



สมรรถนะทางพหุวัฒนธรรม: ความตรงและกลยุทธ์การประเมิน
ในบริบทที่มีความหลากหลายทางวัฒนธรรม*
MULTICULTURAL COMPETENCE: VALIDITY AND ASSESSMENT STRATEGIES
IN DIVERSE CULTURAL CONTEXTS

นลินี ณ นคร^{1*}, สักรณ จัตุระโท¹, กาญจนา วัฒนสุนทร¹, เจนณรงค์ พินลันทุม²

Nalinee Na Nakorn^{1*}, Sungworn Ngudgratoke¹, Kanjana Wattanasuntorn¹, Jennarong Pinlantum²

¹สาขาวิชาศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช นนทบุรี ประเทศไทย

¹School of Educational Studies, Sukhothai Thammathirat Open University, Nonthaburi, Thailand

²คณะศึกษาศาสตร์และศิลปศาสตร์ มหาวิทยาลัยหาดใหญ่ สงขลา ประเทศไทย

²Faculty of Education and Liberal Arts, Hatyai University, Songkhla, Thailand

*Corresponding author E-mail: nalinee_na_nakorn@hotmail.com

บทคัดย่อ

สมรรถนะทางพหุวัฒนธรรมเป็นแนวคิดแบบพหุมิติที่ประกอบด้วยความตระหนักรู้ ความรู้ และทักษะ ซึ่งสะท้อนการเรียนรู้แบบองค์รวมในบริบทที่มีความหลากหลายทางวัฒนธรรม การประเมินสมรรถนะนี้ยังเผชิญข้อจำกัดหลายประการ อาทิ ความไม่ชัดเจนขององค์ประกอบ การใช้คำว่า “สมรรถนะทางพหุวัฒนธรรม” และ “สมรรถนะข้ามวัฒนธรรม” แทนกันโดยไม่แยกความหมายอย่างชัดเจน และความเอนเอียงของการประเมินแบบรายงานตนเองที่อาจไม่สะท้อนพฤติกรรมจริง การประเมินที่มีประสิทธิภาพจึงต้องอาศัยกรอบแนวคิดที่ชัดเจน และมีหลักฐานเชิงประจักษ์รองรับการแปลความหมายผลการวัดอย่างสมเหตุสมผล โดยยึดหลักความตรงแบบองค์รวมของ Messick ที่เน้นว่าความตรงของเครื่องมือวัดควรพิจารณาจากหลายมิติ ทั้งเนื้อหา สาระสำคัญ โครงสร้าง การสรุปอ้างอิง ความสัมพันธ์กับตัวแปรอื่น และผลกระทบจากการใช้คะแนน ร่วมกับกรอบแนวคิดการตรวจสอบความตรงด้วยการโต้แย้งของ Kane ซึ่งเน้นการจัดเรียงข้ออนุมานที่นำไปสู่การตัดสินใจจากผลการประเมิน และการจัดหาหลักฐานเชิงประจักษ์เพื่อรองรับข้ออนุมานในแต่ละขั้นของกระบวนการให้เหตุผลอย่างเป็นระบบ กลยุทธ์การประเมินที่เหมาะสมควรมีความหลากหลายและผสมผสานวิธีการหลายวิธี เช่น การรายงานตนเอง การทดสอบการตัดสินใจตามสถานการณ์ การสังเกตพฤติกรรม การวิเคราะห์ความไม่แปรเปลี่ยนของการวัดระหว่างกลุ่ม การมีส่วนร่วมของผู้มีส่วนได้ส่วนเสีย เพื่อให้ได้ผลการประเมินที่สะท้อนสมรรถนะได้อย่างถูกต้อง เชื่อถือได้ และตอบสนองต่อความหลากหลายทางวัฒนธรรม บทความนี้จึงเสนอแนวทางการประเมินที่รอบด้าน เพื่อสนับสนุนการพัฒนาเครื่องมือที่มีความตรง ใช้งานได้จริง และเชื่อมโยงกับเป้าหมายของสมรรถนะระดับโลก ในศตวรรษที่ 21

คำสำคัญ: สมรรถนะทางพหุวัฒนธรรม, ความตรง, กลยุทธ์การประเมิน, ความหลากหลายทางวัฒนธรรม

Abstract

Multicultural competence is a multidimensional construct encompassing awareness, knowledge, and skills that reflect holistic learning in culturally diverse contexts. However, assessing this competence still faces several limitations, such as unclear conceptual components, the interchangeable use of “multicultural competence” and “cross-cultural competence” without clearly distinguishing their meanings, and the bias in self-report methods that may not accurately reflect actual behavior. Effective assessment thus requires a well-defined conceptual framework and empirical evidence to support valid interpretation. This paper adopts Messick’s unified concept of validity, which emphasizes the consideration of multiple aspects-content, substantive, structural, generalization, relations with other variables, and the consequences of score use-alongside Kane’s argument-based validation framework, which organizes inferences leading to decisions based on assessment results and systematically gathers evidence to support each step of the reasoning chain. Appropriate assessment strategies should be diverse and integrative, incorporating various methods such as self-report, situational judgment test, behavioral observation, measurement invariance analysis across cultural groups, and stakeholder participation. These strategies aim to ensure accurate, reliable, and culturally responsive evaluations. This article proposes a comprehensive assessment approach that reflects the complexity of multicultural competence and supports the development of valid, practical, and applicable tools aligned with global competence goals in the 21st century.

Keywords: Multicultural Competence, Validity, Assessment Strategies, Diverse Cultural Contexts

บทนำ

ในยุคโลกาภิวัตน์ สมรรถนะทางพหุวัฒนธรรม (Multicultural Competence) ได้กลายเป็นองค์ประกอบสำคัญของสมรรถนะระดับโลก (Global Competence) ซึ่งจำเป็นต่อการดำเนินชีวิตและการทำงานในสังคมที่มีความหลากหลายทางวัฒนธรรม โดยเฉพาะในบริบททางการศึกษาที่เปลี่ยนแปลงอย่างรวดเร็ว ทั้งจากการเคลื่อนย้ายของผู้คน ความก้าวหน้าทางเทคโนโลยี และความหลากหลายของผู้เรียนในห้องเรียน (PISA, 2018); (UNESCO, 2021) สมรรถนะนี้เน้นการมีปฏิสัมพันธ์กับผู้คนจากภูมิหลังทางวัฒนธรรมที่แตกต่างอย่างเข้าใจและเคารพ (OECD, 2020) ซึ่งสอดคล้องกับเป้าหมายของการศึกษาเพื่อการอยู่ร่วมกันอย่างสันติในสังคมพหุวัฒนธรรม อย่างไรก็ตาม การส่งเสริมสมรรถนะทางพหุวัฒนธรรมในระดับนานาชาติยังเผชิญข้อจำกัดหลายประการ ไม่ว่าจะเป็นความไม่สอดคล้องกันขององค์ประกอบที่ใช้ในการศึกษา (Deardorff, D. K., 2006) เครื่องมือวัดที่ยังไม่ตอบโจทย์มาตรฐานร่วมสมัย ตลอดจนความคลุมเครือเชิงแนวคิดระหว่างคำว่า “สมรรถนะทางพหุวัฒนธรรม” กับ “สมรรถนะข้ามวัฒนธรรม” ซึ่งมักถูกใช้แทนกันในงานวิจัย แม้ทั้งสองแนวคิดจะมีความแตกต่างกัน ส่งผลให้เกิดความคลาดเคลื่อนในระดับนิยาม องค์ประกอบ และกระบวนการประเมิน ยิ่งไปกว่านั้น การประเมินองค์ประกอบเชิงนามธรรม เช่น การตระหนักรู้ในตนเองเกี่ยวกับสมมติฐาน ค่านิยม และอคติทางวัฒนธรรม มักอาศัยการรายงานตนเอง (Self-report) ซึ่งมีแนวโน้มเกิดอคติในการตอบ

เช่น การให้คำตอบในแบบที่สังคมคาดหวัง หรือการสร้างภาพลักษณ์เชิงบวกเกินจริง (Griffith, R. L. et al., 2016) ภายใต้ข้อจำกัดเหล่านี้ ความตรง (Validity) ของเครื่องมือจึงกลายเป็นประเด็นสำคัญ โดยเฉพาะเมื่อต้องประเมินแนวคิดที่มีหลายมิติและมีลักษณะนามธรรมสูง เครื่องมือจึงควรตั้งอยู่บนกรอบแนวคิดที่ชัดเจน มีหลักฐานเชิงทฤษฎี และเชิงประจักษ์รองรับ สามารถสะท้อนองค์ประกอบของแนวคิดได้อย่างครบถ้วน และให้ผลการวัดที่ตรง เชื่อถือได้ และเป็นธรรมต่อผู้เรียนจากบริบททางวัฒนธรรมที่หลากหลาย

จากสภาพปัญหาและข้อจำกัดที่กล่าวมา บทความนี้มุ่งวิเคราะห์แนวคิด องค์ประกอบ และกลยุทธ์การประเมินสมรรถนะทางพหุวัฒนธรรม โดยให้ความสำคัญกับการตรวจสอบความตรงของเครื่องมือประเมินในมิติต่าง ๆ เพื่อสนับสนุนการพัฒนาเครื่องมือที่มีความเหมาะสม เชื่อถือได้ และสอดคล้องกับบริบทวัฒนธรรมที่หลากหลาย ตลอดจนเชื่อมโยงกับเป้าหมายของสมรรถนะระดับโลกในศตวรรษที่ 21

