สุดโดยการละทิ้งชุดตัวอย่างที่ค่า C มีค่าเป็นช่วง (หรือ Vector-valued) อย่างไรก็ตาม วิธีนี้ทำให้สูญเสียข้อมูลที่สำคัญในการวิเคราะห์ (King and Wand, 2007) และเกิดอคติจากการเลือกตัวอย่าง 2) Uniform allocation within intervals: เป็นวิธีที่เสนอโดย King et al. (2004) โดยการแทนค่า C ที่เป็นค่าช่วงด้วยสัดส่วนที่เท่าๆ กัน ($p_{j, m:n} = 1/(n-m+1)$ สำหรับ $j \in \{m, \dots, n\}$) โดยที่ C มีค่าอยู่ในช่วง $m-n$ อย่างเช่น ในตาราง ที่ 2 ตัวอย่างที่ 7 มีค่า C อยู่ในช่วง 2-4 หรือ $\{2, 3, 4\}$ ดังนั้น $p_{2, 2:4} = 0.33$, $p_{3, 2:4} = 0.33$ และ $p_{4, 2:4} = 0.33$ และค่า C ที่ปรับค่าใหม่มีค่าเท่ากับ 3 เป็นต้น การปรับค่าแบบนี้อาจทำให้ผลลัพธ์ ที่ได้เกิดความอคติจากความหนาแน่นของค่าสเกลที่ปรับใหม่ ทำให้การกระจายของค่า C ไม่สอดคล้อง กับความเป็นจริง ส่วนวิธีที่ 3) Censored ordered probit ซึ่งพัฒนาเพิ่มเติมจากแบบจำลอง Ordered probit ให้สามารถใช้ในกรณีที่ C มีค่าเป็นช่วงได้ และ 4) Minimum entropy ที่ประยุกต์ใช้ Entropy ประมาณค่า $\hat{C}$ จากค่า C ที่เป็นค่าช่วง วิธีที่ 3 และ 4 เป็นวิธีที่เสนอโดย King and

Wand (2007) อย่างไรก็ตาม ในการวิเคราะห์ด้วย Package “anchors” ในโปรแกรม R จะได้ผลลัพธ์ที่แสดงถึงการปรับค่า C ที่เป็นค่าช่วงของทั้งส่ววิธี (Wand, King and Lau, 2011)

วิธีปรับค่าสเกลคำตอบที่เสนอโดย King et al. (2004) และ King and Wand (2007) เรียกโดยทั่วไปว่า C -estimator (หรือ C -scale) ต่อมา Wand (2013) ได้เสนอทางเลือกในการปรับค่าสเกลคำตอบที่ผู้ตอบเป็นผู้ประเมินเทียบกับวินเยตต์ใหม่ที่เรียกว่า B -scale ซึ่งไม่จำเป็นต้องมีข้อตกลงเบื้องต้นเกี่ยวกับความเท่าเทียมกันของวินเยตต์ (Vignette equivalence) และ B -scale ยังมีความน่าเชื่อถือถือว่า C -scale ซึ่งจากภาพ 4 Wand (2013) ตั้งข้อสังเกตถึงความน่าเชื่อถือของ C -scale ที่ว่า หากผู้ตอบคนที่ 1 ให้คะแนนคำตอบเท่ากับ 2 และค่า $v_1 = 1$, $v_2 = 2$ ดังนั้นค่า C -scale ของผู้ตอบคนที่ 1 มีค่าเท่ากับ 4 ($C = 4$ เมื่อ $y = v_2$) ขณะที่ผู้ตอบคนที่ 2 ให้คะแนนคำตอบเท่ากับ 2 และค่า $v_2 = 2$, $v_2 = 3$ ดังนั้นค่า C -scale ของผู้ตอบคนที่ 2 มีค่าเท่ากับ 2 ($C = 2$ เมื่อ $y = v_1$) จากกรณีนี้ข้อตกลงเบื้องต้นทั้งสองข้อมิได้ถูกละเมิด ดังนั้นค่า C -scale ที่ได้หลังจากการปรับค่า พบว่า $C_1 > C_2$ ซึ่งขัดแย้งกับข้อมูลจริงที่ว่า $y_1 < y_2$ ดังนั้น Wand (2013) จึงเสนอค่าสเกลใหม่ที่เรียกว่า B -scale ดังแสดงในตาราง 4 ภายใต้การกำหนดให้ผู้ตอบมีจุดตัด/ขีดแบ่งร่วมกันของผู้ตอบ



