

การศึกษาเปรียบเทียบการตรวจสอบความเที่ยงตรงเชิงเนื้อหา
ของเครื่องมือวิจัยด้วยเทคนิค IOC, CVR และ CVI*

A Comparative Study of Verifying the Content Validity of
a Research Instrument with the IOC, CVR and CVI

วีระยุทธ พรพจน์ธนาต (Wirayut Phonphotthanamat)**

*บทความวิชาการ คณะรัฐศาสตร์ มหาวิทยาลัยรามคำแหง

**ผู้ช่วยศาสตราจารย์, คณะรัฐศาสตร์ มหาวิทยาลัยรามคำแหง, E-mail: p.wirayut@gmail.com

ได้รับบทความ: 2 ก.พ. 65 / แก้ไขปรับปรุง: 26 เม.ย. 65 / อนุมัติให้ตีพิมพ์: 3 พ.ค. 65 / เผยแพร่ออนไลน์: 20 มิ.ย. 65

DOI:

บทคัดย่อ

บทความนี้มีวัตถุประสงค์เพื่ออธิบายและเปรียบเทียบการตรวจสอบความเที่ยงตรงเชิงเนื้อหาของเครื่องมือวิจัยด้วย IOC, CVR และ CVI โดยเปรียบเทียบใน 8 ประเด็น ได้แก่ 1) ผู้ริเริ่มและปีที่ริเริ่ม 2) จำนวนผู้เชี่ยวชาญที่ต้องใช้ในการพิจารณาเครื่องมือ 3) เกณฑ์การให้คะแนน 4) เกณฑ์การตัดสินใจใช้ได้ของข้อคำถาม 5) การใช้ในการตรวจสอบคุณภาพของข้อคำถามเป็นรายข้อ 6) การใช้ในการพิจารณาคุณภาพของเครื่องมือทั้งฉบับ 7) จุดเด่น และ 8) ข้อจำกัด ผลการเปรียบเทียบทั้ง 8 ประเด็น พบว่า ทั้ง IOC, CVR, และ CVI ต่างมีความแตกต่างกันในทุกๆ ประเด็น จึงนำไปสู่การสรุปความเห็นของ

ผู้เขียนว่าทั้ง IOC, CVR และ CVI ต่างมีหลักเกณฑ์ มีจุดเด่น และข้อจำกัดในตัวของตัวเอง มิได้มีวิธีการใดที่ดีหรือเหมาะสมที่สุดในการนำไปใช้ตรวจสอบคุณภาพของเครื่องมือวิจัย หากนักวิจัยเลือกที่จะใช้วิธีการใดแล้วพึงต้องศึกษารายละเอียดของหลักเกณฑ์หรือหลักการของวิธีนั้นๆ เพื่อจะได้อ้างอิงหรือใช้ได้ อย่างถูกต้องตามหลักวิชาการ

คำสำคัญ: เครื่องมือวิจัย ความเที่ยงตรงเชิงเนื้อหา IOC CVR CVI

Abstract

This article aims to describe and compare the content validation of a research instrument with the IOC, CVR, and CVI methods in eight main issues: 1) the pioneers and the proposed years, 2) the number of content experts evaluating an instrument, 3) the values range, 4) the criteria for evaluating the validity of questions, 5) the verification of the validity of each question, 6) the verification of the validity of a questionnaire, 7) the strengths, and 8) the limitations. The results reveal that the three methods are different in every aspect. It can be concluded that the IOC, CVR and CVI have their criteria, strengths, and limitations. Therefore, there is no most suitable method for verifying the validity of research instruments. However, a researcher who chooses to use any technique should study the criteria or principles of that method in detail to be able to cite or use correctly according to academic principles.

Keywords: research instrument; content validity; IOC; CVR; CVI

บทนำ

ในบรรดาการแสวงหาความรู้ ความจริง ในโลกยุคปัจจุบันด้วยวิธีการต่างๆ นั้น การวิจัยถือได้ว่าเป็นวิธีการที่น่าเชื่อถือและเป็นที่ยอมรับมากที่สุด ทำให้ปัจจุบันการวิจัยเป็นกระบวนการหนึ่งของการแสวงหาความรู้ในศาสตร์ทุกสาขา เพราะเป็นกระบวนการที่ดำเนินไปอย่างเป็นระบบ มีระเบียบ กฎเกณฑ์ หรือเป็นวิทยาศาสตร์ หนึ่งในกระบวนการวิจัยนั้นจะมีเรื่องการสร้างเครื่องมือเพื่อใช้ในการเก็บรวบรวมข้อมูล อาทิ แบบสอบถาม แบบสัมภาษณ์ แบบสังเกต และแบบทดสอบ เป็นต้น เครื่องมือต่างๆ เหล่านี้ที่ผู้พัฒนาขึ้นย่อมต้องการตรวจสอบคุณภาพให้เกิดความมั่นใจได้ว่าจะสามารถนำไปใช้ในการเก็บรวบรวมข้อมูลที่นักวิจัยต้องการเพื่อนำมาสู่การตอบคำถามวิจัยได้อย่างถูกต้องตรงตามวัตถุประสงค์วิจัย การตรวจสอบคุณภาพของเครื่องมือวิจัยมีหลากหลายประเภท ได้แก่ การตรวจสอบความเที่ยงตรง การตรวจสอบความเชื่อมั่น การตรวจสอบอำนาจจำแนก และการตรวจสอบความยากง่าย เป็นต้น

บทความนี้มุ่งเน้นความสนใจไปที่การตรวจสอบความเที่ยงตรง (Validity) ซึ่งเป็นเรื่องของการวัดในสิ่งที่ต้องการจะวัดได้อย่างถูกต้องแม่นยำของเครื่องมือวัด (โชติกา ภาชีผล, 2554; สุจิตรา บุญยรัตพันธุ์, 2563; Babbie, 2010) โดยผู้เขียนมุ่งเน้นไปที่การตรวจสอบความเที่ยงตรงเชิงเนื้อหา (Content Validity) ที่มีวิธีการในการตรวจสอบได้หลากหลาย ได้แก่ 1) การหาค่าดัชนีความสอดคล้องของข้อคำถาม (Index of Item Objective Congruence) หรือนิยมเรียกกันสั้นๆ ว่าการหาค่า IOC 2) การหาอัตราส่วนความเที่ยงตรงเชิงเนื้อหาของข้อคำถาม (Content Validity Ratio) หรือนิยมเรียกสั้นๆ ว่า CVR และ 3) การหาค่าดัชนีความเที่ยงตรงเชิงเนื้อหาของเครื่องมือ (Content Validity Index) หรือนิยมเรียกสั้นๆ ว่า CVI โดยบทความนี้ต้องการอธิบายและเปรียบเทียบแต่ละวิธีว่ามีที่มา มีหลักในการคำนวณ และมีจุดเด่น จุดด้อยต่างกันอย่างไร เพื่อเป็นข้อมูลในการตัดสินใจต่อผู้ทำวิจัยโดยเฉพาะอย่างยิ่งการวิจัยเชิงปริมาณ ซึ่งมีความจำเป็นต้องใช้เครื่องมือวิจัยในการเก็บรวบรวมข้อมูล เพื่อให้มั่นใจได้ว่าเครื่องมือวิจัยนั้นจะมีคุณภาพเพียงพอ เพราะที่ผ่านมาจากผู้เขียนได้อ่านงานวิจัยในหลายๆ ชิ้น พบว่า นักวิจัยหน้าใหม่หลายท่าน มักมีความเข้าใจที่คลาดเคลื่อนในการเลือกใช้เทคนิค IOC, CVR และ CVI เพื่อตรวจสอบ

ความเที่ยงตรงเชิงเนื้อหาของเครื่องมือวิจัย ซึ่งอาจเกิดจากความสับสน ทำให้ใช้เทคนิควิธีการดังกล่าว ไม่ถูกต้อง เข้าทำนองการคัดลอกต่อๆ กันมาโดยขาดความเข้าใจที่ถูกต้อง ในส่วนต่อไปของเนื้อหา จะกล่าวถึง การตรวจสอบความเที่ยงตรงเชิงเนื้อหาด้วย IOC, CVR และ CVI จากนั้นจึงนำไปสู่บทสรุป เปรียบเทียบและอภิปรายการตรวจสอบความเที่ยงตรงเชิงเนื้อหาด้วย IOC, CVR และ CVI ในท้ายที่สุด ดังนี้

