



การพัฒนาแบบทดสอบวัดแววความเป็นครูสำหรับผู้ประกอบวิชาชีพครู

The Development of a Teacher Aptitude Test for Prospective Teachers

พิมพ์ทอง สังกุทธิพงษ์^{1*}, พรพิมล ดิษฐสุนนท์² และ ธิดารัตน์ วิชัยรัตน์³

Pimthong Sungsutthipong^{1*}, Bhornpimon Ditsunon² and Thidarat Wichairat³

¹ ดร. อาจารย์, หลักสูตรศึกษาศาสตรมหาบัณฑิต, สาขาวิจัยและเทคโนโลยีการจัดการเรียนรู้, คณะศึกษาศาสตร์ มหาวิทยาลัยฟาอีสเทิร์น

¹ Dr., Teacher, Master of Education Program in Research and Technology in Learning Management, Faculty of Education, The Far Eastern University.

^{2,3} อาจารย์ประจำคณะศึกษาศาสตร์, มหาวิทยาลัยฟาอีสเทิร์น

^{2,3} Teacher, Faculty of Education, The Far Eastern University.

*Corresponding author, E-mail: pimthong@feu.ac.th

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแบบทดสอบวัดแววความเป็นครูสำหรับผู้ประกอบวิชาชีพครู กลุ่มตัวอย่างในการเก็บข้อมูล มี 2 กลุ่ม ได้แก่ 1) กลุ่มตัวอย่างสำหรับตรวจสอบความเที่ยงตรงเชิงเนื้อหา คือ ผู้เชี่ยวชาญ จำนวน 3 คน ได้จากการเลือกแบบเจาะจง และ 2) กลุ่มตัวอย่างสำหรับการวิเคราะห์ ค่าอำนาจจำแนก เป็นนักศึกษาคณะศึกษาศาสตร์ สาขาวิชาชีพครู ระดับบัณฑิตศึกษา ชั้นปีที่ 2 มหาวิทยาลัยฟาอีสเทิร์น จำนวน 102 คน ได้จากการสุ่มอย่างง่าย และ 3) กลุ่มตัวอย่างสำหรับการวิเคราะห์ ค่าความเชื่อมั่นของแบบทดสอบ เป็นนักศึกษาคณะศึกษาศาสตร์ สาขาวิชาชีพครู ระดับบัณฑิตศึกษา ชั้นปีที่ 2 จำนวน 36 คน ซึ่งเป็นคนละกลุ่มกับกลุ่มที่ 2 เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบวัดแวว ความเป็นครูสำหรับผู้ประกอบวิชาชีพครู แบบปรนัย เลือกตอบ 5 ตัวเลือก จำนวน 1 ฉบับ 70 ข้อ สถิติที่ใช้ วิเคราะห์ความเที่ยงตรงเชิงเนื้อหา ได้แก่ ดัชนีความสอดคล้อง (IOC) วิเคราะห์ค่าอำนาจจำแนกดัชนีบี (B-Index) และค่าความเชื่อมั่นทั้งฉบับของโลเวท โดยใช้โปรแกรม RTAP ผลการวิจัย พบว่า แบบทดสอบวัดแวว ความเป็นครูสำหรับผู้ประกอบวิชาชีพครู มีค่าดัชนีความสอดคล้องอยู่ระหว่าง 0.67-1.00 ค่าอำนาจจำแนก อยู่ระหว่าง 0.22 ถึง 1.00 และค่าความเชื่อมั่นทั้งฉบับ เท่ากับ 0.99

คำสำคัญ: แบบทดสอบวัดแววความเป็นครู, วิชาชีพครู, ผู้ประกอบวิชาชีพครู

Abstract

This research aimed to develop a teacher aptitude test for prospective teachers. The data collection sample consisted of two groups: 1) The content validity assessment sample group consisted of three experts selected through purposive sampling; and 2) The discrimination index analysis sample group included 102 graduate students second year in the Teacher Professional Development Program at The Far Eastern University, selected using simple random sampling; and 3) The sample group for reliability analysis sample group included 36 graduate students second year in the Teacher Professional Development Program at The Far Eastern University, selected using simple random sampling, different from Group 2. The research instrument was a multiple-choice test with five options. Statistical analysis of content validity used the index of Item-Objective



Congruence (IOC), discrimination index (B-Index), and reliability (Lovett method) using the RTAP program. The research results showed that The Teacher Aptitude Test for Prospective Teachers had an IOC between 0.67-1.00, discrimination index between 0.22-1.00, and an overall reliability coefficient of 0.99.

Keywords: Teacher Aptitude Test, Teacher Profession, Prospective Teachers

บทนำ

ครูเป็นทรัพยากรมนุษย์ที่มีบทบาทสำคัญอย่างยิ่งต่อการพัฒนาคุณภาพการศึกษาและการพัฒนาประเทศในระยะยาว เนื่องจากครูเป็นผู้มีหน้าที่โดยตรงในการจัดการเรียนรู้ ส่งเสริมพัฒนาการของผู้เรียน และหล่อหลอมคุณลักษณะอันพึงประสงค์ของเยาวชนในสังคม วิชาชีพครูจึงได้รับการยอมรับว่าเป็นวิชาชีพชั้นสูงที่ต้องอาศัยทั้งองค์ความรู้เชิงวิชาการ ความสามารถในการปฏิบัติงาน และคุณลักษณะเชิงจิตวิชาชีพควบคู่กัน (บุญชม ศรีสะอาด, 2560 ข; สมหวัง พิธิยานุวัฒน์, 2558) ทั้งนี้ ความเป็นครูมิได้สะท้อนผ่านผลสัมฤทธิ์ทางวิชาการหรือความรู้เฉพาะสาขาเพียงอย่างเดียว หากแต่รวมถึงคุณธรรม จริยธรรมความรับผิดชอบ และทัศนคติที่เหมาะสมต่อวิชาชีพ ซึ่งสอดคล้องกับกรอบมาตรฐานและจรรยาบรรณวิชาชีพครูของคุรุสภา (คุรุสภา, 2562) ในบริบทการศึกษาในศตวรรษที่ 21 บทบาทของครูได้เปลี่ยนแปลงจากการเป็นผู้ถ่ายทอดความรู้ไปสู่การเป็นผู้ออกแบบการเรียนรู้ ผู้อำนวยการความสะดวกในการเรียนรู้ และผู้ส่งเสริมการพัฒนาผู้เรียนอย่างรอบด้าน ครูต้องสามารถจัดการเรียนรู้ที่ตอบสนองความแตกต่างระหว่างบุคคล ใช้เทคโนโลยีอย่างเหมาะสม และทำงานร่วมกับชุมชนและผู้มีส่วนได้ส่วนเสียทางการศึกษา (อนุวัติ คุณแก้ว, 2563) ความท้าทายดังกล่าวสะท้อนให้เห็นว่า การมีความรู้ทางวิชาการเพียงอย่างเดียวอาจไม่เพียงพอต่อการประกอบวิชาชีพครู หากขาดคุณลักษณะภายในที่สะท้อนความเหมาะสมต่อบทบาทวิชาชีพ