หมายเหตุ: y คือ คำถามที่ผู้ตอบเป็นผู้ประเมิน, v_1 และ v_2 คือ วินเยตต์ โดย $v_1 < v_2$

ที่มา: ดัดแปลงมาจากรูปที่ 4 ของ Wand (2013)

ภาพ 4 ข้อสังเกตถึงความน่าเชื่อถือของ C -scale

ตาราง 4 เกณฑ์การปรับสเกลคำตอบด้วยวิธีไม่อิงพารามิเตอร์ที่เป็น C-scale และ B-scale

ลำดับของค่าสเกล				สัญลักษณ์	C-scale	B-scale
y_i	<	v_1	$\leq$	v_2	$y_i < v_{i1} \leq v_{i2}$	1
y_i	=	v_1	<	v_2	$y_i = v_{i1} \leq v_{i2}$	{1, 2}
v_1	<	y_i	<	v_2	$v_{i1} < y_i \leq v_{i2}$	2
v_1	<	y_i	=	v_2	$v_{i1} < y_i = v_{i2}$	{2, 3}
v_1	$\leq$	v_2	<	y_i	$v_{i1} \leq y_i = v_{i2}$	3

หมายเหตุ: y_i คือ คำถามที่ผู้ตอบเป็นผู้ประเมิน, v_1 และ v_2 คือ วินเยตต์ โดย $v_1 < v_2$

ที่มา: ดัดแปลงมาจากตาราง 1 ของ Wand (2013)

ข. วิธีที่อิงพารามิเตอร์ (Parametric)

วิธีที่อิงพารามิเตอร์ที่นิยมใช้ในการศึกษาที่ผ่านเป็นวิธีที่เสนอโดย King et al. (2004) ซึ่งแก้ไขจุดอ่อนที่สำคัญของวิธีที่ไม่อิงพารามิเตอร์ วิธีนี้ไม่จำเป็นต้องถามวินเยตต์กับชุดตัวอย่างทั้งหมดที่สำรวจ สามารถสอบถามวินเยตต์เฉพาะ Random sample (หรือ Subsample) จากประชากรกลุ่มเป้าหมายที่เป็นกลุ่มเดียวกัน ทำให้สามารถลดต้นทุนในการจัดการได้ และมีการอนุมานทางสถิติรวมทั้งการวัดค่าความคลาดเคลื่อนจากการวัดแบบสุ่ม (Random measurement error) ออกจากคำตอบที่ประเมินด้วยผู้ตอบและวินเยตต์ นอกจากนี้ยังแก้ไขปัญหา DIF ที่เกิดจากการที่ผู้ตอบแต่ละคนมีขีดแบ่งในการให้ค่าสเกลคำตอบที่แตกต่างกัน โดยกำหนดให้ขีดแบ่ง (Thresholds: τ) แตกต่างกันได้ตามลักษณะเฉพาะทางสังคมและเศรษฐกิจของบุคคล กล่าวคือ τ ถูกกำหนดโดยลักษณะเฉพาะทางสังคมและเศรษฐกิจของบุคคล อย่างเช่น เพศ อายุ รายได้ ระดับการศึกษา เป็นต้น [$\tau = f(\text{individual socioeconomic characteristics})$]