การตรวจสอบความเที่ยงตรงเชิงเนื้อหาด้วย IOC, CVR และ CVI

ความเที่ยงตรงเชิงเนื้อหา (Content Validity) หมายถึง การที่เครื่องมือวิจัยที่สร้างขึ้นมามีเนื้อหา ครอบคลุมและเป็นตัวแทนของเนื้อหาสาระของสิ่งที่ต้องการวัดอย่างเพียงพอ (พรเพ็ญ เพชรสุขศิริ, 2541; ศิริชัย กาญจนวาสี, 2552) โดยจะพิจารณาจาก 2 ประเด็น¹ คือ 1) ข้อความในเครื่องมือวัดนั้นๆ เป็นข้อความที่ตรงเป้าหรือตรงประเด็นกับสิ่งที่ต้องการวัด และ 2) ข้อความในเครื่องมือวัดนั้นๆ ครอบคลุม ปรากฏการณ์ที่ต้องการวัดอย่างกว้างขวางเพียงพอ (สุจิตรา บุญยรัตพันธุ์, 2563) กล่าวได้ว่าประเด็นแรก เป็นความเที่ยงตรงจากรูปลักษณ์ (Face Validity) หรือ ความเที่ยงตรงรายข้อ (Item Validity) ซึ่งหมายถึง การพิจารณาจากข้อคำถามแต่ละข้อที่ถูกสร้างขึ้นมามีความเกี่ยวข้องสัมพันธ์กับเนื้อหาที่ต้องการวัด มากน้อยเพียงใด (Gay, Mills, & Airasian, 2014) เช่น หากต้องการวัดระดับการพัฒนาทรัพยากรมนุษย์ ในองค์กรปกครองส่วนท้องถิ่นในด้านการฝึกอบรม² นักวิจัยจึงสร้างข้อคำถามมาประกอบด้วย ข้อที่ 1 “องค์กรของท่านมีการวิเคราะห์ความจำเป็นในการฝึกอบรม” ข้อที่ 2 “องค์กรของท่านมีการออกแบบและ

¹ หนังสือหรือตำราบางเล่มจะแยกระหว่าง Content Validity กับ Face Validity ออกจากกัน โดยที่ Content Validity จะมีความหมายถึงเฉพาะประเด็นเรื่องความครอบคลุมปรากฏการณ์ที่ต้องการวัดอย่างกว้างขวางเพียงพอ (ซึ่งคือ ประเด็นที่ผู้เขียนอธิบายเรื่อง Sampling Validity) สำหรับผู้เขียนขออธิบายรวมๆ เช่นเดียวกับ Gay, Mills, and Airasian (2014) ศุภชัย ยาวะประภาส (2544) สุจิตรา บุญยรัตพันธุ์ (2563) และ สีน พันธุ์พินิจ (2551)

² ข้อคำถามที่นำมายกตัวอย่าง นำมาจากส่วนหนึ่งในแบบสอบถามงานวิจัยของ วงพัทตร์ ภูพันธ์ศรี, และวีระยุทธ พรพจน์ธนาศ (2560)

พัฒนาหลักสูตรการฝึกอบรมที่ดี” ข้อที่ 3 “องค์การของท่านมีการบริหารหลักสูตรการฝึกอบรมที่ดี” ข้อที่ 4 “ผู้นำในองค์การของท่านเปิดกว้างทางความคิดสำหรับบุคลากรทุกคนในองค์การ” และ ข้อที่ 5 “ในองค์การของท่านส่งเสริมให้มีการพูดคุย ถกเถียง ชักถาม อย่างกว้างขวาง” เป็นต้น หากพิจารณาความเที่ยงตรงจากรูปลักษณะพบว่า ข้อที่ 1-3 น่าจะเป็นข้อความที่มีความเที่ยงตรงจากรูปลักษณะเพราะแต่ละข้อถามเกี่ยวกับเรื่องการฝึกอบรม ในขณะที่ข้อที่ 4-5 น่าจะมีปัญหาความเที่ยงตรงจากรูปลักษณะเพราะไม่ได้ถามเรื่องที่เกี่ยวข้องกับการฝึกอบรม

ส่วนประเด็นที่สอง คือเรื่องความเที่ยงตรงเชิงสุ่ม (Sampling Validity) นั่นคือ ข้อคำถามต่างๆ ที่ประกอบเป็นเครื่องมือวัดสามารถเป็นตัวแทนของประเด็นหรือมิติทั้งหมดของปรากฏการณ์ที่ต้องการจะวัดได้หรือไม่ (ศุภชัย ยาวะประภาษ, 2544) ซึ่งกระทำโดยการตรวจสอบเนื้อหาของมาตรวัด (Content Validity) เช่น หากต้องการวัดระดับการพัฒนาทรัพยากรมนุษย์ในองค์กรปกครองส่วนท้องถิ่นในด้านการฝึกอบรม ดังได้กล่าวไปแล้วว่า วัดจาก 5 ข้อ ก็ต้องพิจารณาว่าข้อคำถามทั้งหมดที่สร้างมา มีความครอบคลุมกับปรากฏการณ์ที่ต้องการวัดได้อย่างกว้างขวางเพียงพอแล้วหรือยัง ซึ่งจากตัวอย่างที่ยกมา กล่าวได้ว่ายังขาดความครอบคลุม เพราะขาดการวัดเรื่อง “การประเมินและติดตามผลการฝึกอบรม” “การที่บุคลากรในองค์การมีโอกาสดำเนินการฝึกอบรมอย่างเท่าเทียมกัน” และ “การมีงบประมาณในการจัดฝึกอบรมอย่างเพียงพอ” เป็นต้น

ความเที่ยงตรงเชิงเนื้อหาของนักวิจัยจะต้องดำเนินการ ดังนี้ 1) ต้องสามารถระบุเนื้อหาต่างๆ ที่เกี่ยวข้องกับปรากฏการณ์ที่จะวัดนั้นได้อย่างสมบูรณ์ 2) ต้องเลือกตัวแทนจากเนื้อหาต่างๆ เหล่านี้ แต่ละด้านออกมาได้อย่างดี และ 3) ต้องนำตัวแทนหรือข้อคำถามที่เป็นเนื้อหาเหล่านั้นมาจัดระเบียบให้อยู่ในรูปแบบที่สามารถวัดและทดสอบได้ (สุชาติ ประสิทธิ์รัฐสินธุ์, 2546) วิธีการตรวจสอบความเที่ยงตรงเชิงเนื้อหาโดยทั่วไปนิยมใช้ผู้เชี่ยวชาญ (หรือผู้ทรงคุณวุฒิ) เป็นผู้ตรวจสอบ (Gay, Mills, & Airasian, 2014) วิธีนี้จะให้คุณภาพความเที่ยงตรงได้อย่างดีก็ต่อเมื่อนักวิจัยเลือกผู้เชี่ยวชาญที่เป็นผู้เชี่ยวชาญในเรื่องที่ต้องการวัดจริงๆ (ชูศรี วงศ์รัตน์, 2549) โดยวิธีการในการหาความเที่ยงตรงเชิงเนื้อหาสามารถกระทำได้น้อย 3 วิธี

วิธีแรก การหาดัชนีความสอดคล้องของข้อคำถาม (Index of Item Objective Congruence: IOC) วิธีนี้ถูกพัฒนาโดย Rovinelli, & Hambleton (1977) เป็นวิธีที่เป็นที่นิยมในการวิจัยเชิงปริมาณทางด้านรัฐศาสตร์และรัฐประศาสนศาสตร์มากที่สุดเมื่อเทียบกับวิธีอื่นๆ โดยวิธีนี้นักวิจัยจะนำเครื่องมือวิจัย (แบบสอบถาม หรือแบบทดสอบ) ที่สร้างขึ้น พร้อมกับโครงร่างวิจัยฉบับย่อที่มีรายละเอียดของชื่อเรื่องวิจัย วัตถุประสงค์วิจัย กรอบแนวคิดในการวิจัย นิยามศัพท์ปฏิบัติการ เป็นต้น ไปให้ผู้เชี่ยวชาญจำนวน 3 ท่านขึ้นไปพิจารณาลงความเห็น ถึงความสอดคล้องระหว่างข้อคำถามกับจุดมุ่งหมายที่ต้องการวัด โดยให้ผู้เชี่ยวชาญลงความเห็นในแต่ละข้อคำถามตามความเห็น ใน 3 ระดับ คืออาจให้ +1 0 หรือ -1 โดยที่แต่ละระดับมีความหมาย คือ ให้คะแนน +1 หมายถึง ข้อคำถามวัดคุณลักษณะตรงตามนิยาม ให้คะแนน 0 หมายถึง ไม่แน่ใจว่าข้อคำถามนั้นวัดคุณลักษณะตรงตามนิยามหรือไม่ และให้คะแนน -1 หมายถึง ข้อคำถามวัดคุณลักษณะไม่ตรงตามนิยาม

ตารางที่ 1 ตัวอย่างแบบสอบถามการพัฒนาทรัพยากรมนุษย์เพื่อการหาค่าดัชนีความสอดคล้องของข้อคำถาม (IOC)