แนวคิดเกี่ยวกับ “แนวความเป็นครู” (Teacher Dispositions หรือ Teaching Aptitude) จึงได้รับความสนใจในวงวิชาการด้านการผลิตและพัฒนาครู โดยแนวความเป็นครูเป็นแนวโน้มหรือศักยภาพภายในของบุคคลที่สะท้อนถึงความเหมาะสมในการประกอบวิชาชีพครู ครอบคลุมมิติด้านทัศนคติ ค่านิยม จริยธรรม ความรับผิดชอบ และความสามารถในการตัดสินใจเชิงวิชาชีพ (Messick, 1995; Nitko & Brookhart, 2011) แนวคิดดังกล่าวแตกต่างจากการวัดความรู้หรือทักษะเชิงวิชาการ เนื่องจากมุ่งประเมินคุณลักษณะเชิงนามธรรมที่ส่งผลต่อการปฏิบัติงานในระยะยาว งานวิจัยในบริบทประเทศไทยสะท้อนให้เห็นว่า แนวความเป็นครูเกี่ยวข้องกับคุณลักษณะสำคัญหลายประการ ได้แก่ ความรักและเมตตาต่อผู้เรียน ความรับผิดชอบ ความอดทน ความยุติธรรม และการเสียสละ ซึ่งเป็นรากฐานของจิตวิญญาณความเป็นครู (เอกภูมิ จันทระขันตี, 2557) นอกจากนี้ ยังรวมถึงความสามารถในการจัดการเรียนรู้ การสื่อสาร และการสร้างความสัมพันธ์ที่ดีระหว่างครูกับผู้เรียนและชุมชน (อนุวัติ คุณแก้ว, 2563) สิริมงคล สุวรรณผา (2566) ได้เสนอว่าคุณลักษณะดังกล่าวสามารถพัฒนาเป็นตัวชี้วัดคุณลักษณะความเป็นครูมืออาชีพได้อย่างเป็นระบบ ทั้งในมิติด้านสมรรถนะและมิติด้านคุณค่าเชิงจริยธรรม อย่างไรก็ตาม ในทางปฏิบัติการคัดเลือกและประเมินผู้ประกอบวิชาชีพครู ยังคงให้ความสำคัญกับการวัดผลสัมฤทธิ์ทางการเรียนหรือความรู้เฉพาะด้านเป็นหลัก ขณะที่การประเมินคุณลักษณะเชิงจิตวิชาชีพหรือแนวความเป็นครูยังขาดเครื่องมือที่พัฒนาอย่างเป็นระบบและผ่านกระบวนการตรวจสอบคุณภาพตามหลักการวัดทางการศึกษา (Linn & Gronlund, 2000; Popham, 2014) แม้จะมีงานวิจัยบางส่วนที่พยายามพัฒนาตัวชี้วัดหรือแบบประเมินคุณลักษณะความเป็นครูในประเทศไทย แต่ยังไม่พบข้อจำกัดด้านความครอบคลุมขององค์ประกอบและกระบวนการตรวจสอบคุณภาพเครื่องมือ



ด้วยเหตุนี้ การพัฒนาแบบทดสอบวัดแววความเป็นครูสำหรับผู้ประกอบวิชาชีพครูจึงมีความสำคัญ ทั้งในเชิงวิชาการและเชิงปฏิบัติ เครื่องมือดังกล่าวจะช่วยให้สามารถประเมินแนวโน้มความเหมาะสมของ บุคคลต่อการประกอบวิชาชีพครูอย่างเป็นระบบ มีความน่าเชื่อถือ และสอดคล้องกับบริบทวิชาชีพครู ในประเทศไทย อีกทั้งยังสามารถใช้เป็นเครื่องมือประกอบการคัดเลือก การพัฒนา และการกำหนดนโยบาย ด้านการผลิตและพัฒนาครู อันจะนำไปสู่การยกระดับคุณภาพครูและคุณภาพการศึกษาโดยรวมในระยะยาว

วัตถุประสงค์ของการวิจัย

เพื่อศึกษาคุณภาพของแบบทดสอบวัดแววความเป็นครูสำหรับผู้ประกอบวิชาชีพครู ด้านความเที่ยงตรงเชิงเนื้อหา ค่าอำนาจจำแนก และค่าความเชื่อมั่น

แนวคิด ทฤษฎี กรอบแนวคิด

การวิจัยครั้งนี้ได้ศึกษาแนวคิด ทฤษฎี ดังนี้

แนวคิดเกี่ยวกับวิชาชีพครูและความเป็นครู

วิชาชีพครูจัดเป็นวิชาชีพชั้นสูงที่มีบทบาทสำคัญต่อการพัฒนาทรัพยากรมนุษย์และสังคมโดยรวม เนื่องจากครูเป็นผู้มีหน้าที่โดยตรงในการจัดการเรียนรู้ ส่งเสริมพัฒนาการของผู้เรียน และหล่อหลอม คุณลักษณะอันพึงประสงค์ของเยาวชนในสังคม การประกอบวิชาชีพครูจึงมิได้อาศัยเพียงความรู้ทางวิชาการ เท่านั้น หากแต่ต้องอาศัยความรู้เฉพาะทางด้านการจัดการเรียนรู้ ความสามารถในการปฏิบัติงานจริง และ คุณลักษณะจริยธรรมที่สอดคล้องกับมาตรฐานของวิชาชีพ (บุญชม ศรีสะอาด, 2560 ข; สมหวัง พิธิยานุวัฒน์, 2558) แนวคิดเกี่ยวกับความเป็นครูในเชิงวิชาชีพ มุ่งเน้นการมองครูในฐานะผู้ประกอบวิชา ชีพที่ต้องมีความรับผิดชอบต่อผู้เรียนและสังคม มีทัศนคติที่ดีต่อวิชาชีพ และยึดมั่นในจรรยาบรรณวิชาชีพครู ซึ่งเป็น กรอบสำคัญในการกำกับพฤติกรรมและการตัดสินใจของครูในการปฏิบัติงานจริง (ครูสภา, 2562) คุณลักษณะ ดังกล่าวสะท้อนให้เห็นว่า ความเป็นครู มิได้วัดได้จากผลสัมฤทธิ์ทางวิชาการหรือความรู้ในเนื้อหาวิชาเพียง อย่างเดียว แต่รวมถึงมิติด้านคุณธรรม จริยธรรม และความมุ่งมั่นต่อการพัฒนาผู้เรียนอย่างรอบด้าน นักวิชาการหลายท่านเสนอว่า ครูที่มีคุณภาพควรมีทั้งสมรรถนะทางวิชาชีพและคุณลักษณะ ส่วนบุคคล ที่เหมาะสม เช่น ความเมตตา กรุณา ความเสียสละ ความอดทน ความสามารถในการทำงานร่วมกับผู้อื่น และ ความสามารถในการสื่อสาร คุณลักษณะเหล่านี้เป็นปัจจัยสำคัญที่ส่งผลต่อประสิทธิภาพในการจัดการ เรียนรู้และความไว้วางใจของผู้เรียนและสังคมต่อวิชาชีพครู (Linn & Gronlund, 2000; Popham, 2014) ดังนั้น การพัฒนาครูจึงควรมุ่งเน้นทั้งการเสริมสร้างความรู้ความสามารถและการพัฒนาความเป็นครูในเชิง จิตวิชาชีพควบคู่กันไป