แบบจำลองที่เสนอโดย King et al. (2004) รู้จักกันโดยทั่วไปในชื่อ CHOPIT (Compound Hierarchical Ordinal Probit) ในงานศึกษาบางชิ้นเรียกว่า HOPIT (Hierarchical Ordinal Probit) แบบจำลองดังกล่าวมีพื้นฐานมาจากการพัฒนาแบบจำลองร่วมกันของคำตอบที่ประเมินด้วยผู้ตอบและวินเยตต์ โดยประยุกต์ใช้ Hierarchical ordered probit model (HOPIT) ในการพัฒนาแบบจำลองดังกล่าว

แบบจำลอง HOPIT ทั่วไป ประกอบด้วยสมการที่สำคัญ 2 สมการ คือ 1) สมการการตอบสนอง (Responsiveness equation) ที่เป็นแบบจำลองคำตอบที่ผู้ตอบเป็นผู้ประเมินหรือให้ค่าสเกล ซึ่งเป็นฟังก์ชันของชุดตัวแปรอธิบายของผู้ตอบ อย่างเช่น ตัวแปรด้านสังคมและเศรษฐกิจ

ของผู้ตอบ เป็นต้น และ 2) สมการพฤติกรรมการรายงาน (Reporting behavior equation) ที่เป็นแบบจำลองของขีดแบ่ง (Thresholds) ซึ่งเป็นฟังก์ชันของคุณลักษณะเฉพาะของผู้ตอบ อย่างเช่น เพศ อายุ รายได้ เป็นต้น

ดังนั้นเมื่อกำหนดให้ y_{is} เป็นค่าสเกลคำตอบแบบมาตรวัดประมาณค่า (Rating scale) ที่ผู้ตอบคนที่ i ($i = 1, \dots, n$) ตอบคำถามที่ s ($s = 1, \dots, S$) ในแบบสอบถาม (หรือแบบสำรวจ) ที่เป็น Subjective respondent-assessment ดังนั้นจากสมการการตอบสนอง (หรือสมการพฤติกรรม) สามารถแสดงที่มาของ y_{is} ได้ดังนี้

$$y_{is}^* = \beta X_i + \varepsilon_{is} \quad ; s = 1, \dots, S \quad (2)$$

โดยที่ y_{is}^* คือ คำตอบที่แท้จริงของคำถามที่ s ของผู้ตอบคนที่ i ซึ่งเป็นตัวแปรแฝง β คือ เวกเตอร์ของค่าพารามิเตอร์ (ไม่มีเทอมของค่าคงที่)

X_i คือ เวกเตอร์ของตัวแปรที่สามารถอธิบายคำตอบที่แท้จริงของผู้ตอบคนที่ i

ε_{is} คือ ค่าคลาดเคลื่อนที่มีการแจกแจงเหมือนกันและเป็นอิสระต่อกัน $\varepsilon_{is} \sim N(0, \sigma_{\varepsilon}^2)$ โดย และเป็นอิสระจาก X_i

จากสมการที่ (2) คำตอบที่ผู้ตอบคนที่ i ตอบคำถามที่ s ในแบบสอบถาม (y_{is}) ได้จากการเปรียบเทียบ y_{is}^* กับขีดแบ่งนามธรรม (Subjective thresholds: τ_{is}^j) ของผู้ตอบคนที่ i ดังนี้

$$y_{is} = j \quad \text{if} \quad \tau_{is}^{j-1} < y_{is}^* < \tau_{is}^j \quad ; j = 1, \dots, J \quad (3)$$