การพัฒนาทรัพยากรมนุษย์	ความคิดเห็น			ข้อเสนอแนะ
	เหมาะสม (+1)	ไม่แน่ใจ (0)	ไม่เหมาะสม (-1)	
1. ด้านการฝึกอบรม				
1.1 องค์กรของท่านมีการวิเคราะห์ความจำเป็นในการฝึกอบรม				
1.2 องค์กรของท่านมีการออกแบบและพัฒนาหลักสูตรการฝึกอบรมที่ดีที่สุด				
1.3 องค์กรของท่านมีการบริหารหลักสูตรการฝึกอบรมที่ดีที่สุด				
1.4 องค์กรของท่านมีการประเมินและติดตามผลการฝึกอบรมที่ดำเนินไปแล้ว				
1.5 บุคลากรทุกคนในองค์กรมีโอกาสได้รับการฝึกอบรมอย่างเท่าเทียมกัน				
1.6 องค์กรของท่านมีการจัดฝึกอบรมบุคลากรเป็นประจำ				
1.7 องค์กรของท่านมีงบประมาณในการจัดฝึกอบรมอย่างเพียงพอและเหมาะสม				
1.8 ผู้นำในองค์กรของท่านให้ความสำคัญกับการฝึกอบรม				
1.9 การฝึกอบรมมีส่วนช่วยในการพิจารณาเลื่อนตำแหน่ง				

ที่มา : ปรับจาก วงพักรัตน์ ภูพันธ์ศรี, และวีระยุทธ พรพจน์ธนาศ (2560)

การหาค่า IOC หาได้จากสูตร ดังนี้

$$IOC = \frac{\sum R}{N}$$

IOC แทน ดัชนีความสอดคล้องของข้อคำถาม

$\sum R$ แทน ผลรวมของความคิดเห็นของผู้เชี่ยวชาญทั้งหมด

N แทน จำนวนของผู้เชี่ยวชาญทั้งหมด

ข้อที่มีค่า IOC ผ่านเกณฑ์หรือใช้ได้โดยทั่วไปคือต้องมีค่า IOC ตั้งแต่ .50 ขึ้นไป อย่างไรก็ตาม นักวิจัยสามารถปรับเพิ่มค่า IOC ได้ (เพิ่มได้ลดไม่ได้) ขึ้นอยู่กับการวิจัยเรื่องนั้นๆ ว่าต้องการความเที่ยงตรงเชิงเนื้อหาอย่างน้อยเพียงใด (ประสพชัย พสุนนท์, 2558)

ตัวอย่างการคำนวณค่าดัชนีความสอดคล้องของข้อคำถามแต่ละข้อ สมมติว่านักวิจัยสร้างแบบสอบถามมาจำนวน 6 ข้อ และนำไปให้ผู้เชี่ยวชาญจำนวน 3 ท่านพิจารณา เมื่อผู้เชี่ยวชาญลงความเห็นในแต่ละข้อคำถามแล้ว นักวิจัยจะนำความเห็นจากผู้เชี่ยวชาญมาลงในตาราง แล้วคำนวณค่า IOC ดังตารางที่ 2

ตารางที่ 2 ตัวอย่างการคำนวณค่าดัชนีความสอดคล้องของข้อคำถามแต่ละข้อ

ข้อ	ผู้เชี่ยวชาญคนที่			ΣR	IOC = $\frac{\sum R}{N}$	หมายเหตุ
	1	2	3			
1	1	1	0	2	$\frac{2}{3} = .67$	ใช้ได้
2	1	0	0	1	$\frac{1}{3} = .33$	ใช้ไม่ได้
3	1	-1	0	0	$\frac{0}{3} = 0$	ใช้ไม่ได้
4	1	-1	-1	1	$\frac{1}{3} = .33$	ใช้ไม่ได้
5	0	-1	-1	-2	$\frac{-2}{3} = -.67$	ใช้ไม่ได้
6	1	1	1	3	$\frac{3}{3} = 1$	ใช้ได้

จากตารางที่ 2 จะเห็นได้ว่า ข้อที่มีค่า IOC ผ่านเกณฑ์ (มีค่า IOC ตั้งแต่ .50 ขึ้นไป) มีเพียง 2 ข้อ คือข้อ 1 และข้อ 6 ส่วนข้อคำถามที่เหลือมีค่า IOC ต่ำกว่า .50 ควรต้องปรับปรุง

วิธีที่สอง การหาอัตราส่วนความเที่ยงตรงเชิงเนื้อหาของข้อคำถาม (Content Validity Ratio : CVR) ซึ่ง Lawshe (1975 อ้างถึงในทิพย์สิริ กาญจนวาสี, และศิริชัย กาญจนวาสี, 2561; ศิริชัย กาญจนวาสี, 2552; Ayre, & Scally, 2014; Gilbert, & Prion, 2016) เป็นผู้ริเริ่มในการพัฒนาสูตรขึ้น โดยนำเสนอในรูปแบบบทความในปี 1975 ในบทความชื่อ “A Quantitative Approach to Content Validity” โดยนำเสนอการหาความเที่ยงตรงเชิงเนื้อหาของเครื่องมือวิจัย โดยการหา CVR จะนำเครื่องมือวิจัยไปให้ผู้เชี่ยวชาญพิจารณา โดยจะใช้ผู้เชี่ยวชาญจำนวนขั้นต่ำ 5 คน และสูงสุดไม่เกิน 40 คน หากผู้เชี่ยวชาญพิจารณาข้อคำถามแล้วเห็นว่า ข้อคำถามวัดคุณลักษณะตรงตามนิยามจะให้คะแนน + 1 แต่หากผู้เชี่ยวชาญเห็นว่าข้อคำถามวัดคุณลักษณะไม่ตรงตามนิยามจะให้คะแนน -1

ตารางที่ 3 ตัวอย่างแบบสอบถามการพัฒนาทรัพยากรมนุษย์เพื่อการหาค่าอัตราส่วนความเที่ยงตรงเชิงเนื้อหาของข้อคำถาม (CVR)

การพัฒนาทรัพยากรมนุษย์	ความคิดเห็น		ข้อเสนอแนะ
	เหมาะสม (+1)	ไม่เหมาะสม (-1)	
1. ด้านการฝึกอบรม			
1.1 องค์การของท่านมีการวิเคราะห์ความจำเป็นในการฝึกอบรม			
1.2 องค์การของท่านมีการออกแบบและพัฒนาหลักสูตรการฝึกอบรมที่ดี			
1.3 องค์การของท่านมีการบริหารหลักสูตรการฝึกอบรมที่ดี			
1.4 องค์การของท่านมีการประเมินและติดตามผลการฝึกอบรมที่ดำเนินไปแล้ว			
1.5 บุคลากรทุกคนในองค์กรมีโอกาสได้รับการฝึกอบรมอย่างเท่าเทียมกัน			
1.6 องค์การของท่านมีการจัดการการฝึกอบรมบุคลากรเป็นประจำ			
1.7 องค์การของท่านมีงบประมาณในการจัดฝึกอบรมอย่างเพียงพอและเหมาะสม			
1.8 ผู้นำในองค์กรของท่านให้ความสำคัญกับการฝึกอบรม			
1.9 การฝึกอบรมมีส่วนช่วยในการพิจารณาเลื่อนตำแหน่ง			

ค่า CVR ที่คำนวณได้จะอยู่ระหว่าง -1 ถึง +1 สำหรับเกณฑ์การพิจารณานั้น Lawshe (อ้างถึงในพิสนุ พงศ์ศรี, 2552; Taherdoost, 2014; Zerati, & Alavi, 2014) ได้กำหนดเกณฑ์ ซึ่งข้อคำถามที่มีค่า CVR ผ่านเกณฑ์จะผันแปรไปตามจำนวนผู้เชี่ยวชาญ ดังตารางที่ 4

ตารางที่ 4 ค่าอัตราส่วนความเที่ยงตรงเชิงเนื้อหาต่ำสุดที่ผ่านเกณฑ์ CVR

จำนวนผู้เชี่ยวชาญ	ค่า CVR ต่ำสุดที่มีนัยสำคัญ	จำนวนผู้เชี่ยวชาญ	ค่า CVR ต่ำสุดที่มีนัยสำคัญ
5	.99	13	.54
6	.99	14	.51
7	.99	15	.49
8	.78	20	.42
9	.78	25	.37
10	.62	30	.33
11	.59	35	.31
12	.56	40	.29

ที่มา: Lawshe (อ้างถึงในพิสนุ พงศ์ศรี, 2552, น.159; Taherdoost, 2014, p. 31; Zerati, & Alavi, 2014, p. 464)