แนวคิดเกี่ยวกับแววความเป็นครู

แววความเป็นครู เป็นแนวคิดที่เกี่ยวข้องกับความถนัด และคุณลักษณะภายในของบุคคลที่เอื้อต่อ การประกอบวิชาชีพครูอย่างมีประสิทธิภาพ แนวคิดนี้มุ่งอธิบายลักษณะทางจิตใจ อารมณ์ และทัศนคติ ที่ทำให้บุคคลสามารถปรับตัว ปฏิบัติงาน และตัดสินใจในบทบาทของครูได้อย่างเหมาะสม แม้บุคคลจะมีความรู้ทางวิชาการในระดับสูง แต่หากขาดแววความเป็นครู ก็อาจไม่สามารถปฏิบัติหน้าที่ครูได้อย่างมี คุณภาพ (Nitko & Brookhart, 2011; บุญชม ศรีสะอาด, 2560 ข) นักวิชาการด้านการศึกษาเสนอว่า แววความเป็นครู ควรประกอบด้วยคุณลักษณะสำคัญหลายประการ เช่น ความเอื้ออาทรและความเข้าใจ ผู้เรียน ความอดทนและการควบคุมอารมณ์ ความรับผิดชอบต่อหน้าที่ ความยุติธรรม และความสามารถ ในการตัดสินใจเชิงจริยธรรมในสถานการณ์ทางวิชาชีพ คุณลักษณะเหล่านี้เป็นพื้นฐานสำคัญของการสร้าง ความสัมพันธ์ที่ดีกับผู้เรียน และส่งผลต่อบรรยากาศการเรียนรู้ในห้องเรียน (Messick, 1995; Popham, 2014; เอกภูมิ จันทร์ขันธ์, 2557) ส่วนแนวคิดแววความเป็นครู สอดคล้องกับการมองครูในฐานะผู้ประกอบ



วิชาชีพที่ต้องมีทั้งสมรรถนะ (Competence) และอัตลักษณ์ความเป็นครู (Teacher Identity) กล่าวคือ ครูไม่เพียงแต่ “ทำหน้าที่สอน” ได้อย่างถูกต้องตามหลักวิชาการเท่านั้น แต่ยังต้อง “เป็นครู” ในเชิงคุณค่า ทักษะจิต และจิตสำนึกในวิชาชีพด้วย (Brennan, 2006) การประเมินแนวความเป็นครู จึงเป็นการประเมินแนวโน้มความเหมาะสมของบุคคลต่อการประกอบวิชาชีพครูในระยะยาว มากกว่าการวัดความรู้หรือทักษะเฉพาะด้านเพียงช่วงเวลาใดเวลาหนึ่ง ด้วยเหตุนี้ การพัฒนาแบบทดสอบวัดแนวความเป็นครู จึงมีความสำคัญอย่างยิ่งต่อระบบการผลิตและพัฒนาครู ที่จะช่วยให้สามารถประเมินคุณลักษณะเชิงจิตวิชาชีพที่เป็นรากฐานของความเป็นครู และใช้เป็นข้อมูลประกอบการคัดเลือก การพัฒนาหรือการส่งเสริมผู้ประกอบวิชาชีพครูให้มีคุณภาพตามมาตรฐานที่สังคมคาดหวัง

องค์ประกอบของแนวความเป็นครู

มีงานวิจัยในประเทศไทยที่ศึกษาลักษณะและคุณลักษณะของครูไทยในศตวรรษที่ 21 และการพัฒนาตัวชี้วัดคุณลักษณะวิชาชีพครู สามารถนำมาอธิบายเป็นองค์ประกอบหลักของแนวความเป็นครู ดังนี้ (เอกภูมิ จันทร์ขันธ์, 2557; อนุวัติ คุณแก้ว, 2563; สิริมงคล สุวรรณผา, 2566)

1) ทักษะและความสามารถในการจัดการเรียนรู้และการสอน งานวิจัยที่ศึกษาคุณลักษณะของครูไทยในศตวรรษที่ 21 พบว่าองค์ประกอบที่สำคัญ ครูต้องมีความสามารถในการจัดการเรียนรู้และเป็นผู้ดำเนินการสอนเรียน ซึ่งรวมถึงทักษะการออกแบบและใช้สื่อ การวางแผนสอน และการปรับใช้เทคโนโลยี

2) ความรู้และสมรรถนะเชิงองค์รวม องค์ประกอบนี้ครอบคลุมถึง ความรู้ด้านเนื้อหา ทักษะความสามารถเฉพาะทาง และความเข้าใจบริบทของผู้เรียน ซึ่งช่วยให้ครูสามารถตอบสนองต่อความต้องการของผู้เรียนอย่างมีประสิทธิภาพ

3) ทักษะและคุณค่าเชิงจริยธรรม งานวิจัยในบริบทไทยระบุว่าคุณลักษณะเชิงคุณค่า เช่น ความรับผิดชอบต่อหน้าที่ จริยธรรม ความยุติธรรม และความเสียสละ เป็นองค์ประกอบสำคัญของแนวความเป็นครู ซึ่งสัมพันธ์กับมาตรฐานวิชาชีพครูที่คาดหวังจากสังคม

4) ทักษะสื่อสารและการสร้างความสัมพันธ์ งานวิจัยเกี่ยวกับคุณลักษณะของครูที่ดี พบว่าความสามารถในการสื่อสารอย่างมีประสิทธิภาพ และการสร้างความสัมพันธ์ที่ดีระหว่างครูกับผู้เรียนและชุมชน เป็นอีกมิติหนึ่งของแนวความเป็นครูที่พึงประสงค์ ซึ่งส่งผลต่อการเรียนรู้และสภาพแวดล้อมภายในห้องเรียน (อนุวัติ คุณแก้ว, 2563)

ทฤษฎีการวัดทางการศึกษา (Educational Measurement Theory)

ทฤษฎีการวัดทางการศึกษา เป็นกรอบแนวคิดสำคัญที่อธิบายหลักการและกระบวนการในการแปลงคุณลักษณะเชิงนามธรรมของบุคคล เช่น ความรู้ ความสามารถ ทักษะจิต หรือคุณลักษณะทางวิชาชีพ ให้อยู่ในรูปของค่าคะแนนที่สามารถตีความและนำไปใช้ในการตัดสินใจหรือการวิจัยได้อย่างมีเหตุผลและน่าเชื่อถือ (Linn & Gronlund, 2000; บุญชม ศรีสะอาด, 2560 ก) การเลือกใช้ทฤษฎีการวัดที่เหมาะสมจึงมีผลโดยตรงต่อคุณภาพของแบบทดสอบและความถูกต้องของการสรุปผลการวัด

ในงานวิจัยด้านการพัฒนาแบบทดสอบทางการศึกษา ทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) ถือเป็นทฤษฎีพื้นฐานที่ได้รับการใช้อย่างแพร่หลาย โดยมีสมมติฐานสำคัญว่าคะแนนที่ได้จากการทดสอบ (Observed Score) ประกอบด้วยคะแนนจริง (True Score) และความคลาดเคลื่อนจากการวัด (Error Score) (Crocker & Algina, 1986) แนวคิดดังกล่าวชี้ให้เห็นว่าคะแนนที่ปรากฏไม่สามารถสะท้อนความสามารถที่แท้จริงของผู้สอบได้อย่างสมบูรณ์ จึงจำเป็นต้องควบคุมและลดความคลาดเคลื่อนจากการวัดผ่านการพัฒนาแบบทดสอบที่มีคุณภาพ ทฤษฎีการทดสอบแบบดั้งเดิม มุ่งเน้นการวิเคราะห์คุณภาพของแบบทดสอบจากข้อมูลเชิงประจักษ์ โดยพิจารณาทั้งคุณภาพรายข้อและคุณภาพของแบบทดสอบทั้งฉบับ การวิเคราะห์คุณภาพรายข้อ ประกอบด้วย การหาค่าความยากและค่าอำนาจ