โดยที่ $\tau_{is}^0 = -\infty$, $\tau_{is}^J = \infty$ และ $\tau_{is}^{j-1} < \tau_{is}^j$ ทั้งนี้ขีดแบ่ง (τ) อาจแตกต่างกันตามลักษณะเฉพาะของผู้ตอบที่สามารถสังเกตและวัดค่าได้ (Z_i) (King et al., 2004) และที่ไม่สามารถสังเกตหรือวัดได้ (μ_i) ซึ่งเสนอเพิ่มเติมโดย Kapteyn, Smith and van Soest (2007) ดังนั้นสมการพฤติกรรมการรายงานที่เป็นแบบจำลองขีดแบ่งของผู้ตอบคนที่ i สำหรับคำถาม s สามารถแสดงได้ดังนี้

$$\begin{aligned} \tau_{is}^1 &= \gamma^1 Z_i + \mu_i \\ \tau_{is}^j &= \tau_{is}^{j-1} + e^{\gamma^j Z_i} \quad ; j = 2, \dots, J-1 \end{aligned} \quad (4)$$

โดยที่ μ_i เป็นค่าคลาดเคลื่อน ที่ $\mu_i \sim N(0, \sigma_{\mu}^2)$ และเป็นอิสระจาก X_i , Z_i และ ε_{is} โดย Kapteyn et al. (2007) เสนอเพิ่มเติมต่อจากแบบจำลองของ King et al. (2004) เทอม μ_i เป็น Random individual effect ที่แสดงถึง อิทธิพลอื่นๆ ของผู้ตอบที่มีต่อค่าสเกลคำตอบ ไม่สามารถสังเกตหรือวัดได้ และนอกเหนือจากลักษณะเฉพาะของผู้ตอบที่สามารถวัดได้โดยตรง เช่น เพศ อายุ รายได้ เป็นต้น (Kapteyn et al., 2007) โดย Package “anchors” ในโปรแกรม R จะไม่มีเทอมของค่าคลาดเคลื่อนนี้ (Wand, King and Lau, 2011)

ภายใต้แนวทางเดียวกัน สำหรับค่าสเกลคำตอบของวินยเกตต์ (Y_{iv}) ที่เป็นมาตรวัดประมาณค่า ที่ผู้ตอบคนที่ i ($i = 1, \dots, n$) ตอบวินยเกตต์ที่ v ($v = 1, \dots, V$) สามารถแสดงได้ดังนี้

$$y_{iv}^* = \alpha_v + \varepsilon_{iv} \quad ; v = 1, \dots, V \quad (5)$$

โดยที่ y_{iv}^* คือ คำตอบที่แท้จริงของคำถามวินยเกตต์ที่ v ของผู้ตอบคนที่ i ซึ่งเป็นตัวแปรแฝง

α_v คือ ค่าพารามิเตอร์วินยเกตต์ที่กำหนดเป็นค่าคงที่ (Fixed vignette parameter)

ε_{iv} คือ ค่าคลาดเคลื่อน โดย $\varepsilon_{iv} \sim N(0, \sigma_v^2)$

เช่นเดียวกันกับคำตอบที่ผู้ตอบเป็นผู้ประเมิน ค่าสเกลคำตอบของวินยเกตต์ (Y_{iv}) ที่ผู้ตอบคนที่ i ตอบวินยเกตต์ที่ v ได้จากการเปรียบเทียบ y_{iv}^* กับ Subjective thresholds (τ_{iv}^j) ของผู้ตอบคนที่ i ดังนี้

$$y_{iv} = j \quad \text{if} \quad \tau_{iv}^{j-1} < y_{iv}^* < \tau_{iv}^j \quad ; j = 1, \dots, J \quad (6)$$

โดยที่ $\tau_{iv}^0 = -\infty$, $\tau_{iv}^J = \infty$ และ $\tau_{iv}^{j-1} < \tau_{iv}^j$ ส่วน Thresholds 1 ถึง Thresholds $J-1$ ถูกกำหนดโดย γ_1 และตัวแปรลักษณะเฉพาะของผู้ตอบที่สามารถสังเกตและวัดค่าได้ (Z_i) เช่นเดียวกับสมการที่ (3)

$$\begin{aligned} \tau_{iv}^1 &= \gamma_1 Z_i + \mu_i \\ \tau_{iv}^j &= \tau_{iv}^{j-1} + e^{\gamma^j Z_i} \quad ; j = 2, \dots, J-1 \end{aligned} \quad (7)$$