ความหมายของสมรรถนะทางพหุวัฒนธรรม

สมรรถนะที่เกี่ยวข้องกับวัฒนธรรม มักถูกใช้เพื่ออธิบายความสามารถของบุคคลในการอยู่ร่วมและปฏิบัติงานร่วมกับผู้อื่นในสังคมที่มีความหลากหลายทางวัฒนธรรม อย่างไรก็ตาม ในงานวิจัยระดับนานาชาติ มักพบการใช้คำว่า “สมรรถนะทางพหุวัฒนธรรม” (Multicultural Competence) และ “สมรรถนะข้ามวัฒนธรรม” (Intercultural Competence) แทนกัน โดยไม่มีการแยกแยะความแตกต่างอย่างชัดเจน ทั้งที่แนวคิดทั้งสองนี้มีความต่างกันในด้านนิยาม จุดเน้น และการนำไปใช้ในทางปฏิบัติ การทำความเข้าใจความเหมือนและความต่างระหว่างสองแนวคิดจึงมีความสำคัญ ทั้งในเชิงทฤษฎีและการประยุกต์ใช้ โดยเฉพาะเมื่อนำไปพัฒนาเครื่องมือประเมินที่มีวัตถุประสงค์เฉพาะเจาะจง เพื่อให้การวัดมีความชัดเจน ไม่คลุมเครือ และสามารถตีความผลลัพธ์ได้อย่างถูกต้องตามแนวคิดที่ต้องการ

ในบริบทของโลกยุคปัจจุบัน ซึ่งเต็มไปด้วยความหลากหลายทางเชื้อชาติ ศาสนา ภาษา วัฒนธรรม และวิถีชีวิต แนวคิดเรื่อง “สมรรถนะทางพหุวัฒนธรรม” จึงได้รับความสนใจอย่างแพร่หลายในหลากหลายสาขา โดยเฉพาะวิชาชีพที่เกี่ยวข้องกับการให้บริการมนุษย์ เช่น การให้คำปรึกษา การพยาบาล งานกิจการนักศึกษา การเรียนการสอน การบริการชุมชน ผลงานของ Sue, D. W. et al., Pope, R. L. & Reynolds, A. L. เป็นรากฐานทางวิชาการที่ได้รับการยอมรับและยังคงถูกอ้างอิงอย่างต่อเนื่อง เพื่อใช้ในการพัฒนาแนวคิดและเครื่องมือวัดให้สอดคล้องกับบริบทในยุคโลกาภิวัตน์ จากมุมมองของงานให้คำปรึกษา (Sue, D. W. et al., 1992); (Pope, R. L. & Reynolds, A. L., 1997) ทั้งนี้ Sue, D. W. et al. ได้นิยามสมรรถนะทางพหุวัฒนธรรมว่า เป็นความสามารถของผู้ให้คำปรึกษาในด้านความตระหนักรู้ (Awareness) ความรู้ (Knowledge) และ ทักษะ (Skills) ที่จำเป็นต่อการทำงานในบริบทที่มีความหลากหลายทางวัฒนธรรม (Sue, D. W. et al., 1992) ในทำนองเดียวกัน Pope, R. L. & Reynolds, A. L. ก็เน้นองค์ประกอบทั้งสามนี้ แต่เพิ่มเติมมิติของ “จริยธรรม” ในการปฏิบัติงานเข้าไปด้วย (Pope, R. L. & Reynolds, A. L., 1997) ต่อมาในปี 2019 Pope, R. L. et al. ได้ขยายความหมายของแนวคิดนี้ให้ครอบคลุมมากยิ่งขึ้น โดยชี้ให้เห็นว่าสมรรถนะทางพหุวัฒนธรรม หมายถึง ความสามารถในการมีปฏิสัมพันธ์กับบุคคลจากวัฒนธรรมที่แตกต่างกันอย่างมีประสิทธิภาพ โดยต้องอาศัยการตระหนักรู้ในโลกทัศน์ของตนเอง ทักษะต่อความหลากหลาย ความรู้เกี่ยวกับวัฒนธรรมอื่น และทักษะในการสื่อสารข้ามวัฒนธรรม ทั้งนี้ยังเน้นมิติของการปลูกฝังทัศนคติและพฤติกรรมเพื่อ

ส่งเสริมความยุติธรรมทางสังคมในหมู่นักศึกษา ซึ่งแตกต่างจากแนวคิดดั้งเดิมที่เน้นการพัฒนาบุคลากรระดับปฏิบัติการ เช่น เจ้าหน้าที่และอาจารย์ (Pope, R. L. et al., 2019)

กล่าวโดยสรุป สมรรถนะทางพหุวัฒนธรรม คือ ความสามารถของบุคคลในการรับรู้ เข้าใจ และแสดงพฤติกรรมที่เหมาะสมต่อความหลากหลายทางวัฒนธรรม โดยมีองค์ประกอบสำคัญ คือ ความตระหนักรู้ ความรู้ และทักษะ ซึ่งล้วนเป็นพื้นฐานสำหรับการอยู่ร่วมกันในสังคมพหุวัฒนธรรมอย่างมีความเคารพและเท่าเทียม ทั้งยังเป็นรากฐานสำคัญในการปฏิบัติงานในบริบทการศึกษานานาชาติ การศึกษาวิวัฒนาการของแนวคิดนี้จึงมีประโยชน์ต่อการสังเคราะห์กรอบแนวคิด การออกแบบแบบวัด และการตรวจสอบความตรงเพื่อให้เครื่องมือประเมินที่พัฒนาขึ้นมีความเหมาะสมทั้งในเชิงทฤษฎีและการใช้งาน

ความจำเป็นของการมีสมรรถนะทางพหุวัฒนธรรม

ในยุคโลกาภิวัตน์ที่ความหลากหลายทางวัฒนธรรมขยายตัวและปรากฏชัดในทุกมิติของสังคม สมรรถนะทางพหุวัฒนธรรมจึงกลายเป็นทักษะสำคัญที่บุคคลในหลายบริบทควรมี โดยเฉพาะในภาคการศึกษาและงานบริการมนุษย์ สมรรถนะนี้ไม่เพียงช่วยให้บุคคลเข้าใจความแตกต่างทางวัฒนธรรมเท่านั้น แต่ยังมีบทบาทสำคัญต่อการสร้างสังคมที่เปิดกว้าง เท่าเทียม และอยู่ร่วมกันอย่างสงบสุข บทความนี้ได้นำเสนอประเด็นสำคัญ 5 ประการที่สะท้อนถึงความจำเป็นของสมรรถนะทางพหุวัฒนธรรม ดังนี้

1. การทำงานร่วมกับผู้คนจากหลากหลายวัฒนธรรมอย่างมีประสิทธิภาพ ในบริบทที่มีความหลากหลายทางวัฒนธรรม บุคลากรที่ให้บริการจำเป็นต้องพัฒนาสมรรถนะทางพหุวัฒนธรรม ซึ่งประกอบด้วย การตระหนักรู้ในอคติและสมมติฐานส่วนตัว ความเข้าใจในค่านิยมและบริบทของวัฒนธรรมต่าง ๆ รวมถึงทักษะในการสื่อสารและปฏิบัติงานร่วมกับผู้คนจากวัฒนธรรมที่หลากหลาย สมรรถนะเหล่านี้มีความสำคัญต่อการให้บริการที่มีความเหมาะสมและตอบสนองต่อความต้องการของผู้รับบริการได้อย่างมีประสิทธิภาพ (Sue, D. W. et al., 1992)

2. การรองรับความเปลี่ยนแปลงของสังคมโลกในระบบการศึกษา ด้วยความหลากหลายทางวัฒนธรรมที่เพิ่มขึ้นอย่างรวดเร็วในสังคมโลก ระบบการศึกษาจึงจำเป็นต้องพัฒนาสมรรถนะทางพหุวัฒนธรรมในครูและนักเรียน เพื่อเตรียมความพร้อมรับมือกับความหลากหลายดังกล่าว พร้อมทั้งส่งเสริมการเรียนรู้ที่เป็นธรรม มีคุณภาพ และเปิดกว้างต่อความแตกต่างทางวัฒนธรรม (Banks, J. A., 2014)

3. การเพิ่มประสิทธิภาพในการสื่อสารข้ามวัฒนธรรม สมรรถนะทางพหุวัฒนธรรมช่วยให้บุคคลสามารถสื่อสารกับผู้คนจากวัฒนธรรมต่าง ๆ ได้อย่างเหมาะสม ลดความเข้าใจผิด และสร้างความสัมพันธ์ที่ดีในบริบทที่มีความหลากหลายทางวัฒนธรรม (Ting-Toomey, S., 1999)

4. การสร้างสภาพแวดล้อมการเรียนรู้ที่เปิดกว้างและเท่าเทียม ครูที่มีสมรรถนะทางพหุวัฒนธรรมสามารถออกแบบการเรียนรู้ที่ครอบคลุมมิติทางวัฒนธรรมหลายด้าน ส่งเสริมความเท่าเทียม และลดความเหลื่อมล้ำในชั้นเรียนอย่างมีประสิทธิภาพ (Banks, J. A., 2014)