จำแนก เพื่อประเมินว่าข้อคำถามสามารถวัดคุณลักษณะที่ต้องการ และแยกแยะผู้ที่มีระดับคุณลักษณะแตกต่างกันได้หรือไม่ ขณะที่คุณภาพของแบบทดสอบทั้งฉบับพิจารณาจากค่าความเชื่อมั่น ซึ่งสะท้อนถึงความคงเส้นคงวาและเสถียรภาพของผลการวัด (Nitko & Brookhart, 2011; ล้วน สายยศ และ อังคนา สายยศ, 2560)

นอกจากทฤษฎีการทดสอบแบบดั้งเดิมแล้ว ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ยังเป็นอีกแนวคิดหนึ่งที่ได้รับการพัฒนาเพื่อเพิ่มความแม่นยำในการวัด โดยอธิบายความสัมพันธ์ระหว่างความสามารถแฝงของผู้ตอบกับความน่าจะเป็นในการตอบข้อคำถามถูกต้อง อย่างไรก็ตาม การใช้ IRT ต้องอาศัยขนาดกลุ่มตัวอย่างที่เหมาะสมและกระบวนการวิเคราะห์ที่ซับซ้อนมากขึ้น (Brennan, 2006) ในบริบทของการวิจัยเพื่อพัฒนาแบบทดสอบเชิงวิชาชีพและการนำเสนอในเวทีประชุมวิชาการระดับชาติ การใช้กรอบทฤษฎีการทดสอบแบบดั้งเดิมจึงยังคงมีความเหมาะสม เนื่องจากสามารถอธิบายขั้นตอนการพัฒนาและการตรวจสอบคุณภาพเครื่องมือได้อย่างชัดเจนและเป็นที่ยอมรับโดยทั่วไป สำหรับการวัดคุณลักษณะเชิงวิชาชีพ เช่น แนวความเป็นครู การประยุกต์ใช้ทฤษฎีการวัดทางการศึกษา ต้องคำนึงถึงลักษณะเฉพาะของตัวแปรที่ไม่สามารถสังเกตได้โดยตรง การพัฒนาแบบทดสอบจึงต้องอาศัยการกำหนดกรอบแนวคิดที่ชัดเจน การสร้างข้อคำถามที่สะท้อนพฤติกรรมหรือการตัดสินใจในบริบทวิชาชีพ และการตรวจสอบคุณภาพของเครื่องมืออย่างเป็นระบบ เพื่อให้การวัดมีความน่าเชื่อถือและสามารถนำไปใช้ประกอบการตัดสินใจหรือพัฒนาวิชาชีพครูได้อย่างเหมาะสม (Messick, 1995; Popham, 2014)

โดยสรุป ทฤษฎีการวัดทางการศึกษาเป็นรากฐานสำคัญของการพัฒนาแบบทดสอบที่มีคุณภาพ การเลือกใช้ทฤษฎีการวัดที่สอดคล้องกับวัตถุประสงค์และลักษณะของการวัด จะช่วยให้แบบทดสอบสามารถสะท้อนคุณลักษณะที่ต้องการวัดได้อย่างถูกต้อง ลดความคลาดเคลื่อนจากการวัด และเพิ่มความน่าเชื่อถือของผลการวิจัยในเชิงวิชาการ

แนวคิดและทฤษฎีการสร้างและหาคุณภาพแบบทดสอบ

การสร้างแบบทดสอบทางการศึกษาเป็นกระบวนการที่ต้องอาศัยหลักการและทฤษฎีด้านการวัดและประเมินผลอย่างเป็นระบบ โดยมีเป้าหมายเพื่อให้ได้เครื่องมือวัดที่สามารถสะท้อนคุณลักษณะหรือโครงสร้างเชิงนามธรรมของบุคคลได้อย่างถูกต้องและน่าเชื่อถือ (บุญชม ศรีสะอาด, 2560 ก; Nitko & Brookhart, 2011) การพัฒนาแบบทดสอบที่มีคุณภาพจึงต้องเริ่มจากการกำหนดจุดมุ่งหมายของการวัด การนิยามตัวแปรที่ต้องการวัดอย่างชัดเจน และการกำหนดกรอบเนื้อหาที่สอดคล้องกับวัตถุประสงค์ของการวัด ในทางทฤษฎีการพัฒนาแบบทดสอบมักอาศัยกรอบแนวคิดของทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) ซึ่งคะแนนที่ได้จากการทดสอบประกอบด้วยคะแนนจริงและความคลาดเคลื่อนจากการวัด (Crocker & Algina, 1986) แนวคิดดังกล่าวถูกนำมาใช้ในการวิเคราะห์คุณภาพของข้อคำถามและแบบทดสอบทั้งฉบับ โดยพิจารณาค่าความยาก อำนาจจำแนก และความเชื่อมั่นของแบบทดสอบ ซึ่งเป็นองค์ประกอบพื้นฐานของการประเมินคุณภาพเครื่องมือวัดทางการศึกษา (ล้วน สายยศ และ อังคนา สายยศ, 2560)

คุณภาพของแบบทดสอบเป็นประเด็นสำคัญที่สะท้อนความเหมาะสมและความน่าเชื่อถือของเครื่องมือวัด โดยทั่วไปพิจารณาจากองค์ประกอบหลัก ได้แก่ ความเที่ยงตรง (Validity) ความเชื่อมั่น (Reliability) และคุณภาพรายข้อ (Item Quality) (สมหวัง พิธิยานุวัฒน์, 2558; Linn & Gronlund, 2000) ความเที่ยงตรงของแบบทดสอบ เป็นระดับที่แบบทดสอบสามารถวัดคุณลักษณะตามที่ตั้งใจจะวัด ซึ่งในงานวิจัยด้านการวัดเชิงจิตวิชาชีพ จะให้ความสำคัญกับความเที่ยงตรงเชิงเนื้อหา โดยอาศัยการพิจารณาของผู้ทรงคุณวุฒิในการตรวจสอบความสอดคล้องระหว่างข้อคำถามกับกรอบแนวคิดและมาตรฐานที่เกี่ยวข้อง (AERA, APA, & NCME, 2014) ส่วนความเชื่อมั่นของแบบทดสอบ เป็นความคงเส้นคงวาของผลการวัดสะท้อนถึงระดับความคลาดเคลื่อนที่ยอมรับได้ของเครื่องมือวัด แบบทดสอบที่มีความเชื่อมั่นสูงย่อมให้ผล