จากข้อตกลงความสอดคล้องในการตอบ (Response consistency) ค่าสัมประสิทธิ์ γ_1 ของสมการที่ (4) และ (7) จึงเท่ากัน (หรือ เป็นตัวเดียวกัน) ขณะที่ค่า α_v ที่เป็นค่าพารามิเตอร์วินยเกตต์สมการที่ (5) ถูกกำหนดให้คงที่ภายใต้ข้อตกลงความเท่าเทียมกันของวินยเกตต์ (Vignettes equivalence) ที่ว่า ผู้ตอบทุกคนรับรู้วินยเกตต์ในทางเดียวกัน (Kapteyn et al., 2011; Paccagnell, 2011) นอกจากนี้จะเห็นได้ว่า แบบจำลองข้างต้นที่อยู่ภายใต้วิธี Anchoring vignettes สามารถจัดการกับปัญหา DIF ได้ โดยการประมาณค่าขีดแบ่งภายใต้ความแตกต่างของลักษณะเฉพาะของบุคคล

แบบจำลอง CHOPIT ที่เสนอข้างต้น ประมาณค่าสัมประสิทธิ์ที่ไม่ทราบค่าด้วย Maximum likelihood estimation (MLE) ซึ่งสามารถดูรายละเอียดของ Log-likelihood function ได้ในบทความของ King et al. (2004)

สรุป

บทความนี้มีวัตถุประสงค์เพื่อนำเสนอวิธีแอ่งเคอร์ริง วินเยตต์ (Anchoring vignette) สำหรับแก้ไขอคติของมาตรวัดการรับรู้ในการวัดตัวแปรนามธรรม วิธีดังกล่าวเสนอโดย King et al. (2004) โดยการเพิ่มชุดข้อมูลเพื่อใช้เป็นเกณฑ์มาตรฐานในการยึดและปรับค่าสเกลคำตอบของผู้ตอบ ภายได้ข้อตกลงเบื้องต้นที่สำคัญ 2 ข้อ ได้แก่ ความสอดคล้องในการตอบ และความเท่าเทียมกันของวินเยตต์ การปรับค่าสเกลของผู้ตอบสามารถทำได้ 2 วิธี คือ 1) วิธีที่ไม่อิงพารามิเตอร์ (Non-parametric) ซึ่งง่ายและสะดวกในการดำเนินการ แต่ไม่มีการอนุมานสถิติ และ 2) วิธีที่อิงพารามิเตอร์ (Parametric) ที่พัฒนาจากแบบจำลอง HOPIT (ในบทความของ King et al., 2004 เรียกว่า CHOPIT) วิธีนี้มีประสิทธิภาพมากกว่าวิธีที่ไม่อิงพารามิเตอร์ สามารถจัดการปัญหา DIF ที่เกิดจากการที่ผู้ตอบแต่ละคนมีขีดแบ่ง (Thresholds) ในการให้ค่าสเกลคำตอบที่แตกต่างกัน ทั้งนี้จากการทบทวนการศึกษาในอดีต สามารถกล่าวได้ว่า วิธีแอ่งเคอร์ริง วินเยตต์ เป็นวิธีที่มีประสิทธิภาพในการตรวจสอบความแตกต่างของค่าสเกลคำตอบที่ผู้ตอบเป็นผู้ประเมินภายได้ความแตกต่างของกลุ่มทางเศรษฐกิจและสังคม หรือวัฒนธรรม ซึ่งจะช่วยให้การวิเคราะห์คำตอบที่มีลักษณะแบบเรียงลำดับ (Ordinal) ที่ได้จากการประเมินของผู้ตอบมีการคำนึงถึงความแตกต่างระหว่างบุคคลที่เป็นผู้ประเมิน และสามารถนำสเกลที่ปรับค่าแล้วมาเปรียบเทียบกันได้อย่างถูกต้องและน่าเชื่อถือมากขึ้น ซึ่งควรมีการนำมาประยุกต์ใช้มากขึ้นในประเทศไทย โดยเฉพาะในกลุ่มงานวิจัยทางด้านสังคมที่ใช้แบบสอบถามในการวัดตัวแปรนามธรรม โดยให้ผู้ตอบเป็นผู้ประเมิน อย่างเช่น การวัดความพึงพอใจ การวัดความสุข การวัดทัศนคติ การวัดความคิดเห็น ฯลฯ รวมทั้งการประเมินทางด้านสุขภาพที่ให้ผู้ตอบเป็นผู้ประเมินด้วยตนเอง เป็นต้น