การหา CVR คำนวณจากสูตร

$$CVR = \frac{Ne - N/2}{N/2}$$

โดยที่ CVR = อัตราส่วนความเที่ยงตรงเชิงเนื้อหาของข้อคำถาม

Ne = จำนวนผู้เชี่ยวชาญที่เห็นว่าข้อคำถามนั้นวัดคุณลักษณะตรงตามนิยาม

N = จำนวนผู้เชี่ยวชาญทั้งหมด (N/2 คือครึ่งหนึ่งของผู้เชี่ยวชาญทั้งหมด)

ตัวอย่างการคำนวณค่า CVR สมมติว่านักวิจัยสร้างแบบสอบถามมาจำนวน 6 ข้อ และนำไปให้ผู้เชี่ยวชาญจำนวน 5 ท่านพิจารณา เมื่อผู้เชี่ยวชาญลงความเห็นในแต่ละข้อคำถามแล้ว นักวิจัยจะนำความเห็นจากผู้เชี่ยวชาญมาลงในตาราง แล้วคำนวณค่า CVR ดังตารางที่ 5

ตารางที่ 5 ตัวอย่างการคำนวณค่า CVR

ข้อ \ คน	1	2	3	4	5	Ne	$CVR = \frac{Ne - N/2}{N/2}$	หมายเหตุ
1	1	1	1	1	1	5	$\frac{5-2.5}{2.5} = 1$	ใช้ได้
2	1	1	1	1	-1	4	$\frac{4-2.5}{2.5} = .60$	ใช้ไม่ได้
3	1	1	1	-1	-1	3	$\frac{3-2.5}{2.5} = .20$	ใช้ไม่ได้
4	1	1	-1	-1	-1	2	$\frac{2-2.5}{2.5} = -.20$	ใช้ไม่ได้
5	-1	-1	-1	-1	-1	0	$\frac{0-2.5}{2.5} = -1.0$	ใช้ไม่ได้
6	1	-1	-1	-1	-1	1	$\frac{1-2.5}{2.5} = -.60$	ใช้ไม่ได้

จากตารางที่ 5 จะเห็นได้ว่า ข้อที่มีค่า CVR ผ่านเกณฑ์ (พิจารณาจากตารางที่ 4 จำนวนผู้เชี่ยวชาญ 5 คน ค่า CVR $\geq .99$) มีเพียง 1 ข้อ คือข้อ 1 ส่วนข้อคำถามอื่นๆ ควรต้องปรับปรุง

วิธีที่สาม การหาดัชนีความเที่ยงตรงเชิงเนื้อหาของเครื่องมือ (Content Validity Index: CVI) เป็นวิธีการที่เสนอโดยนักวิชาการทางการศึกษาคือ Hambleton, Swaminathan, Alguna, & Coulson (1978) และพัฒนาต่อโดย Martuza (1977) ; Waltz, & Bausell (1981) โดย CVI จะใช้ในการหาความเที่ยงตรงของเครื่องมือวิจัยทั้งฉบับ (Gilbert, & Prion, 2016) หรือจะหาความเที่ยงตรงเป็นรายข้อก็ได้

โดยนักวิจัยจะนำเครื่องมือวิจัยไปให้ผู้เชี่ยวชาญพิจารณา โดยจำนวนผู้เชี่ยวชาญนั้น Lynn (as cited in Gilbert, & Prion, 2016; Polit, & Beck, 2006) เห็นว่าจำนวนผู้เชี่ยวชาญควรใช้ประมาณ 5-10 คน (ยกเว้นกรณีจำเป็น ไม่ควรต่ำกว่า 3 คน) แต่ไม่ควรเกิน 10 คนเพราะเกินความจำเป็น โดยผู้เชี่ยวชาญจะพิจารณาข้อคำถามและลงความเห็น 4 ระดับ คือ

- 1 หมายถึง ข้อคำถามไม่สามารถวัดคุณลักษณะได้ตรงตามนิยาม
- 2 หมายถึง ข้อคำถามต้องปรับปรุงมากจึงจะวัดคุณลักษณะได้ตรงตามนิยาม
- 3 หมายถึง ข้อคำถามต้องปรับปรุงเล็กน้อยจึงจะวัดคุณลักษณะได้ตรงตามนิยาม
- 4 หมายถึง ข้อคำถามวัดคุณลักษณะตรงตามนิยาม

ค่า CVI ที่นักวิชาการส่วนใหญ่ถือว่าเป็นเกณฑ์มักยึดตาม Davis (as cited in Polit, & Beck, 2006) ซึ่งแนะนำว่าต้องไม่น้อยกว่า .80 แต่ก็มีนักวิชาการจำนวนไม่น้อยที่มีความเห็นต่างอยู่บ้าง เช่น ทิพย์สิริ กาญจนวาสี, และศิริชัย กาญจนวาสี (2561); และ Tiden, Nelson, & May (1990) เห็นสอดคล้องกันว่า ค่า CVI ควรไม่น้อยกว่า .70 ขณะที่ Yaghmaie (อ้างถึงใน ประสพชัย พสุนนท์, 2558) ได้สำรวจบทความวิจัยในวารสารทางด้านการพยาบาลจำนวน 38 บทความ พบว่ามีการเลือกใช้ค่า CVI ไม่น้อยกว่า .75 ร่วมกับข้อเสนอแนะของผู้เชี่ยวชาญ หรือ Beck, & Owen (as cited in Gilbert, & Prion, 2016) แนะนำว่าค่า CVI ควรีค่าเกินกว่า .70 เป็นต้น สำหรับผู้เขียนขอใช้เกณฑ์ .80 ขึ้นไปตาม Davis

จากแนวคิดของ Lynn (as cited in Polit, & Beck, 2006) นั้น CVI สามารถแบ่งได้เป็น 2 ประเภท คือ Scale-level CVI (หรือ S-CVI) ซึ่งเป็นการหาค่าความเที่ยงตรงเชิงเนื้อหาของเครื่องมือทั้งชุด กับ Item-level CVI (หรือ I-CVI) ซึ่งเป็นการหาค่าความเที่ยงตรงเชิงเนื้อหาของข้อคำถามรายข้อ โดยมีสูตรการคำนวณ S-CVI และ I-CVI ดังนี้

การหา S-CVI คำนวณได้จากสูตร

$$S - CVI = \frac{\text{จำนวนข้อคำถามที่ผู้เชี่ยวชาญทุกคนลงความเห็นระดับ 3 และระดับ 4}}{\text{จำนวนข้อคำถามทั้งหมด}}$$

ตารางที่ 6 ตัวอย่างการคำนวณหาค่า S-CVI

ข้อที่	ระดับความเห็นของผู้เชี่ยวชาญ (1-4)					ทุกคนให้ความเห็นระดับ 3 หรือ 4
	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	
1	3	3	3	3	3	✓
2	4	4	4	4	4	✓
3	2	4	3	3	4	✗
4	4	4	3	3	4	✓
5	4	3	4	3	4	✓
6	3	3	4	4	3	✓
7	1	1	2	3	2	✗
8	3	4	4	3	4	✓
9	4	4	3	4	4	✓
10	3	4	3	3	3	✓
สรุปผล	$S - CVI = \frac{8}{10} = .80$ (ผ่านเกณฑ์ $\geq .80$)					

จากตารางที่ 6 จะเห็นได้ว่าจำนวนข้อคำถามที่ผู้เชี่ยวชาญทุกคนลงความเห็นระดับ 3 หรือ 4 นั้น มีจำนวน 8 ข้อ (คือข้อ 1, 2, 4, 5, 6, 8, 9 และ 10) เมื่อนำไปหารด้วยจำนวนข้อคำถามทั้งหมด 10 ข้อ จึงได้เท่ากับ .80 ซึ่งผ่านเกณฑ์ที่กำหนด อย่างไรก็ตาม แม้จะผ่านเกณฑ์ แต่ในทางปฏิบัติข้อเสนอแนะของผู้เชี่ยวชาญในข้อต่างๆ นักวิจัยสามารถนำมาใช้ในการปรับปรุงข้อคำถามเพื่อให้มีความเที่ยงตรงเชิงเนื้อหาได้

การหา I-CVI คำนวณได้จากสูตร

$$I - CVI = \frac{\text{จำนวนผู้เชี่ยวชาญที่ลงความเห็นระดับ 3 และระดับ 4 ในแต่ละข้อ}}{\text{จำนวนผู้เชี่ยวชาญทั้งหมด}}$$