การวัดที่มีเสถียรภาพและสามารถนำไปใช้ในการตัดสินหรือการวิจัยได้อย่างเหมาะสม (บุญชม ศรีสะอาด, 2560 ก; Popham, 2014) การตรวจสอบความเชื่อมั่น จะใช้ค่าสัมประสิทธิ์ความเชื่อมั่น ดังนั้น การพิจารณาคุณภาพรายข้อเป็นขั้นตอนสำคัญในการพัฒนาแบบทดสอบ เนื่องจากช่วยให้ทราบว่าข้อคำถามแต่ละข้อมีความเหมาะสมในการวัดคุณลักษณะที่ต้องการหรือไม่ การวิเคราะห์ค่าความยากและอำนาจจำแนกของข้อคำถามช่วยให้สามารถคัดเลือกหรือปรับปรุงข้อคำถามให้มีประสิทธิภาพในการแยกแยะผู้ที่มีระดับคุณลักษณะแตกต่างกันได้อย่างเหมาะสม (Crocker & Algina, 1986; Nitko & Brookhart, 2011) โดยสรุป การสร้างและหาคุณภาพแบบทดสอบจำเป็นต้องอาศัยทั้งกรอบแนวคิดทางทฤษฎีและกระบวนการตรวจสอบคุณภาพอย่างเป็นระบบ เพื่อให้ได้เครื่องมือวัดที่มีความเที่ยงตรง ความเชื่อมั่น และสามารถนำผลการวัดไปใช้ในการตัดสินหรือพัฒนาทางการศึกษาได้อย่างมีความน่าเชื่อถือ (Messick, 1995)

แนวคิดเกี่ยวกับคุณภาพของแบบทดสอบ

คุณภาพของแบบทดสอบเป็นประเด็นสำคัญที่สะท้อนความน่าเชื่อถือและความเหมาะสมของเครื่องมือวัดในการตีความผลคะแนน โดยทั่วไป การพิจารณาคุณภาพของแบบทดสอบอาศัยองค์ประกอบหลักได้แก่ ความเที่ยงตรง (Validity) ความเชื่อมั่น (Reliability) และคุณภาพรายข้อ (Item Quality) ซึ่งเป็นแนวคิดพื้นฐานในทฤษฎีการวัดทางการศึกษา (AERA, APA, & NCME, 2014; Linn & Gronlund, 2000)

ความเที่ยงตรง (Validity)

ความเที่ยงตรง หมายถึง ระดับที่หลักฐานและทฤษฎีสนับสนุนการตีความคะแนนจากแบบทดสอบให้สอดคล้องกับวัตถุประสงค์ของการวัด (AERA, APA, & NCME, 2014) กล่าวคือ แบบทดสอบที่มีความเที่ยงตรงสูงต้องสามารถวัดคุณลักษณะตามที่ตั้งใจจะวัดได้อย่างเหมาะสม ในการพัฒนาแบบทดสอบวัดแววความเป็นครู ควรให้ความสำคัญกับความเที่ยงตรงเชิงเนื้อหา (Content Validity) โดยพิจารณาความสอดคล้องระหว่างข้อคำถามกับองค์ประกอบของแววความเป็นครูตามกรอบแนวคิดและมาตรฐานวิชาชีพครู ทั้งนี้ การให้ผู้ทรงคุณวุฒิประเมินความสอดคล้องของข้อคำถามกับนิยามเชิงปฏิบัติการของตัวแปร เป็นแนวทางที่ได้รับการยอมรับอย่างกว้างขวางในการเสริมสร้างความเที่ยงตรงเชิงเนื้อหา (บุญชม ศรีสะอาด, 2560 ก; Nitko & Brookhart, 2011)

ความเชื่อมั่น (Reliability)

ความเชื่อมั่น หมายถึง ความคงเส้นคงวาและความสม่ำเสมอของผลการวัด หากนำแบบทดสอบไปใช้ในเงื่อนไขที่คล้ายคลึงกัน ผลคะแนนควรมีเสถียรภาพและมีความคลาดเคลื่อนจากการวัดต่ำ (Popham, 2014) การตรวจสอบความเชื่อมั่นของแบบทดสอบที่วัดคุณลักษณะเชิงนามธรรม เช่น แววความเป็นครู มักใช้ค่าสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's alpha) ซึ่งสะท้อนความสอดคล้องภายในของข้อคำถาม (Cronbach, 1951) ค่าสัมประสิทธิ์ที่อยู่ในระดับ .70 ขึ้นไปมักถือว่ายอมรับได้สำหรับการวิจัยเชิงพัฒนาเครื่องมือ (Nunnally & Bernstein, 1994)

คุณภาพรายข้อ (Item Quality)

คุณภาพรายข้อเป็นการพิจารณาคุณลักษณะของข้อคำถามแต่ละข้อ เช่น ค่าความยาก (Item Difficulty) และค่าอำนาจจำแนก (Item Discrimination) เพื่อประเมินว่าข้อคำถามสามารถแยกแยะผู้ที่มีระดับคุณลักษณะแตกต่างกันได้หรือไม่ (Linn & Gronlund, 2000) การวิเคราะห์คุณภาพรายข้อช่วยให้สามารถคัดเลือก ปรับปรุง หรือยกเลิกข้อคำถามก่อนนำไปใช้จริง ซึ่งเป็นขั้นตอนสำคัญในการพัฒนาแบบทดสอบให้มีคุณภาพ



การหาคุณภาพแบบทดสอบแบบอิงเกณฑ์

แบบทดสอบแบบอิงเกณฑ์ (Criterion-Referenced Test: CRT) เป็นเครื่องมือวัดที่มุ่งประเมินระดับความรู้ ความสามารถ หรือคุณลักษณะของผู้ตอบ โดยเปรียบเทียบกับเกณฑ์หรือมาตรฐานที่กำหนดไว้ล่วงหน้า มิใช่การเปรียบเทียบกับกลุ่มผู้สอบอื่น (Glaser, 1963; Popham, 2014) แบบทดสอบลักษณะนี้เหมาะสมกับการประเมินสมรรถนะขั้นต่ำหรือความพร้อมในการประกอบวิชาชีพ เช่น การประเมินแว่วความเป็นครู แนวคิดพื้นฐานของแบบทดสอบอิงเกณฑ์ คือ การจำแนกผู้ตอบออกเป็น 2 กลุ่ม ได้แก่ ผู้รอบรู้ (Masters) และผู้ไม่รอบรู้ (Nonmasters) ตามเกณฑ์การตัดสิน (Cut Score) ที่กำหนด (Hambleton & Novick, 1973) เกณฑ์ดังกล่าวอาจกำหนดจากมาตรฐานวิชาชีพ การวิเคราะห์เนื้อหา หรือการพิจารณาของผู้เชี่ยวชาญ การตรวจสอบคุณภาพของแบบทดสอบแบบอิงเกณฑ์จึงเน้นว่าข้อคำถามสามารถจำแนกผู้ที่ “ผ่านเกณฑ์” หรือ “ไม่ผ่านเกณฑ์” ได้อย่างถูกต้องและสม่าเสมอ มากกว่าการกระจายคะแนนเหมือนแบบอิงกลุ่ม

การหาค่าอำนาจจำแนกในแบบทดสอบแบบอิงเกณฑ์

แนวคิดของอำนาจจำแนกแบบอิงเกณฑ์

อำนาจจำแนกของข้อคำถามในแบบทดสอบแบบอิงเกณฑ์ หมายถึง ความสามารถของข้อคำถามในการแยกผู้ที่มีคุณลักษณะตามเกณฑ์ออกจากผู้ที่ไม่เป็นไปตามเกณฑ์ (Hambleton & Novick, 1973) กล่าวคือ สัดส่วนการตอบถูกของกลุ่มผู้รอบรู้ควรสูงกว่ากลุ่มผู้ไม่รอบรู้อย่างชัดเจน ต่างจากแบบอิงกลุ่มที่มุ่งแยกผู้มีความสามารถสูง-ต่ำ อำนาจจำแนกแบบอิงเกณฑ์มุ่งแยกตาม “สถานะการผ่านเกณฑ์” เป็นสำคัญ