5. การส่งเสริมความเข้าใจและความเคารพในความแตกต่างทางวัฒนธรรมเพื่อสังคมที่ยั่งยืน สมรรถนะทางพหุวัฒนธรรมมีบทบาทสำคัญในการสร้างความเข้าใจและความเคารพในความแตกต่างระหว่างวัฒนธรรม ซึ่งเป็นพื้นฐานสำคัญในการส่งเสริมสังคมที่สงบสุขและยั่งยืน ตามแนวทางของการศึกษาสำหรับการพัฒนาที่ยั่งยืน



(Education for Sustainable Development) ที่เน้นการเรียนรู้เพื่อเสริมสร้างความเคารพในความหลากหลาย และการอยู่ร่วมกันอย่างสันติในสังคม (UNESCO, 2021)

สมรรถนะทางพหุวัฒนธรรมเป็นพื้นฐานสำคัญสำหรับการเรียนรู้ การทำงาน และการอยู่ร่วมกันในสังคมที่หลากหลาย บุคคลที่มีสมรรถนะนี้จะสามารถสื่อสารและปฏิบัติตนได้อย่างเหมาะสมในบริบทพหุวัฒนธรรม การประเมินสมรรถนะจึงควรมีความตรง ครบคลุม และเชื่อมโยงกับบริบท เพื่อใช้ในการส่งเสริมและพัฒนาผู้เรียนได้อย่างมีประสิทธิภาพ ทั้งนี้ การเข้าใจองค์ประกอบของสมรรถนะเป็นก้าวแรกสู่การออกแบบและประเมินอย่างมีประสิทธิภาพ

องค์ประกอบของสมรรถนะทางพหุวัฒนธรรม

สมรรถนะทางพหุวัฒนธรรมไม่ได้หมายถึงเพียงความรู้ความเข้าใจเชิงทฤษฎีเท่านั้น หากแต่เป็นคุณลักษณะที่สามารถพัฒนาได้ และมีความจำเป็นอย่างยิ่งต่อการปฏิบัติงานในสภาพแวดล้อมที่มีความหลากหลายทางวัฒนธรรม การทำความเข้าใจองค์ประกอบของสมรรถนะทางพหุวัฒนธรรมจึงเป็นสิ่งสำคัญ โดยทั่วไปกรอบแนวคิดพื้นฐาน จะประกอบด้วย 3 องค์ประกอบหลัก ได้แก่ ความตระหนักรู้ ความรู้ และทักษะ อย่างไรก็ตาม นักวิชาการบางท่านได้ขยายหรือปรับกรอบแนวคิดดังกล่าว เพื่อเพิ่มมิติทางความคิดและความเข้าใจที่ลึกซึ้งยิ่งขึ้น

Sue, D. W. et al. เสนอโมเดลพื้นฐานของสมรรถนะทางพหุวัฒนธรรม (Multicultural Competence) ประกอบด้วย 3 องค์ประกอบหลัก ได้แก่ ความตระหนักรู้ (Awareness) ความรู้ (Knowledge) และทักษะ (Skills) ซึ่งถือเป็นแนวคิดคลาสสิกที่ยังถูกศึกษาและพัฒนาอย่างต่อเนื่องในปัจจุบัน (Sue, D. W. et al., 1992) ขณะที่ Pope, R. L. & Reynolds, A. L. ได้ใช้กรอบแนวคิดเดียวกัน ในการอธิบายสมรรถนะทางพหุวัฒนธรรมสำหรับบุคลากรงานกิจการนักศึกษา โดยมุ่งเน้นการพัฒนา 3 องค์ประกอบข้างต้นเพื่อให้สามารถปฏิบัติงานอย่างมีประสิทธิภาพและมีจริยธรรมในบริบททางวัฒนธรรมที่หลากหลาย (Pope, R. L. & Reynolds, A. L., 1997)

Campinha-Bacote, J. ได้พัฒนาและขยายกรอบแนวคิดจากโครงสร้างพื้นฐานเดิม “awareness-knowledge-skills” ซึ่งได้รับความนิยมในวงการจิตวิทยาและการให้คำปรึกษา โดยเฉพาะจากโมเดลของ Sue, D. W. et al. ปี 1992 อย่างไรก็ตาม Campinha-Bacote, J. มองว่า กรอบแนวคิดเดิมนั้นเน้นการแยกองค์ประกอบออกจากกันเป็นส่วน ๆ โดยไม่สะท้อนถึงลักษณะของสมรรถนะทางพหุวัฒนธรรมในฐานะ กระบวนการพัฒนาอย่างต่อเนื่องและองค์รวม ดังนั้น จึงได้เสนอโมเดลที่ชื่อว่า “กระบวนการของสมรรถนะทางวัฒนธรรมในการให้บริการด้านสุขภาพ” (The Process of Cultural Competence in the Delivery of Healthcare Services) ซึ่งมองว่าสมรรถนะทางวัฒนธรรมเป็นกระบวนการที่ดำเนินไปอย่างต่อเนื่อง ที่ผู้ให้บริการต้องพยายามพัฒนาตนเองอย่างสม่ำเสมอเพื่อสามารถทำงานได้อย่างเหมาะสมในบริบททางวัฒนธรรมของผู้รับบริการ โมเดลนี้ ประกอบด้วย 5 องค์ประกอบที่เชื่อมโยงกัน ได้แก่ การตระหนักรู้ทางวัฒนธรรม (Cultural Awareness) ความรู้ทางวัฒนธรรม (Cultural Knowledge) ทักษะทางวัฒนธรรม (Cultural Skill) การปฏิสัมพันธ์กับบุคคลที่มาจากวัฒนธรรมที่ต่างกัน (Cultural Encounters) และความปรารถนาในการพัฒนาสมรรถนะทางวัฒนธรรม (Cultural Desire) โมเดลใหม่นี้เพิ่มเติม 2 องค์ประกอบสำคัญ คือ การปฏิสัมพันธ์กับบุคคลที่มาจากวัฒนธรรมที่ต่างกัน และความปรารถนาในการพัฒนาสมรรถนะทางวัฒนธรรม เพื่อสะท้อนให้เห็นว่าการพัฒนา สมรรถนะทางพหุวัฒนธรรม ไม่ได้เป็น



เพียงเรื่องของความรู้หรือทักษะเท่านั้น แต่ยังเชื่อมโยงกับการมีปฏิสัมพันธ์จริงกับคนจากวัฒนธรรมต่าง ๆ และแรงจูงใจเชิงจิตใจกับการเรียนรู้และปรับตัวต่อวัฒนธรรมที่หลากหลาย (Campinha-Bacote, J., 2002)

ต่อมาปี 2019 Pope, R. L. et al. ได้พัฒนาแนวคิดเรื่องสมรรถนะทางพหุวัฒนธรรม โดยอิงจากกรอบเดิมเมื่อปี 2004 ซึ่งประกอบด้วย 3 องค์ประกอบหลัก ได้แก่ ความตระหนักรู้ ความรู้ และทักษะ โดยสมรรถนะทางพหุวัฒนธรรม หมายถึง ความสามารถในการปฏิสัมพันธ์อย่างมีประสิทธิภาพกับผู้คนจากวัฒนธรรมที่แตกต่าง ในฉบับปรับปรุงล่าสุด Pope, R. L. et al. ได้ขยายโมเดลเป็น 5 องค์ประกอบ โดยเพิ่มการลงมือปฏิบัติ (Multicultural Action) และการสนับสนุนความยุติธรรมทางสังคม (Advocacy/Social Justice) เพื่อเน้นย้ำว่า บุคลากรด้านกิจการนักศึกษาไม่เพียงแต่ต้องมีความรู้และทักษะ แต่ยังควรลงมือทำและมีบทบาทเชิงรุกในการเปลี่ยนแปลงเชิงระบบ เพื่อสร้างความครอบคลุมและความเป็นธรรมในสภาพแวดล้อมการศึกษา ขณะที่องค์ประกอบเดิมทั้งสามยังคงเป็นรากฐานสำคัญ โดยมีความหมายเฉพาะ ดังนี้ (Pope, R. L. et al., 2019)

1. ความตระหนักรู้ในโลกทัศน์ของตนเอง (Self-awareness) คือ การเข้าใจค่านิยม ความเชื่อ และอคติที่กำหนดมุมมองของตนเอง ซึ่งอาจส่งผลต่อการปฏิสัมพันธ์กับผู้คนจากวัฒนธรรมที่แตกต่าง
 2. ความรู้เกี่ยวกับวัฒนธรรมต่าง ๆ (Cultural Knowledge) คือ ความเข้าใจเกี่ยวกับผู้คน วัฒนธรรม ประวัติศาสตร์ และบริบททางสังคม เพื่อหลีกเลี่ยงการเหมารวม การตีตรา หรือการใช้สื่ออคติในการสื่อสาร
 3. ทักษะระหว่างวัฒนธรรม (Intercultural Skills) คือ ความสามารถในการนำความตระหนักรู้และความรู้ไปประยุกต์ใช้จริง เช่น การสื่อสาร การให้บริการ หรือการออกแบบกิจกรรมที่เหมาะสมกับกลุ่มเป้าหมายที่หลากหลาย
- แม้แนวคิดพื้นฐานทั้งสามนี้ยังเป็นแกนหลักของโมเดล แต่ในฉบับล่าสุด ได้มีการขยายแนวคิดเพื่อรวมแรงจูงใจ ประสบการณ์จริง และบทบาทด้านความยุติธรรมทางสังคม เพื่อให้สมรรถนะทางพหุวัฒนธรรมมีความลึกและครอบคลุมมากขึ้น ทั้งนี้ ข้อมูลเหล่านี้เป็นรากฐานสำคัญในการออกแบบ กระบวนการอนุมาน (Inference Chain) สำหรับการประเมินสมรรถนะทางพหุวัฒนธรรม ซึ่งช่วยเชื่อมโยงองค์ความรู้ ทักษะ และพฤติกรรมที่ต้องการวัดอย่างเป็นระบบ