ตารางที่ 7 ตัวอย่างการคำนวณหาค่า I-CVI

ข้อที่	ระดับความเห็นของผู้เชี่ยวชาญ (1-4)					จำนวนคนที่ลงความเห็นระดับ 3-4	การคำนวณ	แปลผล
	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5			
1	3	3	3	3	3	5	$I - CVI = \frac{3}{5} = 1.0$	ใช้ได้
2	4	4	4	4	4	5	$I - CVI = \frac{3}{5} = 1.0$	ใช้ได้
3	2	4	3	3	4	4	$I - CVI = \frac{4}{5} = .8$	ใช้ได้
4	4	4	3	3	4	5	$I - CVI = \frac{3}{5} = 1.0$	ใช้ได้
5	4	3	4	3	4	5	$I - CVI = \frac{3}{5} = 1.0$	ใช้ได้
6	3	3	4	4	3	5	$I - CVI = \frac{3}{5} = 1.0$	ใช้ได้
7	1	1	2	3	2	1	$I - CVI = \frac{1}{5} = .2$	ใช้ไม่ได้
8	3	4	4	3	4	5	$I - CVI = \frac{3}{5} = 1.0$	ใช้ได้
9	4	4	3	4	4	5	$I - CVI = \frac{3}{5} = 1.0$	ใช้ได้
10	3	4	3	3	3	5	$I - CVI = \frac{3}{5} = 1.0$	ใช้ได้

จากตารางที่ 7 จะเห็นได้ว่ามีเพียงข้อคำถามที่ 7 เท่านั้นที่ไม่ผ่านเกณฑ์ หรือใช้ไม่ได้ เนื่องจากค่า I-CVI ที่คำนวณได้มีค่าเท่ากับ .20 ซึ่งต่ำกว่าเกณฑ์ (I-CVI ต้องไม่น้อยกว่า .80)

สรุปเปรียบเทียบและอภิปรายการตรวจสอบความเที่ยงตรงเชิงเนื้อหาด้วย IOC, CVR และ CVI

ในส่วนนี้ผู้เขียนจะสรุปเปรียบเทียบและอภิปรายการตรวจสอบความเที่ยงตรงเชิงเนื้อหาด้วย IOC, CVR และ CVI ใน 8 ประเด็น ได้แก่ 1) ผู้ริเริ่มและปีที่ริเริ่ม 2) จำนวนผู้เชี่ยวชาญที่ต้องใช้ในการพิจารณาเครื่องมือ 3) เกณฑ์การให้คะแนน 4) เกณฑ์ตัดสินการใช้ได้ของข้อคำถาม 5) การใช้ในการตรวจสอบคุณภาพของข้อคำถามเป็นรายข้อ 6) การใช้ในการพิจารณาเปรียบเทียบคุณภาพของเครื่องมือทั้งฉบับ 7) จุดเด่น และ 8) ข้อจำกัด

ประเด็นแรก ผู้ริเริ่มและปีที่ริเริ่ม ผลการเปรียบเทียบพบว่าทั้ง IOC, CVR และ CVI มาจากการริเริ่มของนักวิชาการที่ต่างคนและต่างปีกัน โดย IOC ริเริ่มขึ้นในปี 1977 โดย Rovinelli, & Hambleton (1977) ขณะที่ CVR ริเริ่มในปี 1975 โดย Lawshe (อ้างถึงในทิพย์สิริ กาญจนวาสิ, และศิริชัย กาญจนวาสิ, 2561; ศิริชัย กาญจนวาสิ, 2552; Ayre, & Scally, 2014; Gilbert, & Prion, 2016) ส่วน CVI นั้นริเริ่มโดยนักวิชาการด้านการศึกษาคือ Hambleton, Swaminathan, Alguna, & Coulson, (1978) ดังนั้นในประเด็นแรก กล่าวได้ว่า ทั้ง IOC, CVR และ CVI ต่างถูกพัฒนามาจากนักวิชาการต่างคน โดย CVR ถูกพัฒนามาก่อน IOC 2 ปี และก่อน CVI 3 ปี ซึ่งถือเป็นช่วงเวลาที่ไม่ต่างกันมากนัก

ประเด็นที่สอง จำนวนผู้เชี่ยวชาญที่ต้องใช้ในการพิจารณาเครื่องมือ ผลการเปรียบเทียบพบว่า IOC ต้องใช้ผู้เชี่ยวชาญในการพิจารณาเครื่องมืออย่างน้อย 3 คนขึ้นไป ขณะที่ CVR ต้องใช้ผู้เชี่ยวชาญ 5-40 คน ส่วน CVI โดยทั่วไปจะใช้ผู้เชี่ยวชาญ 5-10 คน เว้นแต่เป็นกรณีจำเป็น เช่นหาผู้เชี่ยวชาญได้ยากจึงอาจใช้ผู้เชี่ยวชาญ 3 คนได้ หากพิจารณาในแง่นี้จะพบว่า IOC มีความง่ายในทางปฏิบัติมากที่สุดในการหาผู้เชี่ยวชาญเพื่อพิจารณา (โดยเฉพาะอย่างยิ่งกรณีที่นักวิจัยมีเวลาในการทำวิจัยจำกัด และหาผู้เชี่ยวชาญได้ยาก) รองลงมาคือ CVI เพราะแม้จะกำหนดโดยทั่วไปว่าต้องใช้ผู้เชี่ยวชาญ 5-10 คนแต่ก็มีการเปิดช่องให้ในกรณีจำเป็นให้ใช้ผู้เชี่ยวชาญ 3 คนได้ ขณะที่ CVR ถือว่าเป็นการกำหนดจำนวนผู้เชี่ยวชาญที่ค่อนข้างเข้มงวดในปริมาณคนที่ต้องไม่ต่ำกว่า 5 คน

เป็นอย่างน้อยซึ่งอาจเป็นการยากโดยเฉพาะในกรณีที่งานวิจัยในเรื่องนั้นๆ หาผู้เชี่ยวชาญได้ยาก อย่างไรก็ตามในกรณีปกติทั่วไปคงไม่เป็นการยากเกินไปนักในการหาผู้เชี่ยวชาญ 5 คนเพื่อพิจารณา

ประเด็นที่สาม เกณฑ์การให้คะแนน ผลการเปรียบเทียบพบว่า IOC กำหนดเกณฑ์การให้คะแนนเป็น 3 ค่า ได้แก่ +1 หมายถึง ข้อคำถามสามารถวัดได้ตรงตามนิยาม 0 หมายถึง ไม่แน่ใจว่า ข้อคำถามสามารถวัดได้ตรงตามนิยามหรือไม่ และ -1 หมายถึง ข้อคำถามไม่สามารถวัดได้ตรงตามนิยาม ส่วน CVR กำหนดเกณฑ์การให้คะแนนเพียง 2 ค่า คือ +1 หมายถึง ข้อคำถามสามารถวัดได้ตรงตามนิยาม และ -1 หมายถึง ข้อคำถามไม่สามารถวัดได้ตรงตามนิยาม ส่วน CVI กำหนดเกณฑ์การให้คะแนนมากถึง 4 ค่า คือ 1 หมายถึง ข้อคำถามไม่สามารถวัดได้ตรงตามนิยาม 2 หมายถึง ข้อคำถามต้องปรับปรุงมาก 3 หมายถึง ข้อคำถามต้องปรับปรุงเล็กน้อย และ 4 หมายถึง ข้อคำถามวัดตรงตามนิยาม หากเปรียบเทียบในแง่นี้จะพบว่า CVR กำหนดเกณฑ์การให้คะแนนที่บังคับให้ผู้เชี่ยวชาญต้องตัดสินใจแบบฟันธงว่า ข้อคำถามนั้นจะสามารถวัดได้หรือวัดไม่ได้ตามนิยาม และไม่เปิดช่องให้กับความลังเลสงสัยว่าข้อคำถามนั้นจะใช้ได้หรือไม่ ขณะที่ IOC มีการเปิดช่องให้ผู้เชี่ยวชาญที่ไม่แน่ใจในข้อคำถามนั้นสามารถให้ค่าเป็น 0 ได้ ส่วน CVI นั้นถือได้ว่าเป็นการกำหนดเกณฑ์การให้คะแนนที่ละเอียดที่สุดเพราะมีการกำหนดไว้ถึง 4 ระดับ แต่เป็นการพิจารณาในแง่ของการใช้ได้ว่าข้อคำถามนั้นต้องมีการปรับปรุงเล็กน้อยเพียงใด ทั้งนี้ผู้เขียนเห็นว่า ณ จุดนี้เองถือเป็นการตอกย้ำกับนักวิจัยว่า การเลือกผู้เชี่ยวชาญในการพิจารณาเครื่องมือวิจัยนั้นจำเป็นต้องเลือกผู้เชี่ยวชาญที่มีความเชี่ยวชาญในเรื่องนั้นๆ จริงๆ เพราะผู้เชี่ยวชาญย่อมรู้ว่าข้อคำถามนั้นจะใช้ได้หรือไม่ หรือต้องปรับปรุงแก้ไขเล็กน้อยเพียงใด