วิธีการหาค่าอำนาจจำแนกแบบอิงเกณฑ์

1. ดัชนีความแตกต่างระหว่างกลุ่มผู้รอบรู้-ไม่รอบรู้ (Difference Index) คำนวณจาก

$$D = P_m - P_{nm} \text{ โดยที่}$$

P_m = สัดส่วนการตอบถูกของกลุ่มผู้รอบรู้

P_{nm} = สัดส่วนการตอบถูกของกลุ่มผู้ไม่รอบรู้

ค่า D ที่เป็นบวกและมีค่าสูง แสดงว่าข้อคำถามสามารถจำแนกได้ดี โดยงานวิจัยด้านการวัดเสนอว่าค่าอำนาจจำแนกควรมีอย่างน้อย .20 ขึ้นไปจึงถือว่าอยู่ในระดับยอมรับได้ (Ebel & Frisbie, 1991)

2. ดัชนีการจำแนกแบบ Brennan (Brennan's Index)

Brennan (1972) เสนอแนวคิดการพิจารณาอำนาจจำแนกในบริบทแบบอิงเกณฑ์ โดยเน้นความแม่นยำในการจำแนกผู้ตอบตามเกณฑ์มากกว่าการกระจายคะแนน ดัชนีนี้สะท้อนประสิทธิภาพของข้อคำถามในการตัดสินผู้รอบรู้และไม่รอบรู้โดยตรง

3. การวิเคราะห์ความถูกต้องของการจำแนก (Classification Accuracy)

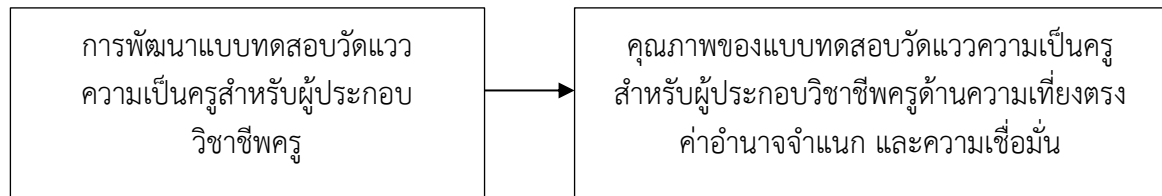
อีกแนวทางหนึ่งคือการสร้างตารางจำแนก (Classification Table) เพื่อพิจารณาสัดส่วนของการจำแนกถูกต้องและคลาดเคลื่อน ซึ่งมีความสำคัญอย่างยิ่งเมื่อผลการทดสอบใช้ในการตัดสินเชิงนโยบายหรือวิชาชีพ (Hambleton & Pitoniak, 2006)

ความสำคัญของการพัฒนาแบบทดสอบวัดแว่วความเป็นครูแบบอิงเกณฑ์

แบบทดสอบวัดแว่วความเป็นครู มีเป้าหมายเพื่อประเมินความเหมาะสมเชิงวิชาชีพ มิใช่เพื่อจัดลำดับคะแนน ดังนั้น ข้อคำถามที่ดีควรสามารถแยกผู้ที่มีคุณลักษณะตามเกณฑ์ออกจากผู้ที่ยังไม่บรรลุเกณฑ์ได้อย่างชัดเจน การตรวจสอบอำนาจจำแนกแบบอิงเกณฑ์จึงเป็นกลไกสำคัญในการยืนยันว่าเครื่องมือสามารถใช้ประกอบการตัดสินหรือพัฒนาวิชาชีพครูได้อย่างมีความน่าเชื่อถือ (Popham, 2014; AERA, APA, & NCME, 2014)



แบบทดสอบวัดแววจความเป็นครูที่พัฒนาขึ้นในงานวิจัยนี้ เป็นแบบทดสอบประเภท Maximum Performance Test ซึ่งเป็นการวัดความสามารถของผู้ตอบในการเลือกคำตอบที่เหมาะสมที่สุดในสถานการณ์ที่กำหนด โดยมีเกณฑ์การให้คะแนนที่ชัดเจนว่าคำตอบใดถูกหรือไม่ถูก (Nitko & Brookhart, 2011) การวิเคราะห์คุณภาพรายข้อใช้ดัชนีอำนาจจำแนกตามแนวคิดของแบบทดสอบอิงเกณฑ์ ดังนั้น งานวิจัยนี้จึงมุ่งพัฒนาแบบทดสอบวัดแววจความเป็นครูในลักษณะแบบอิงเกณฑ์ เพื่อให้สามารถประเมินคุณลักษณะเชิงพฤติกรรมและการตัดสินใจในสถานการณ์วิชาชีพได้อย่างมีประสิทธิภาพ



ภาพที่ 1 กรอบการวิจัย

วิธีดำเนินการวิจัย

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้วิจัย ดังนี้

1. ผู้เชี่ยวชาญ สำหรับการประเมินความเที่ยงตรงเชิงเนื้อหาของแบบทดสอบ จำนวน 3 ท่าน ได้มาโดยการเลือกแบบเจาะจง ได้แก่ คณบดีคณะศึกษาศาสตร์ ผู้เชี่ยวชาญทางการศึกษาและการวัดผล อาจารย์ประจำคณะศึกษาศาสตร์ผู้เชี่ยวชาญทางการศึกษาและเทคโนโลยีการจัดการเรียนรู้ และอาจารย์ประจำคณะศึกษาศาสตร์ผู้เชี่ยวชาญทางการศึกษาและด้านภาษาไทย มหาวิทยาลัยฟาร์อีสเทอร์น

2. นักศึกษากลุ่มที่ 1 สำหรับการหาค่าอำนาจจำแนกของแบบทดสอบ จำนวน 102 คน ได้มาโดยการสุ่มอย่างง่าย (Simple Random Sampling) จากหลักสูตรประกาศนียบัตรบัณฑิตวิชาชีพครู ชั้นปีที่ 2 คณะศึกษาศาสตร์ มหาวิทยาลัยฟาร์อีสเทอร์น

3. นักศึกษากลุ่มที่ 2 สำหรับการหาค่าความเชื่อมั่นของแบบทดสอบ จำนวน 36 คน ได้มาโดยการสุ่มอย่างง่าย จากหลักสูตรการวิจัยและเทคโนโลยีการจัดการเรียนรู้ และหลักสูตรประกาศนียบัตรบัณฑิตวิชาชีพครู ชั้นปีที่ 2 คณะศึกษาศาสตร์ มหาวิทยาลัยฟาร์อีสเทอร์น ซึ่งเป็นคนละกลุ่มกับกลุ่มที่ 1 โดยใช้เกณฑ์จำนวนขั้นต่ำสำหรับการวิเคราะห์ความเชื่อมั่นของแบบทดสอบตามแนวคิดของ Ebel & Frisbie (1991) ที่ระบุว่าควรใช้กลุ่มตัวอย่างขั้นต่ำ 30 คน