การตรวจสอบความตรง

การตรวจสอบความตรง (Validation) เป็นกระบวนการสำคัญในการประเมินคุณภาพของเครื่องมือวัด โดยเฉพาะในกรณีที่ต้องวัดสมรรถนะหรือแนวคิดเชิงนามธรรม เช่น “สมรรถนะทางพหุวัฒนธรรม” ซึ่งไม่สามารถวัดได้โดยตรง ความตรง หมายถึง ระดับที่หลักฐานเชิงประจักษ์และทฤษฎี สนับสนุนการตีความคะแนนสอบและการนำคะแนนไปใช้ให้สอดคล้องกับวัตถุประสงค์ของการวัดในบริบทที่เฉพาะเจาะจง (American Educational Research Association (AERA), American Psychological Association (APA), And National Council on Measurement in Education (NCME), 2014)

แนวคิดดั้งเดิมเกี่ยวกับความตรงมักจำแนกเป็นประเภท ได้แก่ ความตรงเชิงเนื้อหา ความตรงเชิงเกณฑ์ และความตรงเชิงโครงสร้าง ซึ่ง Messick, S. ได้วิพากษ์แนวคิดดังกล่าวว่าเป็นการมองแบบแยกส่วน และละเลยประเด็นสำคัญเกี่ยวกับคุณค่าทางสังคมและผลกระทบจากการใช้คะแนน ความตรงมิใช่คุณลักษณะของแบบทดสอบโดยตัวมันเอง แต่เป็นคุณภาพของความหมายที่สรุปได้จากคะแนนสอบ และผลจากการนำคะแนนไปใช้ในการตัดสินใจ

Messick, S. จึงเสนอแนวคิด “ความตรงแบบองค์รวม” (Unified Concept of Validity) แนวคิดนี้ ประกอบด้วย แหล่งของหลักฐาน 6 ด้าน (Messick, S., 1995) ได้แก่

1. เนื้อหา (Content) หมายถึง หลักฐานที่แสดงถึงความครอบคลุมและสอดคล้องของข้อสอบกับโครงสร้างแนวคิดที่ต้องการวัด
2. สารสำคัญ (Substantive) คือ หลักฐานที่แสดงว่าแบบทดสอบสามารถกระตุ้นหรือสะท้อน กระบวนการคิดของผู้ตอบ ตามกรอบทฤษฎีที่รองรับคุณลักษณะที่ต้องการวัด
3. โครงสร้าง (Structural) คือ หลักฐานที่แสดงว่าโครงสร้างของคะแนนสอดคล้องกับโครงสร้างของคุณลักษณะที่ต้องการวัด ตามกรอบแนวคิดหรือทฤษฎีที่รองรับ
4. การสรุปอ้างอิง (Generalizability) คือ ความสามารถในการใช้สมบัติของคะแนน (Score Properties) และการตีความคะแนน (Score Interpretations) สรุปผลการวัดไปยังบริบทอื่น
5. ความสัมพันธ์กับตัวแปรอื่น (External) คือ หลักฐานความตรงที่แสดงถึงผลการวัดมีความสัมพันธ์ในเชิงบวกกับตัวแปรที่เกี่ยวข้อง (Convergent Validity) และไม่มีความสัมพันธ์กับตัวแปรที่ไม่เกี่ยวข้อง (Discriminant Validity) รวมถึงความสัมพันธ์กับเกณฑ์ภายนอก (Criterion-related Validity)
6. ผลกระทบจากการใช้คะแนน (Consequential) คือ การประเมินคุณค่าหรือผลกระทบจากการตีความคะแนนที่นำไปสู่การดำเนินการต่าง ๆ รวมถึง ผลลัพธ์จริงหรือที่อาจเกิดขึ้นจากการใช้แบบทดสอบ โดยเฉพาะในเรื่องของอคติ ความเป็นธรรม และความยุติธรรม

ต่อมา มาตรฐานการวัดทางการศึกษาและจิตวิทยาฉบับล่าสุด (AERA, APA & NCME, 2014) มีมุมมองว่า ความตรงเป็นกระบวนการรวบรวมหลักฐานเพื่อตรวจสอบการตีความคะแนนสอบตามวัตถุประสงค์ของการใช้งาน และได้จัดระบบแหล่งหลักฐานสำหรับการประเมินความตรงไว้ 5 ด้าน ซึ่งสามารถเชื่อมโยงกับกรอบแนวคิดของ Messick, S. ได้อย่างสอดคล้อง ได้แก่

1. เนื้อหาของแบบทดสอบ (Test Content) คือ การตรวจสอบความเกี่ยวข้องของข้อสอบกับเนื้อหาที่ต้องการวัด
2. กระบวนการตอบของผู้สอบ (Response Processes) คือ การพิจารณากลยุทธ์หรือการคิดของผู้เข้าสอบ ในการตอบคำถาม
3. โครงสร้างภายใน (Internal Structure) คือ ความสอดคล้องของโครงสร้างข้อสอบกับทฤษฎี เช่น การวิเคราะห์องค์ประกอบหรือความเที่ยง (Reliability)
4. ความสัมพันธ์กับตัวแปรอื่น (Relations to Other Variables) คือ การพิจารณาความสัมพันธ์ของคะแนน กับเกณฑ์ภายนอกหรือการศึกษาแบบหลายลักษณะหลายวิธี (Multitrait Multimethod)
5. ผลกระทบจากการใช้ผลการทดสอบ (Consequences of Testing) คือ การตรวจสอบผลกระทบทั้งที่ตั้งใจและไม่ตั้งใจจากการใช้คะแนนหรือผลลัพธ์จากการทดสอบต่อบุคคลและสังคม

นอกจากแนวคิดข้างต้น Kane, M. T. ยังเสนอแนวทางที่สอดคล้องกันเรียกว่า “การตรวจสอบความตรงด้วยการโต้แย้ง” (Argument-based Validation) ซึ่งมุ่งเน้นการสร้างกรอบเหตุผลที่เชื่อมโยงข้อมูลจากการวัดกับข้อสรุปที่ได้จากการใช้คะแนน โดยแบ่งออกเป็น 2 องค์ประกอบหลัก (Kane, M. T., 1992); (Kane, M. T., 2006) ได้แก่



1. การให้เหตุผลในการตีความ (Interpretive Argument) คือ การจัดลำดับข้ออ้าง (Claims) ที่แสดงความสัมพันธ์ระหว่างข้อมูลจากการวัดกับการตีความและการใช้คะแนน เช่น การให้คะแนน การสรุปผล และการตัดสินใจ

2. การโต้แย้งเชิงความตรง (Validity Argument) คือ การสนับสนุนคำโต้แย้งเหล่านั้นด้วยหลักฐานเชิงประจักษ์และเหตุผลทางทฤษฎีที่เหมาะสม เช่น ความเที่ยง (Reliability) ความเป็นปรนัย (Objectivity) ความคงที่ของโครงสร้าง (Structural Consistency) และความยุติธรรม (Fairness)

ภายใต้แนวคิดของ Kane, M. T. ความตรงจึงเป็นการตรวจสอบระดับความสมเหตุสมผลของการให้เหตุผลที่เชื่อมโยงกับคำโต้แย้งและการอนุมานต่าง ๆ ซึ่งรองรับการตีความและการใช้คะแนน และได้ระบุ 4 ขั้นตอนเชิงอนุมาน (Inferences) ที่เป็นแกนกลางของการสร้างและตรวจสอบกรอบการให้เหตุผลในการตีความ (Kane, M. T., 2006); (Kane, M. T., 2013) ดังนี้

1. การตีความคะแนน (Scoring Inference) คือ การให้คะแนนต้องมีความถูกต้องสอดคล้องกับการตอบของผู้สอบและเชื่อถือได้

2. การสรุปอ้างอิงไปยังชุดข้อมูลอื่น (Generalization Inference) คือ การตรวจสอบว่า คะแนนที่ได้จากแบบทดสอบสามารถแทนความสามารถของผู้สอบในเรื่องนั้นอย่างคงที่และเชื่อถือได้หรือไม่ เช่น ถ้าทำข้อสอบคล้ายกันอีกครั้ง ผู้สอบจะได้คะแนนใกล้เคียงเดิม การอนุมานนี้จึงเน้นที่ความคงที่ของคะแนน (Reliability) และความเป็นตัวแทนของชุดข้อสอบ

3. การอนุมานคะแนนไปยังคุณลักษณะที่ต้องการวัด (Extrapolation Inference) คือ การตรวจสอบว่า คะแนนที่ได้ควรสะท้อนถึงความสามารถหรือพฤติกรรมจริงในสถานการณ์จริง เช่น สมรรถนะในการทำงาน การอนุมานนี้จึงเน้นที่ความถูกต้องของการตีความว่า คะแนนนั้นสะท้อนสมรรถนะจริงได้มากน้อยเพียงใด