ประเด็นที่สี่ เกณฑ์ตัดสินการใช้ได้ของข้อคำถาม ผลการเปรียบเทียบพบว่า IOC กำหนดเกณฑ์ขั้นต่ำในการพิจารณาการใช้ได้ของข้อคำถามไว้ที่ .50 อย่างชัดเจนและเป็นที่ยอมรับร่วมกันของนักวิชาการโดยทั่วไป ขณะที่ CVR การพิจารณาว่าข้อคำถามใช้ได้หรือไม่นั้น เกณฑ์ในการพิจารณาจะผูกผันไปตามจำนวนผู้เชี่ยวชาญ หากใช้ผู้เชี่ยวชาญจำนวนน้อย เช่น 5 คน จะตัดสินการใช้ได้ของข้อคำถามที่ .99 แต่หากนักวิจัยใช้ผู้เชี่ยวชาญมากถึง 40 คน จะตัดสินการใช้ได้ของข้อคำถามเพียง .29

นั่นคือยิ่งใช้จำนวนผู้เชี่ยวชาญมากขึ้นเท่าไร เกณฑ์ตัดสินการใช้ได้ของข้อคำถามจะยิ่งลดลง ส่วน CVI การพิจารณาตัดสินการใช้ได้ของข้อคำถามมีหลากหลาย เช่น .80, .75 และ .70 เป็นต้น มิได้มีเกณฑ์ใดที่เป็นที่ยอมรับร่วมกันทั้งหมด สุดแต่แต่นักวิจัยจะอ้างอิงเกณฑ์การตัดสินจากใคร หากพิจารณาในแง่มุมนี้ ผู้เขียนเห็นว่าเกณฑ์ของ IOC ค่อนข้างชัดเจนมากที่สุด คือ ไม่ว่าจะใช้ผู้เชี่ยวชาญจำนวนกี่คนก็ยังคงใช้เกณฑ์เดียวกันคือ .50 ซึ่งต่างจาก CVR ที่ผกผันไปตามจำนวนผู้เชี่ยวชาญ อีกทั้งความเห็นต่อเกณฑ์การตัดสินการใช้ได้ของข้อคำถามก็ชัดเจนตรงกัน ไม่หลากหลายแบบ CVI อีกด้วย

ประเด็นที่ห้า การใช้ในการตรวจสอบคุณภาพของข้อคำถามเป็นรายชื่อ
ผลการเปรียบเทียบพบว่าทั้ง IOC, CVR และ CVI สามารถใช้ในการตรวจสอบคุณภาพของข้อคำถามเป็นรายชื่อได้ทั้งสิ้น เพียงแต่จะแตกต่างกันในแง่ของค่าเกณฑ์ที่ใช้ในการตัดสินซึ่ง IOC และ CVI (หมายถึง I-CVI) จะง่ายในการผ่านเกณฑ์มากกว่า CVR ตัวอย่างเช่น สมมตินักวิจัยใช้ผู้เชี่ยวชาญ 5 คน ในการพิจารณาข้อคำถาม หากมีผู้เชี่ยวชาญ 1 คน เห็นว่าข้อคำถามนั้นไม่สามารถวัดได้ตรงตามนิยาม ขณะที่อีก 4 คน เห็นว่าสามารถวัดได้ตรงตามนิยาม หากเป็นวิธีการของ IOC ข้อคำถามข้อนี้จะได้ค่า IOC เท่ากับ .60 (เกิดจากการนำคะแนน -1, +1, +1, +1 และ +1 หารด้วยจำนวนผู้เชี่ยวชาญทั้งหมด) ซึ่งถือว่าผ่านเกณฑ์ที่ .50 ขณะที่ถ้าใช้ CVR แม้จะได้ค่า .60 เท่ากัน (เกิดจากการนำจำนวนผู้เชี่ยวชาญที่เห็นว่ ข้อคำถามนั้นวัดคุณลักษณะตรงตามนิยามไปด้วยครึ่งหนึ่งของผู้เชี่ยวชาญทั้งหมด และนำค่าที่ได้มา หารด้วยครึ่งหนึ่งของผู้เชี่ยวชาญทั้งหมด) แต่ไม่ผ่านเกณฑ์เนื่องจากเกณฑ์ตัดสินที่ .99 ส่วนถ้าใช้ CVI ข้อคำถามนี้จะได้ค่า I-CVI เท่ากับ .80 (เกิดจากการนำจำนวนผู้เชี่ยวชาญที่ลงความเห็นระดับ 3 และ 4 หารด้วยจำนวนผู้เชี่ยวชาญทั้งหมด) ซึ่งถือว่าผ่านเกณฑ์

ประเด็นที่หก การใช้ในการพิจารณาคุณภาพของเครื่องมือทั้งฉบับ ผลการเปรียบเทียบพบว่า IOC ไม่สามารถนำไปใช้ในการพิจารณาตัดสินคุณภาพของเครื่องมือทั้งฉบับ ขณะที่ CVR สามารถทำได้ โดยการรวมค่า CVR ที่ได้จากข้อคำถามในแต่ละข้อและนำไปหารด้วยจำนวนข้อทั้งหมด แต่การรวมดังกล่าวอาจถูกพิจารณาเรียกชื่อเป็นอย่างอื่น เช่น ทิพย์สิริ กาญจนवासี, และศิริชัย กาญจนवासี (2561) ได้อธิบายวิธีการดังกล่าวว่าเป็นการหา CVI) ส่วน CVI นั้นสามารถหาความเที่ยงตรงเชิงเนื้อหาของ

แบบสอบถามทั้งฉบับได้ ซึ่งกรณีนี้จะเรียกว่า S-CVI โดยที่นักวิจัยสามารถเลือกได้อีกว่าจะพิจารณาจากความเห็นพ้องต้องกัน (Universal Agreement) ที่เรียกว่า S-CVI/UA แบบที่ผู้เขียนได้นำมาใช้ในการอธิบาย หรือ จะใช้การพิจารณาจากการเฉลี่ย (Average) ที่เรียกว่า S-CVI/Ave ซึ่งสามารถกระทำได้หลากหลายแบบ

ประเด็นที่เจ็ด จุดเด่น ผลการเปรียบเทียบพบว่า IOC มีจุดเด่นคือ เป็นที่นิยมใช้ในการตรวจสอบความเที่ยงตรงเชิงเนื้อหาในการวิจัยทางสังคมศาสตร์โดยทั่วไป (รวมถึงการวิจัยทางรัฐประศาสนศาสตร์) ซึ่งหนังสือหรือตำราทางด้านระเบียบวิธีวิจัยส่วนใหญ่มักกล่าวถึงมากกว่าวิธีอื่นๆ อีกทั้งการนำ IOC ไปใช้ยังง่ายและสะดวกในทางปฏิบัติมากกว่า ทั้งในแง่การทำความเข้าใจเพราะมีวิธีการในการคิดคำนวณที่ไม่ยุ่งยากซับซ้อน ในแง่จำนวนผู้เชี่ยวชาญที่ต้องใช้ก็กำหนดไว้เพียง 3 คน ขึ้นไป และในแง่เกณฑ์ที่ใช้ในการตัดสินก็กำหนดไว้เพียง .50 เท่านั้น ขณะที่ CVR มีจุดเด่นคือ เนื่องจากความยากในเกณฑ์ที่ใช้ในการตัดสินว่าข้อคำถามใดใช้ได้นั้นกำหนดเกณฑ์ไว้สูงอย่างมาก หรือหากต้องการลดเกณฑ์ลง ก็ต้องเพิ่มจำนวนผู้เชี่ยวชาญในการพิจารณาซึ่งอาจมากได้ถึง 40 คน ซึ่งหมายความว่าคุณภาพของข้อคำถามข้อนั้นที่ผ่านเกณฑ์ของ CVR ย่อมมีความมั่นใจได้ต่อความเที่ยงตรงเชิงเนื้อหาเป็นอย่างมาก ส่วน CVI มีจุดเด่นที่นักวิจัยสามารถคำนวณหาค่าความเที่ยงตรงของเครื่องมือได้ทั้งรายข้อและทั้งฉบับ ซึ่งเป็นข้อจำกัดที่ IOC ไม่สามารถทำได้ นอกจากนี้ CVI ยังมักเป็นวิธีการหาความเที่ยงตรงเชิงเนื้อหาของการวิจัยบางสาขา เช่น สาขาการพยาบาลซึ่งนิยมใช้วิธีการดังกล่าวมากที่สุด (สุจิตรา เทียนสวัสดิ์, 2562) อีกด้วย