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นแบบทดสอบวัดแววจความเป็นครู ที่มีลักษณะเป็นแบบทดสอบปรนัยเลือกตอบ 5 ตัวเลือก เชิงสถานการณ์ จำนวน 70 ข้อ

ขั้นตอนการสร้างเครื่องมือ

- 1) กำหนดจุดมุ่งหมายของการสร้างแบบทดสอบเพื่อวัดแววจความเป็นครู
- 2) ศึกษาแนวคิดเกี่ยวกับวิชาชีพครูและความเป็นครู แววจความเป็นครู องค์ประกอบของแววจความเป็นครู ทฤษฎีการวัดทางการศึกษา ทฤษฎีการสร้างและหาคุณภาพแบบทดสอบ การหาคุณภาพแบบทดสอบแบบอิงเกณฑ์และค่าอำนาจจำแนก
- 3) สังเคราะห์ความรู้แล้วกำหนดองค์ประกอบของแววจความเป็นครู และนิยามตัวแปร 5 ด้าน เพื่อนำไปเป็นกรอบในการสร้างข้อคำถามในแบบทดสอบได้ดังนี้



ด้านที่ 1 เจตคติและคุณธรรมความเป็นครู หมายถึง ความคิด ความรู้สึก และค่านิยมของบุคคลที่มีต่อวิชาชีพครู ผู้เรียน และสังคม ซึ่งสะท้อนผ่านความรักศรัทธาในวิชาชีพ ความเสียสละ ความเมตตา กรุณา ความยุติธรรม ความซื่อสัตย์สุจริต และการยึดมั่นในจรรยาบรรณวิชาชีพครู ตลอดจนความมุ่งมั่นในการส่งเสริมศักยภาพผู้เรียนอย่างเสมอภาค โดยแสดงออกผ่านพฤติกรรมที่สอดคล้องกับมาตรฐานวิชาชีพและจรรยาบรรณครูอย่างต่อเนื่อง

ด้านที่ 2 ความรับผิดชอบต่อวิชาชีพและการพัฒนาตนเอง หมายถึง ความตระหนักและความมุ่งมั่นของบุคคลในการปฏิบัติหน้าที่ครูอย่างเต็มความสามารถ มีวินัยในตนเอง รับผิดชอบต่อผลลัพธ์การเรียนรู้ของผู้เรียน และแสวงหาความรู้ใหม่ ๆ เพื่อพัฒนาสมรรถนะทางวิชาชีพอย่างต่อเนื่อง รวมทั้งเปิดรับการประเมินและสะท้อนคิด (Reflection) เพื่อนำไปสู่การปรับปรุงตนเอง

ด้านที่ 3 ความสามารถในการจัดการเรียนรู้ หมายถึง ศักยภาพของบุคคลในการออกแบบ วางแผน ดำเนินการ และประเมินผลการจัดการเรียนรู้ให้สอดคล้องกับหลักสูตร มาตรฐานการเรียนรู้ และความแตกต่างระหว่างบุคคลของผู้เรียน โดยใช้กลยุทธ์ วิธีสอน และสื่อการเรียนรู้ที่เหมาะสม เพื่อส่งเสริมการคิดวิเคราะห์ การแก้ปัญหา และการเรียนรู้ที่มีความหมาย

ด้านที่ 4 ทักษะการสื่อสารและปฏิสัมพันธ์ หมายถึง ความสามารถในการถ่ายทอดความรู้ ความคิด และความรู้สึกได้อย่างชัดเจน เหมาะสม และสร้างสรรค์ รวมทั้งการสร้างความสัมพันธ์เชิงบวกกับผู้เรียน เพื่อนร่วมงาน ผู้ปกครอง และชุมชน โดยอาศัยทักษะการฟังอย่างลึกซึ้ง การใช้ภาษาที่เหมาะสม การควบคุมอารมณ์ และการจัดการความขัดแย้งอย่างมืออาชีพ

ด้านที่ 5 บุคลิกภาพและความเป็นมืออาชีพ หมายถึง ลักษณะภายนอกและภายในของบุคคลที่สะท้อนความน่าเชื่อถือ ความมั่นคงทางอารมณ์ ความเหมาะสมตามบทบาทครู และความสามารถในการประพฤติตนเป็นแบบอย่างที่ดีแก่ผู้เรียนและสังคม โดยครอบคลุมการแต่งกาย กิริยามารยาท การควบคุมตนเอง ความมั่นใจในตนเอง และการวางตัวอย่างเหมาะสมตามกาลเทศะ

4) สร้างข้อคำถามเป็นแบบทดสอบปรนัยเลือกตอบ 5 ตัวเลือก เชิงสถานการณ์ ตามองค์ประกอบ 5 ด้าน ๆ ละ 25 ข้อ รวม 125 ข้อ ๆ ละ 1 คะแนน (ถูก 1 ผิด 0)

5) นำข้อคำถามตามองค์ประกอบ 5 ด้าน เสนอผู้เชี่ยวชาญ 3 ท่าน พิจารณาความเที่ยงตรงเชิงเนื้อหาตามนิยามตัวแปร ภาษา และรูปแบบการสร้างเครื่องมือ แล้วนำผลการพิจารณามาวิเคราะห์ค่าดัชนีความสอดคล้อง (IOC)

6) ปรับข้อสอบตามคำแนะนำของผู้เชี่ยวชาญ เช่น การปรับถ้อยคำ ภาษา ให้ชัดเจน การลดความกำกวมของตัวเลือก การปรับให้สอดคล้องกับนิยามเชิงปฏิบัติการ ความถูกต้องของเฉลย จากนั้นจัดทำแบบทดสอบ 125 ข้อ โดยใช้ Google Form นำไปทดลองใช้กับกลุ่มตัวอย่างนักศึกษา กลุ่มที่ 1 จำนวน 112 คน เพื่อนำข้อมูลมาวิเคราะห์ค่าอำนาจจำแนกแบบอิงเกณฑ์

7) นำผลการตอบมาวิเคราะห์ค่าอำนาจจำแนกแบบอิงเกณฑ์ โดยใช้โปรแกรม RTAP (Research Tools Analysis Program) (ศูนย์บริการวิชาการและเผยแพร่ นวัตกรรมทางการศึกษา และภาคีวิจัย และพัฒนาการศึกษา, 2565) กำหนดเกณฑ์การผ่านร้อยละ 60 (42 คะแนน) ให้สอดคล้องกับเกณฑ์การประเมินสมรรถนะด้านการปฏิบัติงานและการปฏิบัติตนของครูสภา พบว่า ข้อสอบ 125 ข้อ มีค่าอำนาจจำแนกอยู่ระหว่าง 0.14 - 1.00 จากนั้นคัดเลือกข้อสอบที่ผ่านเกณฑ์ค่าอำนาจจำแนกตั้งแต่ 0.40 ขึ้นไป ได้ข้อสอบที่ผ่านเกณฑ์ จำนวน 70 ข้อ ตามองค์ประกอบ 5 ด้านดังตารางที่ 1