4. การตัดสินใจ/การอนุมานคะแนน (Decision/Implication Inference) คือ การใช้คะแนนเพื่อการตัดสินใจต่าง ๆ เช่น การคัดเลือกหรือจัดระดับควรได้รับการพิจารณาว่าเหมาะสมและเป็นธรรม

กล่าวได้ว่า กรอบความตรงแบบองค์รวมของ Messick, S. ซึ่งเสนอให้พิจารณาความตรงผ่านหลักฐาน 6 ประเภท ได้แก่ เนื้อหา สำคัญ โครงสร้าง การสรุปอ้างอิง ความสัมพันธ์กับตัวแปรอื่น และผลกระทบจากการใช้คะแนน เป็นรากฐานสำคัญของแนวทางการประเมินความตรงตามข้อกำหนดของมาตรฐานการวัดทางการศึกษาและจิตวิทยา ที่จัดระบบหลักฐานความตรงไว้ 5 ด้าน ซึ่งสอดคล้องและสามารถจัดวางสัมพันธ์กับ 4 ขั้นตอนของกรอบการอนุมานของ Kane, M. T. ได้อย่างเหมาะสม ดังนี้

1. เนื้อหาและสำคัญ สนับสนุนการตีความคะแนน โดยช่วยยืนยันว่าข้อสอบครอบคลุมและสะท้อนกระบวนการตอบของผู้สอบตามทฤษฎีอย่างถูกต้อง

2. กระบวนการตอบของผู้สอบ ซึ่งสอดคล้องกับประเด็นสำคัญของ Messick, S. สนับสนุนการตรวจสอบว่าวิธีการตอบของผู้สอบสอดคล้องกับการตีความคะแนนอย่างมีเหตุผล

3. โครงสร้างภายใน และ ความสัมพันธ์กับตัวแปรอื่น สนับสนุนการสรุปอ้างอิง และการอนุมานคะแนนไปยังคุณลักษณะที่ต้องการวัด โดยยืนยันว่าโครงสร้างคะแนนและความสัมพันธ์กับตัวแปรอื่น ๆ สอดคล้องกับคุณลักษณะที่วัด

4. ผลกระทบจากการใช้ผลการทดสอบ สนับสนุนการตัดสินใจและการอนุมานคะแนนโดยพิจารณาผลลัพธ์และผลกระทบจากการนำคะแนนไปใช้ รวมทั้งความยุติธรรมและอคติที่อาจเกิดขึ้น

ดังนั้น การบูรณาการแนวคิดของ Messick, S., Kane, M. T. เป็ดมมมมมมที่ครอบคลุมทั้งในเชิงทฤษฎีและเชิงปฏิบัติ (Messick, S., 1995); (Kane, M. T., 1992) สอดคล้องกับมาตรฐานการวัดทางการศึกษาและจิตวิทยา ซึ่งช่วยให้สามารถพิจารณาความตรงของการประเมินได้อย่างรอบด้าน มีเหตุผล และเหมาะสมต่อการใช้ในบริบทที่หลากหลาย สำหรับการประเมินสมรรถนะทางพหุวัฒนธรรมในบริบทที่มีความหลากหลายทางวัฒนธรรม การตรวจสอบความตรงจึงควรดำเนินการอย่างครบถ้วน ไม่จำกัดเพียงการวิเคราะห์เชิงสถิติของข้อสอบเท่านั้น แต่ต้องพิจารณาถึงกระบวนการตอบของผู้เข้าสอบจากวัฒนธรรมที่แตกต่างกันอย่างละเอียด รวมถึงการตรวจสอบความไม่แปรเปลี่ยนของการวัด (Measurement Invariance) และการประเมินความยุติธรรมของข้อสอบผ่านการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning: DIF) เพื่อให้มั่นใจว่าการวัดสะท้อนสมรรถนะที่แท้จริงโดยปราศจากอคติทางเชื้อชาติ ภาษา หรือวัฒนธรรม ซึ่งเป็นเงื่อนไขสำคัญของความตรงที่รองรับการใช้แบบวัดอย่างเป็นธรรม

กลยุทธ์การประเมินสมรรถนะทางพหุวัฒนธรรมในบริบทที่มีความหลากหลายทางวัฒนธรรม

ในยุคที่การศึกษามีความหลากหลายทางวัฒนธรรมมากขึ้น การพัฒนาสมรรถนะทางพหุวัฒนธรรมจึงเป็นเป้าหมายสำคัญของการเรียนรู้ระดับโลก อย่างไรก็ตาม การประเมินสมรรถนะดังกล่าวยังเผชิญข้อจำกัดหลายประการ ทั้งในแง่ความไม่ชัดเจนของกรอบแนวคิด และความยากในการแปลงคุณลักษณะเชิงนามธรรม เช่น การตระหนักรู้ตนเองและทางวัฒนธรรม ให้เป็นตัวชี้วัดที่สังเกตหรือวัดได้ โดยเฉพาะเมื่อยังพึ่งพาวิธีการเชิงปริมาณหรือแบบรายงานตนเอง ซึ่งอาจสะท้อนสมรรถนะได้เพียงบางมิติ และมีความเสี่ยงต่ออคติของผู้ตอบ (Griffith, R. L. et al., 2016); (ASCD, 2024) นอกจากนี้ เครื่องมือเหล่านี้ยังอาจขาดความเป็นธรรมในการวัดระหว่างกลุ่มชาติพันธุ์หรือวัฒนธรรม หากไม่มีการตรวจสอบด้านการวิเคราะห์ความไม่แปรเปลี่ยนของการวัด และการทำหน้าที่ต่างกันของข้อสอบ (ASCD, 2024) ดังนั้น การออกแบบกลยุทธ์การประเมินที่มีประสิทธิภาพจึงควรยึดหลักวิชาการที่ชัดเจน และสอดคล้องกับบริบททางวัฒนธรรม เพื่อให้การประเมินมีความตรง เทียบ เป็นธรรม และสามารถนำไปใช้ได้จริงในสังคมที่มีความหลากหลายทางวัฒนธรรม

กลยุทธ์การประเมินสมรรถนะทางพหุวัฒนธรรม ดำเนินการ ดังนี้

1. กำหนดกรอบแนวคิดเพื่อการประเมิน การประเมินสมรรถนะทางพหุวัฒนธรรมจำเป็นต้องเริ่มต้นจากกรอบแนวคิดที่ชัดเจน โดยหนึ่งในโมเดลที่ได้รับความนิยมและมีการนำไปใช้ต่อยอดในการวิจัย คือ โมเดลของ Sue, D. W. et al. ปี 1992 ซึ่งประกอบด้วย องค์ประกอบหลัก 3 ด้าน ได้แก่ ความตระหนักรู้ทางวัฒนธรรม (Cultural Awareness) ความรู้เกี่ยวกับวัฒนธรรมอื่น (Cultural Knowledge) และทักษะในการสื่อสารและมีปฏิสัมพันธ์ระหว่างวัฒนธรรม (Cultural Skills) ขณะที่ Deardorff, D. K. ได้พัฒนาโครงสร้างพหุมิติของ “สมรรถนะข้ามวัฒนธรรม” (Intercultural Competence) ซึ่งมีแนวคิดพื้นฐานและลำดับขั้นแตกต่างกัน (Deardorff, D. K., 2006) อย่างไรก็ตาม การนำกรอบแนวคิดของ “สมรรถนะข้ามวัฒนธรรม” มาใช้แทน “สมรรถนะทางพหุวัฒนธรรม” โดยไม่แยกแยะความแตกต่าง อาจทำให้การประเมินขาดความตรงเชิงโครงสร้างได้ อีกทั้ง แม้จะมุ่งเน้นศึกษา



สมรรถนะทางพหุวัฒนธรรมโดยเฉพาะ แต่ก็มีการวิจัยเสนอกรอบแนวคิดที่แตกต่างกัน เช่น Sue, D. W. et al., Pope, R. L. & Reynolds, A. L. ซึ่งแต่ละแนวทางให้ความสำคัญกับองค์ประกอบเฉพาะที่สะท้อนมุมมองต่างบริบท (Sue, D. W. et al., 1992); (Pope, R. L. & Reynolds, A. L., 1997)