ประเด็นสุดท้าย ข้อจำกัด ผลการเปรียบเทียบพบว่า IOC มีข้อจำกัดคือ ไม่สามารถหาความเที่ยงตรงเชิงเนื้อหาของเครื่องมือวิจัยแบบทั้งฉบับได้ ทำได้เพียงการรายงานเป็นรายข้อว่ามีข้อใดผ่านเกณฑ์หรือไม่ผ่านเกณฑ์ ขณะที่ CVR นั้นข้อจำกัดที่เห็นได้ชัดคือ ความยากในการปฏิบัติมากกว่าวิธีการอื่นๆ เนื่องจากต้องใช้จำนวนผู้เชี่ยวชาญมาก (5-40 คน) โดยไม่มีการยืดหยุ่นให้ในกรณีจำเป็นที่หาผู้เชี่ยวชาญได้ยาก อีกทั้งยังกำหนดค่าที่ผ่านเกณฑ์ไว้สูงมากเมื่อเทียบกับวิธีอื่นๆ นอกจากนี้ในการวิจัยทางสังคมศาสตร์โดยทั่วไปโดยเฉพาะการวิจัยทางรัฐประศาสนศาสตร์แล้ว CVR มักไม่ค่อยปรากฏว่า

มีการนำไปใช้ในการตรวจสอบความเที่ยงตรงเชิงเนื้อหา หนังสือหรือตำราการวิจัยที่อธิบายเรื่องนี้มีจำนวนน้อยกว่า กล่าวอีกนัยหนึ่งได้ว่า CVR ไม่ค่อยเป็นที่แพร่หลายหรือไม่ค่อยเป็นที่รู้จักมากนักโดยเฉพาะเมื่อเทียบกับ IOC ส่วนกรณีของ CVI มีข้อจำกัดคือ มีความหลากหลายยุ่งยากต่อการทำความเข้าใจมากกว่า IOC และ CVR เพราะมีทั้งการหาความเที่ยงตรงเชิงเนื้อหาของเครื่องมือทั้งฉบับ และแบบรายข้อ และในกรณีของการหาความเที่ยงตรงเชิงเนื้อหาของทั้งฉบับก็สามารถทำได้หลากหลายวิธีการ อีกทั้งเกณฑ์ในการพิจารณาตัดสินว่าข้อคำถามหรือเครื่องมือทั้งฉบับนั้นผ่านเกณฑ์หรือไม่ยังไม่ชัดเจนที่ใดเป็นที่ยอมรับร่วมกันว่าจะตัดสินผ่านที่ค่าใด จะใช้เกณฑ์ .80, .75 หรือ .70 เป็นต้น นอกจากนี้ยังเป็นที่แพร่หลายในการวิจัยทางรัฐประศาสนศาสตร์น้อยกว่า IOC อีกด้วย³

ตารางที่ 8 เปรียบเทียบการตรวจสอบความเที่ยงตรงเชิงเนื้อหาด้วย IOC, CVR และ CVI

ประเด็นเปรียบเทียบ	วิธีการในการตรวจสอบความเที่ยงตรงเชิงเนื้อหา		
	IOC	CVR	CVI
1. ผู้ริเริ่ม (ปี)	Rovinelli, & Hambleton (1977)	Lawshe (1975)	Hambleton, Swaminathan, Alguna, & Coulson (1978)
2. จำนวนผู้เชี่ยวชาญที่ต้องใช้ในการพิจารณาเครื่องมือ	3 คนขึ้นไป	5 – 40 คน	5 - 10 คน (ยกเว้นกรณีจำเป็นต้องไม่ต่ำกว่า 3 คน)

³ การอธิบายว่า IOC เป็นที่นิยมแพร่หลายกว่า CVR และ CVI นั้น เป็นเพียงการกล่าวอ้างของผู้เขียนซึ่งได้มาซึ่งความรู้ ความจริง โดยอาศัยประสบการณ์ส่วนตัวที่เป็นนักวิจัย เป็นอาจารย์ที่ปรึกษาวิจัย เป็นผู้ทรงคุณวุฒิอ่านงานวิจัยของผู้อื่น และจากการค้นคว้าหนังสือและตำราต่างๆ เพื่อประกอบการเขียนหนังสือระเบียบวิธีวิจัยทางรัฐศาสตร์และรัฐประศาสนศาสตร์ มิได้มาจากข้อมูลผลการวิจัยเชิงประจักษ์ ดังนั้นจุดนี้อาจตั้งไว้เป็นข้อสังเกต ซึ่งอาจต้องรอการพิสูจน์ยืนยันกันต่อไปในอนาคต

ประเด็น เปรียบเทียบ	วิธีการในการตรวจสอบความเที่ยงตรงเชิงเนื้อหา		
	IOC	CVR	CVI
3. เกณฑ์การให้ คะแนน	+1 = ข้อคำถามสามารถ วัดได้ตรงตามนิยาม 0 = ไม่แน่ใจว่าข้อคำถาม สามารถวัดได้ตรงตาม นิยามหรือไม่ -1 = ข้อคำถามไม่ สามารถวัดได้ตรงตาม นิยาม	+1 = ข้อคำถามสามารถ วัดได้ตรงตามนิยาม -1 = ข้อคำถามไม่สามารถ วัดได้ตรงตามนิยาม	1 = ข้อคำถามไม่สามารถ วัดได้ตรงตามนิยาม 2 = ข้อคำถามต้อง ปรับปรุงมาก 3 = ข้อคำถามต้อง ปรับปรุงเล็กน้อย 4 = ข้อคำถามวัดตรงตาม นิยาม
4. เกณฑ์ตัดสิน การใช้ได้ของข้อ คำถาม	.50 ขึ้นไป (ปรับเกณฑ์ขึ้น ได้ แต่ลดไม่ได้)	ผันแปรตามจำนวน ผู้เชี่ยวชาญ เช่น ผู้เชี่ยวชาญ 5 คน ค่า CVR ต้องไม่ต่ำกว่า .99 แต่ถ้า ผู้เชี่ยวชาญ 40 คน ค่า CVR ต้องไม่ต่ำกว่า .29 เป็นต้น	หลากหลาย เช่น .80, .75 และ .70
5. การใช้ในการ ตรวจสอบ คุณภาพของข้อ คำถามเป็นรายข้อ	ทำได้	ทำได้	ทำได้
6. การใช้ในการ พิจารณาคุณภาพ ของเครื่องมือทั้ง ฉบับ	ทำไม่ได้	ทำได้ แต่อาจเรียกชื่อเป็น อย่างอื่น (งานของ ทิพย์สิริ กาญจนวาสี, และศิริชัย กาญจนวาสี ใช้การหา ค่าเฉลี่ยของ CVR ในแต่ละ	ทำได้

ประเด็น เปรียบเทียบ	วิธีการในการตรวจสอบความเที่ยงตรงเชิงเนื้อหา		
	IOC	CVR	CVI
		ข้อ แต่เรียกว่าเป็นการหา CVI)	
7. จุดเด่น	<ul style="list-style-type: none"> - นิยมใช้ในการวิจัยทางสังคมศาสตร์โดยทั่วไป รวมถึงการวิจัยทางรัฐประศาสนศาสตร์ - ง่ายในทางปฏิบัติกว่า CVR และ CVI ทั้งในแง่การทำความเข้าใจ จำนวนผู้เชี่ยวชาญที่ใช้ และ เกณฑ์การตัดสินว่า ข้อคำถามใช้ได้ 	<ul style="list-style-type: none"> - มั่นใจได้ใน ความเที่ยงตรงเชิงเนื้อหาของข้อคำถามที่ผ่านเกณฑ์มากกว่าวิธีการอื่น ๆ (เนื่องจากความยากในเกณฑ์ที่ใช้ในการตัดสินว่า ข้อคำถามใดใช้ได้ นั้น กำหนดเกณฑ์ไว้สูง) 	<ul style="list-style-type: none"> - สามารถหาความเที่ยงตรงเชิงเนื้อหาได้ทั้งรายข้อและทั้งฉบับ - นิยมใช้ในการวิจัยบางสาขา เช่น การพยาบาล
8. ข้อจำกัด	<ul style="list-style-type: none"> - ไม่สามารถหาความเที่ยงตรงของเครื่องมือวิจัยทั้งฉบับได้ 	<ul style="list-style-type: none"> - ความยากในการปฏิบัติมากกว่าวิธีการอื่น ๆ เนื่องจากต้องใช้จำนวนผู้เชี่ยวชาญมาก (5-40 คน) อีกทั้งยังกำหนดค่าที่ผ่านเกณฑ์ไว้สูง - มีความแพร่หลายในการใช้น้อยกว่าวิธีการอื่น ๆ (โดยเฉพาะในการวิจัยทางรัฐประศาสนศาสตร์) 	<ul style="list-style-type: none"> - มีความหลากหลายยุ่งยากต่อการทำความเข้าใจมากกว่า IOC และ CVR - เกณฑ์ที่ใช้ในการตัดสินว่า ข้อคำถาม หรือ เครื่องมือทั้งฉบับผ่านเกณฑ์หรือไม่นั้น ยังมี ความหลากหลาย - การนำไปใช้ในการวิจัยทางรัฐประศาสนศาสตร์ ยังน้อยกว่า IOC