กิตติกรรมประกาศ

บทความนี้เป็นส่วนหนึ่งของโครงการ “ผลกระทบของการท่องเที่ยวต่อประชาชนในท้องถิ่นและศักยภาพการเก็บค่าธรรมเนียม” ภายได้ชุดโครงการเรื่อง ประสิทธิภาพและธรรมาภิบาลของนโยบายสาธารณะไทย การวิจัยเพื่อเสริมสร้างนักนโยบายสาธารณะที่ดี ได้รับการสนับสนุนจากหน่วยวิจัยมนุษยศาสตร์ สำนักงานกองทุนสนับสนุนการวิจัย (สกว.) ผู้เขียนขอขอบคุณ ศ.ดร.มิ่งสรรพ ขาวสอาด หัวหน้าโครงการ ผู้มองเห็นความสำคัญของเครื่องมือนี้ในการวิจัยเศรษฐศาสตร์เชิงสังคม และผลักดันให้มีการเรียนรู้และใช้โดยนักวิจัยในโครงการ

รายการอ้างอิง

- มณิการ์ ชูทอง (2557). การพัฒนาเครื่องมือประเมินแรงจูงใจใฝ่สัมฤทธิ์ของนักเรียนโดยใช้แองเคอร์ริง วินเยตต์. วิทยานิพนธ์ (สาขาวิชาวิธีวิทยาการวิจัยการศึกษา). กรุงเทพฯ, จุฬาลงกรณ์มหาวิทยาลัย, 170.
- มณิการ์ ชูทอง และชยุดม ภิรมย์สมบัติ (2558). การพัฒนาเครื่องมือประเมินแรงจูงใจใฝ่สัมฤทธิ์ของนักเรียนโดยใช้แองเคอร์ริง วินเยตต์. *วารสารอิเล็กทรอนิกส์ทางการศึกษา*, 10(4), 309–323.
- Angelini, V., Cavapozzi, D., Corazzine, L. and Paccagnella, O. (2014). Do Danes and Italians rate life satisfaction in the same way? using vignettes to correct for individual-specific scale biases. *Oxford Bulletin of Economics and Statistics*, 76(5), 643–666.
- Araña, J.E., Leon, C.J. (2013). Correcting for scale perception bias in tourist satisfaction surveys. *Journal of Travel Research*, 52(6), 772–788.
- Bago d’Uva, T., Lindeboom, M., O’Donnell, O. and van Doorslaer, E. (2011). Slipping anchor? testing the vignettes approach to identification and correction of reporting heterogeneity. *Journal of Human Resources*, 46(4), 875–906.
- Bago d’Uva, T., van Doorslaer, E., Lindeboom, M., and O’Donnell, O. (2008). Does reporting heterogeneity bias the measurement of health disparities? *Health Economics*, 17(3), 351–375.
- Chevalier, A. and Fielding, A. (2011). An introduction to anchoring vignettes. *Journal of the Royal Statistical Society. Series A (General)*, 17(3), 569–574.
- Crane, M., Rissel, C., Greaves, S. and Gebel, K. (2016). Correcting bias in self-rated quality of life: an application of anchoring vignettes and ordinal regression models to better understand QoL differences across commuting modes. *Quality of Life Research*, 25, 257–266.
- Dolnicar, S. and B. Grun. (2009). Response style contamination of student evaluation data. *Journal of Marketing Education*, 31(2), 160–172.
- Grol-Prokopczyk, H. (2014). Age and sex effects in anchoring vignette studies: methodological and empirical contributions. *Survey Research Methods*, 8(1), 1–17.
- Grol-Prokopczyk, H., Freese, J. and Hauser, R.M. (2011). Using anchoring vignettes to assess group differences in general self-rated health. *Journal of Health and Social Behavior*, 52(2), 246–261.