2. กำหนดรูปแบบและวิธีการประเมินแบบผสมผสาน สมรรถนะทางพหุวัฒนธรรมเป็นแนวคิดที่ซับซ้อนครอบคลุมทั้งพฤติกรรมที่สามารถสังเกตได้และกระบวนการคิดภายในที่จับต้องได้ยาก การประเมินจึงควรใช้แนวทางแบบผสมผสาน (Mixed Methods) เพื่อให้สามารถเข้าใจผู้เรียนได้อย่างรอบด้าน งานวิจัยหลายเรื่อง เช่น Lie, D. A. Et Al., Griffith, R. L. et al. สนับสนุนการใช้ทั้งเครื่องมือเชิงปริมาณ (เช่น แบบสอบถามและการประเมินตามสถานการณ์) และเครื่องมือเชิงคุณภาพ (เช่น การเขียนสะท้อนตนเองและการสังเกตพฤติกรรมในสถานการณ์จริง) วิธีการเหล่านี้ช่วยให้สามารถประเมินองค์ประกอบสำคัญของสมรรถนะ ได้แก่ การรับรู้ ทัศนคติ การให้เหตุผลเชิงวัฒนธรรม และพฤติกรรมที่แสดงออก ซึ่งไม่สามารถประเมินได้อย่างครบถ้วนด้วยวิธีใดวิธีหนึ่งเพียงลำพัง (Lie, D. A. et al., 2011); (Griffith, R. L. et al., 2016) งานล่าสุดของ Luo, J. & Chan, C. K. Y., Wilcox, M. M. et al. ยังตอกย้ำความสำคัญของเครื่องมือเหล่านี้ โดยเฉพาะในบริบทที่มีความหลากหลายทางวัฒนธรรมสูง ดังนั้นการออกแบบการประเมินสมรรถนะทางพหุวัฒนธรรมควรพิจารณาอย่างรอบด้าน ทั้งในด้านวัตถุประสงค์ของการวัด ความตรง ความเชื่อถือได้ ความเป็นธรรม และความเหมาะสมกับบริบทของผู้เรียน เพื่อให้สามารถสะท้อนสมรรถนะได้อย่างแม่นยำทั้งในเชิงวิชาการและเชิงปฏิบัติ (Luo, J. & Chan, C. K. Y., 2022); (Wilcox, M. M. et al., 2020)

3. การตรวจสอบความตรงของเครื่องมือประเมิน การออกแบบและใช้เครื่องมือประเมินสมรรถนะพหุวัฒนธรรมจำเป็นต้องพิจารณาความตรงอย่างรอบด้าน Messick, S. มองว่าความตรงเป็นแนวคิดองค์รวมที่ครอบคลุมโครงสร้างของเครื่องมือ ความสอดคล้องกับทฤษฎี และการแปลผลอย่างถูกต้อง (Messick, S., 1989) ในขณะที่ Kane, M. T. เสนอกรอบการตรวจสอบความตรงด้วยการโต้แย้ง โดยเน้นการตั้งข้อสันนิษฐาน (Inference) และการจัดหาหลักฐานสนับสนุนตลอดกระบวนการ ตั้งแต่การเก็บข้อมูล การตีความผล ไปจนถึงการนำไปใช้ในการตัดสินใจ (Kane, M. T., 1992); (Kane, M. T., 2013) หากไม่มีหลักฐานรองรับอย่างเพียงพอ การประเมินอาจนำไปสู่ข้อสรุปที่คลาดเคลื่อนและผลกระทบที่ไม่เป็นธรรม โดยเฉพาะในบริบทที่มีความหลากหลายทางวัฒนธรรม

4. การประกันความเป็นธรรมและความเท่าเทียม เมื่อเครื่องมือถูกนำไปใช้ในบริบทนานาชาติ ต้องตรวจสอบความเป็นธรรมในการประเมิน (Assessment Fairness) โดยพิจารณาว่าแบบสอบถามหรือกิจกรรมไม่ก่อให้เกิดความลำเอียงทางวัฒนธรรม (Cultural Bias) เช่น การใช้สำนวน ภาษาท้องถิ่น หรือค่านิยมที่เฉพาะเจาะจงกับบางวัฒนธรรม (Van de Vijver, F. & Leung, K., 1997) กระบวนการปรับเครื่องมือควรใช้การแปลแบบกลับ (Back-translation) และการปรับทางวัฒนธรรม (Cultural Adaptation) เพื่อลดความเหลื่อมล้ำในการทำ ความเข้าใจโจทย์และคำถาม นอกจากนี้ ควรมีการวิเคราะห์ความไม่แปรเปลี่ยนของการวัด และการทำหน้าที่ต่างกันของข้อสอบเพื่อยืนยันว่าเครื่องมือไม่ลำเอียงต่อกลุ่มใดกลุ่มหนึ่ง (Zumbo, B. D., 1999)

5. การประเมินเพื่อส่งเสริมการเรียนรู้ แนวทางสำคัญในการประเมินสมรรถนะทางพหุวัฒนธรรม คือ การใช้เพื่อพัฒนาหรือการประเมินเพื่อการเรียนรู้ (Assessment for Learning) มากกว่าการคัดกรองหรือการประเมินเพื่อตัดสิน (Assessment of Learning) การให้ข้อมูลป้อนกลับอย่างต่อเนื่อง การตั้งเป้าหมายการพัฒนาตน

(Developmental Goals) และการสะท้อนตนเอง (Self-reflection) จะช่วยให้ผู้เรียนเข้าใจและปรับปรุงทักษะของตนเองอย่างเป็นระบบ (ASCD, 2024)

ข้อเสนอแนะในการพัฒนาและประเมินเครื่องมือ

1. เริ่มต้นด้วยกรอบแนวคิดที่ชัดเจน คือ นิยามองค์ประกอบสมรรถนะทางพหุวัฒนธรรมที่ต้องการวัดอย่างละเอียด เชื่อมโยงองค์ประกอบกับตัวชี้วัดเชิงพฤติกรรม
2. เลือกใช้วิธีประเมินแบบผสม โดยผสมผสานเครื่องมือ เช่น แบบรายงานตนเองกับการประเมินพฤติกรรมจริงและสะท้อนตน หรืออาจเพิ่มความน่าเชื่อถือด้วยการใช้หลายวิธีหรือหลายแหล่งข้อมูล (Triangulation)
3. สร้างและตรวจสอบหลักฐานความตรงหลายมิติ เช่น ทบทวนว่าตัวชี้วัดครอบคลุมสาระสำคัญที่ต้องการวัด (Substantive Validity) ศึกษากระบวนการตอบของผู้สอบ (Response Process Validity) อาทิ ผ่านการสัมภาษณ์หรือ Think-aloud Protocol วิเคราะห์โครงสร้างของเครื่องมือ (Internal Structure Validity) ตรวจสอบผลกระทบที่เกิดจากการใช้เครื่องมือ (Consequential Validity) อาทิ ความเป็นธรรม ความเท่าเทียม
4. ตรวจสอบความเป็นธรรมทางวัฒนธรรม (Fairness) เช่น ใช้การวิเคราะห์ความไม่แปรเปลี่ยนของการวัด การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ เพื่อยืนยันว่าไม่มีข้อสอบหรือตัวชี้วัดใดลำเอียงกับกลุ่มวัฒนธรรมใด ใช้การแปลและปรับวัฒนธรรม (Back-translation & Cultural Adaptation)
5. จัดทำคู่มือการประเมินและอบรมผู้ประเมิน เพื่อให้ทุกฝ่ายเข้าใจวิธีการใช้เครื่องมือและเกณฑ์การประเมินอย่างเหมาะสม
6. สรุปและรายงานผลอย่างโปร่งใส จัดทำรายงานหลักฐานความตรงทั้งหมดและข้อจำกัดของเครื่องมือเสนอแนวทางพัฒนาต่อไปสำหรับการใช้งานในบริบทต่าง ๆ

ข้อเสนอเชิงสังเคราะห์และกลยุทธ์ประเมินสมรรถนะทางพหุวัฒนธรรม

จากการศึกษาวรรณกรรมอย่างเป็นระบบเกี่ยวกับแนวคิดสมรรถนะทางพหุวัฒนธรรม และแนวทางการพัฒนาเครื่องมือประเมินที่มีความตรงบริบทการศึกษานานาชาติ พบว่า สมรรถนะทางพหุวัฒนธรรมเป็นแนวคิดที่มีมิติสลับซับซ้อน ประกอบด้วย ความตระหนักรู้ ความรู้ และทักษะ ซึ่งสะท้อนลักษณะการเรียนรู้แบบองค์รวม (Holistic Learning) ที่ครอบคลุมทั้งความเข้าใจ ความรู้สึกนึกคิด และความสามารถในการประยุกต์ใช้ในการสื่อสารข้ามวัฒนธรรมอย่างเหมาะสมตามบริบท อย่างไรก็ตาม การประเมินสมรรถนะทางพหุวัฒนธรรมในเชิงลึกจำเป็นต้องมีกรอบแนวคิดที่ชัดเจนเกี่ยวกับนิยามและองค์ประกอบ ตลอดจนแนวทางการอนุมานจากผลการวัดที่สามารถตรวจสอบและให้เหตุผลได้ (Valid Inferences) เพื่อหลีกเลี่ยงความเสี่ยงจากการใช้เครื่องมือที่เบี่ยงเบนจากแนวคิดเดิม ดังนั้น การพัฒนาเครื่องมือประเมินควรให้ความสำคัญกับความตรงเชิงโครงสร้าง โดยต้องสะท้อนลักษณะพหุมิติของแนวคิด และสอดคล้องกับบริบททางวัฒนธรรมของผู้เรียนที่หลากหลาย เพื่อให้ผลการประเมินสามารถนำไปใช้ในการออกแบบหลักสูตรและนโยบายทางการศึกษาได้อย่างมีประสิทธิภาพ สอดคล้องกับเป้าหมายของการพัฒนาสมรรถนะระดับโลก (Global Competence) และสามารถใช้งานได้จริงในทางปฏิบัติ โดยมีหลักฐานเชิงประจักษ์รองรับ



แนวทางการพัฒนาเครื่องมือประเมิน

เพื่อให้เครื่องมือประเมินสมรรถนะทางพฤติกรรมมีความน่าเชื่อถือและสามารถใช้งานได้ในบริบทสากล ควรพิจารณาแนวทาง ดังต่อไปนี้:

1. พัฒนาเครื่องมือประเมินโดยใช้กรอบแนวคิดความตรงแบบองค์รวม Messick, S. เสนอว่าความตรงของเครื่องมือวัดควรถูกพิจารณาจากหลากหลายมิติ ได้แก่ เนื้อหา สาระสำคัญที่สะท้อนกรอบแนวคิดของผู้สอบ โครงสร้าง การสรุปอ้างอิง ความสัมพันธ์กับตัวแปรอื่น และผลกระทบที่เกิดจากการตีความหรือการใช้คะแนน โดยเฉพาะ “ความเป็นธรรม” (Fairness) ซึ่งถือเป็นประเด็นสำคัญในมิติของผลกระทบที่เกิดจากการใช้คะแนน (Messick, S., 1995) ข้อเสนอแนะสำหรับการออกแบบและพัฒนาเครื่องมือประเมิน โดยอ้างอิงจากกรอบแนวคิดความตรงแบบองค์รวม ดังแสดงในตารางที่ 1

ตารางที่ 1 ข้อเสนอแนะสำหรับการออกแบบและพัฒนาเครื่องมือประเมิน โดยอ้างอิงจากกรอบแนวคิดความตรงแบบองค์รวม

ด้านความตรงของ Messick	คำอธิบาย	ข้อเสนอแนะเพื่อการออกแบบและพัฒนาเครื่องมือ
1. ความตรงเชิงเนื้อหา (Content Validity)	เครื่องมือวัดครอบคลุมเนื้อหาหรือองค์ประกอบของสมรรถนะตามที่ระบุไว้ในกรอบแนวคิดหรือทฤษฎีหรือไม่	- นิยามกรอบแนวคิดสมรรถนะทางพฤติกรรมให้ชัดเจน เช่น กรอบแนวคิดของ Sue et al. - ใช้ผู้เชี่ยวชาญด้านพฤติกรรมร่วมพิจารณาความครอบคลุมของข้อคำถามในแต่ละมิติ - ใช้การวิเคราะห์เชิงเนื้อหา
2. ความตรงด้านสาระสำคัญ (Substantive Validity)	รายการข้อคำถามสอดคล้องสะท้อนกระบวนการคิดหรือพฤติกรรมภายในของผู้ตอบหรือไม่	- วิเคราะห์กระบวนการรับรู้/ตัดสินใจของผู้ตอบ (think-aloud protocols) - ออกแบบข้อคำถามโดยใช้สถานการณ์จำลองหรือคำถามเปิดเพื่อให้ผู้ตอบแสดงเหตุผลหรือเจตคติจริง แล้วนำมาสร้างเป็นคำถามปลายปิด
3. ความตรงเชิงโครงสร้าง (Structural Validity)	โครงสร้างของเครื่องมือสอดคล้องกับโครงสร้างแนวคิดของสมรรถนะหรือไม่	- ใช้การวิเคราะห์องค์ประกอบ (EFA/CFA) ตรวจสอบมิติของแบบวัด - ตรวจสอบว่าองค์ประกอบต่าง ๆ สะท้อนโครงสร้างตามทฤษฎี เช่น ความตระหนักรู้ ความรู้ทักษะ
4. การสรุปอ้างอิง (Generalizability)	เครื่องมือสามารถใช้ได้ในกลุ่มตัวอย่างหรือบริบทที่หลากหลายหรือไม่	- ศึกษาความไม่แปรเปลี่ยนของการวัดระหว่างกลุ่ม เช่น เชื้อชาติ ภาษา ประเทศ - ใช้วิธีการสุ่มกลุ่มตัวอย่างจากบริบทต่าง ๆ
5. ความตรงภายนอก (External Validity)	คะแนนจากเครื่องมือสัมพันธ์กับผลลัพธ์หรือพฤติกรรมจริงที่คาดว่าจะเกิดขึ้นหรือไม่	- ตรวจสอบความสัมพันธ์ระหว่างคะแนนแบบวัดกับพฤติกรรมที่ประเมินจากอาจารย์ เพื่อน หรือผลสัมฤทธิ์ด้านการสื่อสาร - ทดสอบสหสัมพันธ์กับเครื่องมืออื่นที่วัดสมรรถนะคล้ายกัน (Convergent Validity)



ตารางที่ 1 ข้อเสนอแนะสำหรับการออกแบบและพัฒนาเครื่องมือประเมิน โดยอ้างอิงจากกรอบแนวคิดความตรงแบบองค์รวม (ต่อ)

ด้านความตรงของ Messick	คำอธิบาย	ข้อเสนอแนะเพื่อการออกแบบและพัฒนาเครื่องมือ
6. ความตรงเชิงผลกระทบ (Consequential Validity)	การใช้เครื่องมือวัดเกิดผล กระทบเชิงบวกหรือเชิงลบต่อผู้เรียนหรือกระบวนการทางการศึกษาอย่างไร	- พิจารณาว่าเครื่องมือสร้างความเหลื่อมล้ำทางวัฒนธรรมหรือไม่ - ตรวจสอบว่าการใช้ผลคะแนนนำไปสู่การพัฒนาหรือการตัดสินใจที่ยุติธรรม - สร้างระบบป้อนกลับเพื่อใช้ผลการประเมินในการพัฒนา

2. พัฒนาเครื่องมือประเมินโดยยึดหลักการตรวจสอบความตรงด้วยการโต้แย้ง Kane, M. T. เสนอว่าการประเมินที่มีความตรงควรอิงกับการจัดเรียง “ข้ออ้าง” (Inferences) ที่นำไปสู่การตัดสินใจจากผลการวัด และต้องจัดหาหลักฐานเชิงประจักษ์เพื่อรองรับข้ออ้างเหล่านั้นอย่างเป็นระบบ โดยกำหนดข้ออนุมานในแต่ละขั้น ได้แก่ พฤติกรรม → คะแนน → การตีความ → การตัดสินใจ และวางแผนการเก็บข้อมูลให้สอดคล้องกับแต่ละขั้นตอน ตัวอย่างแนวทาง ได้แก่ การตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ การวิเคราะห์โครงสร้างภายใน (Internal Structure) การตรวจสอบความไม่แปรเปลี่ยนของการวัดระหว่างกลุ่มต่างวัฒนธรรม การวิเคราะห์หลักฐานภายนอก (External Evidence) โดยพิจารณาความสัมพันธ์กับตัวแปรอื่น ตลอดจนการประเมินผลกระทบจากการใช้เครื่องมือ (Kane, M. T., 2013) ข้อเสนอแนะสำหรับการออกแบบและพัฒนาเครื่องมือประเมินโดยใช้กรอบแนวคิดนี้ แสดงไว้ในตารางที่ 2

ตารางที่ 2 ข้อเสนอแนะสำหรับการออกแบบและพัฒนาเครื่องมือประเมินโดยใช้หลักการตรวจสอบความตรงด้วยการโต้แย้ง

ขั้นตอนการอนุมาน (Inference)	คำอธิบาย	ข้อเสนอแนะแนวทางการออกแบบและตรวจสอบความตรง
1. การให้คะแนน (Scoring)	วิธีการให้คะแนนถูกต้อง สอดคล้องกับพฤติกรรมหรือคำตอบที่แสดง สมรรถนะหรือไม่	- สร้างเกณฑ์การให้คะแนน (Rubric) อย่างชัดเจน สอดคล้องกับสมรรถนะที่ต้องการวัด - ตรวจสอบความตรงของข้อคำถาม (Content or Construct Validity) - ตรวจสอบความสามารถในการจำแนก (Item Discrimination) - ตรวจสอบความเที่ยงระหว่างผู้ให้คะแนน (Inter-rater Reliability)
2. การสรุปอ้างอิง (Generalization)	คะแนนที่ได้สะท้อนสมรรถนะของผู้เรียนในสถานการณ์อื่น ๆ หรือไม่	- ใช้แบบวัดในหลายสถานการณ์หรือหลายเงื่อนไข - ตรวจสอบความเที่ยงโดยทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด - ใช้วิธีสมผสานเพื่อขยายขอบเขตการตีความ

ตารางที่ 2 ข้อเสนอแนะสำหรับการออกแบบและพัฒนาเครื่องมือประเมินโดยใช้หลักการตรวจสอบความตรงด้วยการโต้แย้ง

ขั้นตอนการอนุมาน (Inference)	คำอธิบาย	ข้อเสนอแนะแนวทางการออกแบบและตรวจสอบความตรง
3. การอนุมานคะแนนไป ยังคุณลักษณะที่ต้องการ วัด (Extrapolation)	คะแนนจากการประเมิน สะท้อนความสามารถ ที่เกี่ยวข้องได้จริงหรือไม่	- การวิเคราะห์สหสัมพันธ์ (Correlation) กับตัวแปรเกณฑ์ (Criterion-Related Validity) - การสังเกตพฤติกรรมจริง (Real-World Observation)
4. การใช้ผลการประเมิน เพื่อตัดสินใจ (Decision/ Utilization)	ผลการประเมินถูก นำไปใช้อย่างเหมาะสม และไม่ก่อให้เกิด ผลกระทบเชิงลบ	- วิเคราะห์ผลกระทบจากการใช้คะแนน (Washback) - ใช้โมเดลการวินิจฉัย (Diagnostic Classification Models) - ใช้การวิเคราะห์ DIF - การใช้ propensity สกอร์ (Propensity Score) เพื่อสนับสนุน การประเมินผลลัพธ์ที่ตามมาจากการใช้เครื่องมือวัด - ใช้เพื่อการพัฒนา (Assessment for Learning)