จากที่กล่าวมาทั้งหมดจึงนำไปสู่การสรุปปิดท้ายเรื่องของผู้เขียนว่า การตรวจสอบความเที่ยงตรงเชิงเนื้อหาของเครื่องมือวิจัย ไม่ว่าจะโดยการใช้ IOC, CVR หรือ CVI ต่างมีจุดเด่นและข้อจำกัดในตัวของตัวเอง ซึ่งผู้เขียนไม่ประสงค์บ่งชี้ว่าวิธีการใดดีที่สุด เพราะเห็นว่าไม่มีวิธีการใดที่ดีที่สุด (No One Best Way) นักวิจัยที่ต้องการใช้วิธีการเหล่านี้พึงตระหนักในจุดเด่นและข้อจำกัดดังได้กล่าวไปแล้ว และที่สำคัญคือ แต่ละวิธีการล้วนมีหลักการหรือเกณฑ์ต่างๆ ในการคิดคำนวณค่าที่แตกต่างกัน หากเลือกใช้วิธีการใดแล้วคงต้องศึกษาและอ้างอิงเกณฑ์นั้นๆ ให้ถูกต้องแม่นยำ ซึ่งจะทำให้การอ้างอิงผลการตรวจสอบความเที่ยงตรงเชิงเนื้อหาของนักวิจัยได้รับการยอมรับมากกว่าการขาดการตรวจสอบคุณภาพดังกล่าวเลย อย่างไรก็ตามพึงตระหนักว่าผู้เขียนได้กล่าวไปทั้งหมดเป็นเพียงวิธีการที่นำมาใช้ในการตรวจสอบความเที่ยงตรงเชิงเนื้อหาเท่านั้น การตรวจสอบความเที่ยงตรงของเครื่องมือวิจัยยังมีกรณีอื่นๆ อีกด้วย อีกทั้งหากพิจารณาให้กว้างกว่านั้น เครื่องมือวิจัยยังจำเป็นต้องมีการตรวจสอบคุณภาพในแง่มุมอื่นๆ ทั้งความเชื่อมั่น อำนาจจำแนก หรือความยากง่ายอีกด้วย ซึ่งทั้งหมดเป็นเรื่องที่อยู่นอกเหนือขอบเขตของบทความชิ้นนี้แต่เป็นเรื่องสำคัญมิใช่น้อยต่อนักวิจัยโดยเฉพาะอย่างยิ่งต่อนักวิจัยเชิงปริมาณ

เอกสารอ้างอิง

- ชูศรี วงศ์รัตน์. (2549). เทคนิคการเขียนเค้าโครงการวิจัย: แนวทางสู่ความสำเร็จ. กรุงเทพฯ: ไทเนรมิตกิจ
อินเตอร์ โปรดักส์.
- โชติกา ภาณีผล. (2554). การสร้างและพัฒนาเครื่องมือในการวัดและประเมินผลการศึกษา (พิมพ์ครั้งที่ 2).
กรุงเทพฯ: คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ทิพย์สิริ กาญจนวาสี, และศิริชัย กาญจนวาสี. (2561). วิธีวิทยาการวิจัย. กรุงเทพฯ: พิมพ์หัวใจ.
- บุญใจ ศรีสถิตยัณรากร. (2555). การพัฒนาและตรวจสอบคุณภาพเครื่องมือวิจัย: คุณสมบัติการวัดเชิง
จิตวิทยา. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ประสพชัย พสุนนท์. (2558). ความเที่ยงตรงของแบบสอบถามสำหรับงานวิจัยทางสังคมศาสตร์. วารสาร
สังคมศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ, 18(1), 375-396.
- พรเพ็ญ เพชรสุขศิริ. (2541). การสร้างมาตรฐานวัด (พิมพ์ครั้งที่ 3). กรุงเทพฯ: คณะรัฐประศาสนศาสตร์
สถาบันบัณฑิตพัฒนบริหารศาสตร์.
- พิสนุ พงศ์ศรี. (2552). การสร้างและพัฒนาเครื่องมือวิจัย. กรุงเทพฯ: ด้านสุทธากการพิมพ์.
- วงพักตร์ ภูพันธ์ศรี, และวีระยุทธ พรพจน์ธนาศ. (2560). การพัฒนาทรัพยากรมนุษย์ที่ส่งผลต่อ
ประสิทธิภาพองค์กร: กรณีศึกษาองค์กรปกครองส่วนท้องถิ่นในจังหวัดระยอง. ชลบุรี:
คณะรัฐศาสตร์และนิติศาสตร์ มหาวิทยาลัยบูรพา.
- ศิริชัย กาญจนวาสี. (2552). ทฤษฎีการทดสอบแบบดั้งเดิม (พิมพ์ครั้งที่ 6). กรุงเทพฯ: โรงพิมพ์แห่ง
จุฬาลงกรณ์มหาวิทยาลัย.

ศุภชัย ยาวะประภาษ. (2544). *การวิจัยทางรัฐประศาสนศาสตร์* (พิมพ์ครั้งที่ 3). กรุงเทพฯ: คณะรัฐศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

สิน พันธุ์พินิจ. (2551). *เทคนิคการวิจัยทางสังคมศาสตร์* (พิมพ์ครั้งที่ 1). กรุงเทพฯ: วิทญพัฒน์.

สุจิตรา เทียนสวัสดิ์. (2562). *การพัฒนาเครื่องมือสำหรับการวิจัยทางการพยาบาล*. เชียงใหม่: สยามพิมพ์ นานา.

สุจิตรา บุญยรัตพันธุ์. (2563). *ระเบียบวิธีวิจัยสำหรับรัฐประศาสนศาสตร์* (พิมพ์ครั้งที่ 19). กรุงเทพฯ : คณะรัฐประศาสนศาสตร์ สถาบันบัณฑิตพัฒนบริหารศาสตร์.

สุชาติ ประสิทธิ์รัฐสินธุ์. (2546). *ระเบียบวิธีการวิจัยทางสังคมศาสตร์* (พิมพ์ครั้งที่ 12). กรุงเทพฯ: สามลดา.

Ayre, C., & Scally, A. J. (2014). Critical Values for Lawshe's Content Validity Ratio: Revisiting the Original Methods of Calculation. *Measurement and Evaluation in Counseling and Development*, 47(1), 79-86.

Babbie, E. (2010). *The Practice of Social Research* (12th ed.). Canada: Wadsworth, Cengage Learning.

Gay, L. R., Mills, G. E., & Airasian, P. W. (2014). *Educational Research: Competencies for Analysis and Applications* (10th ed). U.S.A.: Pearson education.

Gilbert, G. E., & Prion, S. (2016). Making Sense of Methods and Measurement: Lawshe's Content Validity Index. *Clinical Simulation in Nursing*, 12(1), 530-531.

Hambleton, R. K., Swaminathan, H., Algina, J. Y., & Coulson, D. B. (1978). Criterion-referenced testing and measurement: A review of Technical Issues and developments. *Review of Educational Research*, 48(1), 1-47.

- Martuza, V. R. (1977). *Applying norm-referenced measurement in education*. Boston: Allyn & Bacon.
- Polit, D. F., & Beck, C. T. (2006). The Content Validity Index: Are You Sure You Know What's Being Reported? Critique and Recommendations. *Research in Nursing & Health*, 29(1), 489-497.
- Rovinelli, R. J., & Hambleton, R. K. (1977). On the use of content specialists in the assessment of criterion-referenced test item validity. *Tijdschrift voor Onderwijsresearch*, 2(2), 49-60.
- Taherdoost, H. (2014). Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research. *International Journal of Academic Research in Management*, 5(3), 28-36.
- Tiden, V. P., Nelson, C. A., & May, B. A. (1990). Use of qualitative methods to enhance content validity. *Nursing research*, 39(3), 172-175.
- Waltz, C., & Bausell, R. B. (1981). *Nursing research: Designs statistics and computer analysis*. Philadelphia: F.A. Davis.
- Yousoff, M.S.B. (2019). ABC of Content Validation and Content Validity Index Calculation. *Education in Medicine Journal*, 11(2), 49-54.
- Zerati, M., & Alavi, N. (2014). Designing and Validity Evaluation of Quality of Nursing Care Scale in Intensive Care Units. *Journal of Nursing Measurement*, 22(3), 461-471.