- Holland, P.W. and Wainer, H. (1993). *Differential Item Functioning*. Hillsdale, New Jersey: Lawrence Erlbaum.
- Kapteyn, A., Smith, J.P. and van Soest, A. (2007). Vignettes and self-reports of work disability in the United States and the Netherlands. *American Economic Review*, 97(1), 461–473.
- Kapteyn, A., Smith, J.P., van Soest, A. and Yonkova, H. (2011). *Anchoring vignettes and response consistency*. RAND Population Research Center, 33.
- King, G. and Wand, J. (2007). Comparing incomparable survey responses: evaluating and selecting anchoring vignettes. *Political Analysis*, 15(1), 46–66.
- King, G., Murray, C.J.L., Salomon, J.A. and Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, 98(1), 191–207.
- Kristensen, N. and Johansson, E. (2008). New evidence on cross-country differences in job satisfaction using anchoring vignettes. *Labour Economics*, 15(1), 96–117.
- Lindeboom M. and van Doorslaer E. (2004). Cut-point shift and index shift in self-reported health. *Journal of Health Economics*, 23(6), 1083–1099.
- OECD (2014). Questionnaire development issues. In *PISA 2012 Technical Report*. 361–388.
- Paccagnella, O. (2011). Anchoring vignettes with sample selection due to non-response. *Journal of the Royal Statistical Society. Series A (General)*, 174(3), 665–687.
- Paccagnella, O. (2016). *Anchoring Vignettes with Sample Selection*. Department of Statistical Sciences, University of Padua: 41.
- Paulhus, D. (1991). Measurement and control of response bias. In Robinson, J., Shaver, P. and Wrightsman, L. (eds), *Measures of Personality and Social Psychological Attitudes*, 17–59. San Diego, CA: Academic Press, Inc.
- Rabe-Hesketh, S. and Skrondal, A. (2002). *Estimating Chopit models in gllam Political Efficacy Example from King et al. (2002)*. 8. Retrieved from <http://www.gllamm.org/chopit.pdf> (24 June 2016).
- van Soest, A., Delaney, L., Harmon, C., Kapteyn, A. and Smith, J.P. (2011). Validating the use of anchoring vignettes for the correction of response scale differences in subjective questions. *Journal of the Royal Statistical Society: Series A (General)*, 174(3), 575–595.

- Wand, J. (2013). Credible comparisons using interpersonally incomparable data: nonparametric scales with anchoring vignettes. *American Journal of Political Science*, 57(1), 249–262.
- Wand, J., King, G. and Lau, O. (2011). anchors: software for anchoring vignette data. *Journal of Statistical Software*, 42(3), 1–25.
- Wilgenburg, K.V. (2010). *The Validity of anchoring vignettes: Testing response consistency with an experiment*. Master Health Economics, Policy and Law. Instiute of Healthe Policy&Management. Rotterdam, Erasmus University, 30.

Translated Thai References

- Choothong, M. (2014). *A development of students' achievement motivation assessment tool using anchoring vignette* (master degree thesis). Chulalongkorn University, Bangkok.
- Choothong, M., & Piromsombat, C. (2015). A development of students' achievement motivation assessment tool using anchoring vignette. *An Online Journal of Education*, 10, 309–323.