3. ออกแบบเครื่องมือให้สะท้อนสภาพจริง โดยผสมผสานรูปแบบต่าง ๆ เช่น แบบรายงานตนเอง แบบทดสอบการตัดสินใจตามสถานการณ์ (SJTs) การสังเกตพฤติกรรม หรือการวิเคราะห์วาทกรรมข้ามวัฒนธรรม เพื่อให้ได้ข้อมูลที่รอบด้าน

ข้อเสนอแนะการนำไปใช้ในบริบทที่มีความหลากหลายทางวัฒนธรรม

ในบริบทการศึกษา การประเมินสมรรถนะทางพหุวัฒนธรรมควรคำนึงถึงความแตกต่างของบริบท วัฒนธรรมของผู้เรียนในแต่ละประเทศหรือภูมิภาคอย่างรอบด้าน โดยเฉพาะประเด็นต่อไปนี้

1. ความเป็นธรรมทางวัฒนธรรม (Cultural Fairness) เครื่องมือควรได้รับการออกแบบให้เหมาะสมกับผู้เรียนหลากหลายวัฒนธรรม ไม่ใช่ภาษาหรือกรณีตัวอย่างที่ลำเอียง (Bias) ทางวัฒนธรรมหรือมีอคติ

2. ความไม่แปรเปลี่ยนของการวัด ก่อนการนำเครื่องมือไปใช้กับกลุ่มตัวอย่างที่แตกต่างกันในด้านสังคม วัฒนธรรม ภาษา หรือมาจากกลุ่มประชากรที่แตกต่างกัน การวิเคราะห์กลุ่มพหุ (Multi-sample Analysis) เพื่อวิเคราะห์ความไม่แปรเปลี่ยนของการวัดจากโมเดลสมการโครงสร้างนั้นถือเป็นสิ่งสำคัญ กล่าวคือ มีการตีความการวัดด้วยแบบวัดที่ตรงกันหรือไม่ อย่างไร หรือมีความแตกต่างของคะแนนในแต่ละองค์ประกอบเหมือนกันหรือไม่

3. การมีส่วนร่วมของผู้มีส่วนได้ส่วนเสีย การพัฒนาเครื่องมือควรเปิดโอกาสให้ผู้เชี่ยวชาญด้านวัฒนธรรม นักการศึกษา และผู้เรียนจากหลากหลายบริบทเข้ามามีส่วนร่วมเพื่อให้เกิดความหลากหลายทางวาทกรรมและมุมมองในกระบวนการพัฒนา

ข้อเสนอแนะเชิงวิจัย

1. การวิเคราะห์ความตรง ควรออกแบบการตรวจสอบความตรงโดยใช้แนวทางการตรวจสอบความตรงด้วยการโต้แย้งที่นำไปสู่การตัดสินใจจากผลการประเมินและการจัดหาหลักฐานเชิงประจักษ์เพื่อรองรับข้ออนุมานเหล่านั้นอย่างเป็นระบบ หรือใช้กรอบแนวคิดความตรงแบบองค์รวม เช่น การตรวจสอบผลกระทบจากการใช้ผลการทดสอบสมรรถนะทางพหุวัฒนธรรม

2. การศึกษาความเชื่อมโยงกับผลลัพธ์การเรียนรู้ ควรมีการศึกษาว่า สมรรถนะทางพหุวัฒนธรรม มีความสัมพันธ์กับผลสัมฤทธิ์ทางการเรียน การปรับตัวในต่างวัฒนธรรม หรือความสามารถในการทำงานใน สภาพแวดล้อมพหุวัฒนธรรมหรือไม่ เพื่อยืนยันความสำคัญของสมรรถนะนี้

3. การพัฒนาเครื่องมือแบบข้ามสาขา เครื่องมือวัดในอนาคตควรบูรณาการองค์ความรู้จากจิตวิทยา การศึกษา มานุษยวิทยา ภาษาศาสตร์สังคม และเทคโนโลยีการเรียนรู้ เช่น การวิเคราะห์ข้อมูลจากกิจกรรม ออนไลน์ หรือการประเมินโดยใช้ปัญญาประดิษฐ์

4. การใช้วิธีผสมผสาน งานวิจัยที่ใช้ทั้งข้อมูลเชิงปริมาณและเชิงคุณภาพควรได้รับการส่งเสริม เพื่อเข้าใจ ความตระหนักรู้ ความรู้ ทักษะ และการตีความทางวัฒนธรรมได้ลึกซึ้งยิ่งขึ้น

สรุป

สมรรถนะทางพหุวัฒนธรรมเป็นคุณลักษณะที่จำเป็นต่อการดำรงชีวิตในโลกยุคใหม่ ซึ่งต้องการมากกว่า ความรู้พื้นฐานทางวิชาการ แต่รวมถึงความสามารถในการเข้าใจ ยอมรับ และอยู่ร่วมกับความแตกต่างอย่างสร้างสรรค์ เพื่อให้การประเมินสมรรถนะนี้มีประสิทธิภาพ เครื่องมือวัดต้องอิงแนวคิดความตรงแบบองค์รวมที่พิจารณาทั้งความ ตรงของเนื้อหา สาระสำคัญ โครงสร้างตามทฤษฎี การสรุปอ้างอิง ความสัมพันธ์กับตัวแปรอื่น และผลกระทบที่ตามมา นอกจากนี้ กรอบการตรวจสอบความตรงด้วยการโต้แย้งยังมีบทบาทสำคัญในการสร้างข้ออ้างที่มีหลักฐานรองรับ อย่างเป็นระบบ สำหรับการใช้คะแนนในเชิงนโยบายหรือการเรียนรู้ หากการพัฒนาเครื่องมือดำเนินไปอย่างรอบด้าน มีความเป็นธรรม และตอบสนองความหลากหลายทางวัฒนธรรม จะสามารถสะท้อนสมรรถนะที่แท้จริงของผู้เรียน และสนับสนุนการออกแบบการเรียนรู้ที่พัฒนา “สมรรถนะเพื่อการอยู่ร่วมกัน” (Competence for Living Together) ซึ่งเป็นหัวใจของการศึกษาที่มีคุณภาพในศตวรรษที่ 21 ตามเป้าหมายของ UNESCO

เอกสารอ้างอิง

- American Educational Research Association (AERA), American Psychological Association (APA), And National Council on Measurement in Education (NCME). (2014). Standards for educational and psychological testing. (7th ed.). United States of America: American Psychological Association.
- ASCD. (2024). How to give effective feedback to your students [Webinar]. ASCD. Retrieved June 25, 2025, from <https://www.ascd.org/webinars/how-to-give-effective-feedback-to-your-students>
- Banks, J. A. (2014). An introduction to multicultural education. (5th ed.). London: Pearson.
- Campinha-Bacote, J. (2002). The process of cultural competence in the delivery of healthcare services: A model of care. *Journal of Transcultural Nursing*, 13(3), 181-184.
- Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education*, 10(3), 241-266.



- Griffith, R. L. et al. (2016). Assessing Intercultural Competence in Higher Education: Existing Research and Future Directions. Retrieved June 25, 2025, from <https://www.mccc.edu/~lyncha/documents/Assessinginterculturalcompetence-ets212112.pdf>
- Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112(3), 527-535.
- _____. (2006). *Validation. Educational measurement*. (4th ed.). Westport, CT: Praeger.
- _____. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1-73.
- Lie, D. A. et al. (2011). Does cultural competency training of health professionals improve patient outcomes? A systematic review and proposed algorithm for future research. *Journal of General Internal Medicine*, 26(3), 317-325.
- Luo, J. & Chan, C. K. Y. (2022). Qualitative methods to assess intercultural competence in higher education research: A systematic review with practical implications. *Educational Research Review*, 37(2022), 100476. <https://doi.org/10.1016/j.edurev.2022.100476> .
- Messick, S. (1989). *Validity. Educational Measurement*. (3rd ed.). New York: American Council on Education and Macmillan.
- _____. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749.
- OECD. (2020). *Global competence for an inclusive world*. Paris: OECD Publishing.
- PISA. (2018). *PISA 2018 assessment and analytical framework*. Paris: OECD Publishing.
- Pope, R. L. & Reynolds, A. L. (1997). *Multicultural competence in student affairs*. New Jersey: Jossey-Bass.
- Pope, R. L. et al. (2019). *Multicultural Competence in Student Affairs: Advancing Social Justice and Inclusion*. San Francisco: John Wiley & Sons.
- Sue, D. W. et al. (1992). Multicultural counseling competencies and standards: A call to the profession. *Journal of Counseling & Development*, 70(4), 477-486.
- Ting-Toomey, S. (1999). *Communicating across cultures*. New York: The Guilford Press.
- UNESCO. (2021). *Education for sustainable development: Towards achieving the SDGs*. Paris: UNESCO Publishing.
- Van de Vijver, F. & Leung, K. (1997). *Methods and data analysis for cross cultural research*. Thousand Oaks, CA: Sage Publications.
- Wilcox, M. M. et al. (2020). Who's multiculturally competent? Everybody and nobody: A multimethod examination. *The Counseling Psychologist*, 48(4), 369-401.

Zumbo, B. D. (1999). A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.