ตารางที่ 1 Test Blueprint

องค์ประกอบ	จำนวนข้อ ก่อนตัด	จำนวนข้อ หลังตัด	ข้อที่คัดเลือกได้
ด้านที่ 1	25	9	2, 9, 10, 11, 12, 13, 14, 15, 23
ด้านที่ 2	25	11	28, 29, 31, 34, 36, 38, 39, 40, 44, 47, 49
ด้านที่ 3	25	18	52, 53, 54, 56, 57, 59, 60, 61, 62, 63, 64, 65, 66, 67, 71, 72, 74, 75
ด้านที่ 4	25	19	76, 77, 78, 79, 81, 82, 83, 84, 85, 87, 89, 90, 91, 92, 93, 94, 95, 96, 98
ด้านที่ 5	25	13	101, 102, 103, 106, 108, 111, 112, 114, 116, 120, 121, 122, 123
รวม	125	70	

8) นำข้อสอบที่ผ่านเกณฑ์การคัดเลือก จำนวน 70 ข้อ จัดทำฉบับสมบูรณ์แล้วนำไปทดสอบกับ นักศึกษากลุ่มที่ 2 จำนวน 36 คน เพื่อหาค่าความเชื่อมั่นทั้งฉบับตามวิธีของโลเวท (Lovett Method) โดยใช้โปรแกรม RTAP (Research Tools Analysis Program) (ศูนย์บริการวิชาการและเผยแพร่ นวัตกรรม ทางการศึกษา และภาควิชาวิจัยและพัฒนาการศึกษา, 2565)

การเก็บรวบรวมข้อมูล

ในการเก็บรวบรวมข้อมูลวิจัยได้ดำเนินการตามขั้นตอนดังนี้

- 1) ข้อมูลความเที่ยงตรงเชิงเนื้อหา เก็บข้อมูลจากผู้เชี่ยวชาญ 3 ท่าน โดยใช้แบบประเมินความ เที่ยงตรงของแบบทดสอบวัดแววจความเป็นครูสำหรับผู้เชี่ยวชาญ โดยจัดทำเป็นจดหมายอิเล็กทรอนิกส์
- 2) ข้อมูลค่าอำนาจจำแนก เก็บข้อมูลจากนักศึกษากลุ่มที่ 1 จำนวน 112 คน โดยใช้ Google Form จัดทำเป็นแบบทดสอบออนไลน์ จำนวน 125 ข้อ ดำเนินการเมื่อสิ้นภาคเรียนที่ 1 ปีการศึกษา 2568
- 3) ข้อมูลความเชื่อมั่น เก็บข้อมูลจากนักศึกษากลุ่มที่ 2 จำนวน 36 คน โดยใช้ Google Form จัดทำเป็นแบบทดสอบออนไลน์ จำนวน 70 ข้อ ดำเนินการเมื่อสิ้นภาคเรียนที่ 2 ปีการศึกษา 2568

การวิเคราะห์ข้อมูล

- 1) ข้อมูลความเที่ยงตรงเชิงเนื้อหา ใช้การวิเคราะห์ค่าดัชนีความสอดคล้องของผู้เชี่ยวชาญ (IOC)
- 2) ข้อมูลค่าอำนาจจำแนก ใช้การวิเคราะห์ค่าดัชนีบี (B-Index หรือ Brennan Index) โดยใช้โปรแกรม RTAP (ศูนย์บริการวิชาการและเผยแพร่ นวัตกรรมทางการศึกษา และภาควิชาวิจัยและพัฒนาการศึกษา, 2565)
- 3) ข้อมูลความเชื่อมั่น ใช้การวิเคราะห์ของโลเวท (Lovett Method) (คณาจารย์ภาควิชาวิจัยและพัฒนา การศึกษา, 2558) โดยใช้โปรแกรม RTAP (ศูนย์บริการวิชาการและเผยแพร่ นวัตกรรมทางการศึกษาและ ภาควิชาวิจัยและพัฒนาการศึกษา, 2565)

ผลการวิจัย

- 1) คุณภาพของแบบทดสอบวัดแววจความเป็นครูในด้านความเที่ยงตรงเชิงเนื้อหา ผลการวิเคราะห์ความเที่ยงตรงเชิงเนื้อหาของแบบทดสอบวัดแววจความเป็นครู จากผู้เชี่ยวชาญ จำนวน 3 ท่าน แสดงดังตารางที่ 2



ตารางที่ 2 ผลการวิเคราะห์ค่าดัชนีความสอดคล้องของผู้เชี่ยวชาญ (IOC) ของแบบทดสอบวัดแววจิต
ความเป็นครู

องค์ประกอบ	จำนวน ข้อสอบ	ค่า IOC	แปลผล
ด้านที่ 1 เจตคติและคุณธรรมความเป็นครู	25	0.67-1.00	นำไปใช้ได้ทุกข้อ
ด้านที่ 2 ความรับผิดชอบต่อวิชาชีพและการพัฒนาตนเอง	25	0.67-1.00	นำไปใช้ได้ทุกข้อ
ด้านที่ 3 ความสามารถในการจัดการเรียนรู้	25	0.67-1.00	นำไปใช้ได้ทุกข้อ
ด้านที่ 4 ทักษะการสื่อสารและปฏิสัมพันธ์	25	0.67-1.00	นำไปใช้ได้ทุกข้อ
ด้านที่ 5 บุคลิกภาพและความเป็นมืออาชีพ	25	0.67-1.00	นำไปใช้ได้ทุกข้อ
รวม	125	0.67-1.00	นำไปใช้ได้ทุกข้อ

จากตารางที่ 2 ผลการวิเคราะห์ค่าดัชนีความสอดคล้องของผู้เชี่ยวชาญ (IOC) ของแบบทดสอบวัดแววจิตความเป็นครู จำนวน 125 ข้อ มีค่าตั้งแต่ 0.67 – 1.00 สามารถนำไปใช้ได้ทุกข้อ

2) คุณภาพของแบบทดสอบวัดแววจิตความเป็นครูในด้านค่าอำนาจจำแนก

ผลการวิเคราะห์ค่าอำนาจจำแนกของแบบทดสอบวัดแววจิตความเป็นครู จำนวน 70 ข้อ จากการทดลองใช้กับนักศึกษาในกลุ่มที่ 2 แสดงดังตารางที่ 3

ตารางที่ 3 ผลการวิเคราะห์ค่าอำนาจจำแนกของแบบทดสอบวัดแววจิตความเป็นครู จำแนกตามองค์ประกอบ 5 ด้าน

ด้าน/ข้อที่	B-Index	แปลผล	ด้าน/ข้อที่	B-Index	แปลผล
ด้านที่ 1 (9 ข้อ)			ด้านที่ 3 (ต่อ)		
ข้อที่ 1	0.97	อำนาจจำแนกสูง	ข้อที่ 36	0.97	อำนาจจำแนกสูง
ข้อที่ 2	0.97	อำนาจจำแนกสูง	ข้อที่ 37	1.00	อำนาจจำแนกสูง
ข้อที่ 3	0.86	อำนาจจำแนกสูง	ข้อที่ 38	0.89	อำนาจจำแนกสูง
ข้อที่ 4	0.92	อำนาจจำแนกสูง	ด้านที่ 4 (19 ข้อ)		
ข้อที่ 5	1.00	อำนาจจำแนกสูง	ข้อที่ 39	1.00	อำนาจจำแนกสูง
ข้อที่ 6	0.89	อำนาจจำแนกสูง	ข้อที่ 40	0.97	อำนาจจำแนกสูง
ข้อที่ 7	0.83	อำนาจจำแนกสูง	ข้อที่ 41	0.97	อำนาจจำแนกสูง
ข้อที่ 8	0.94	อำนาจจำแนกสูง	ข้อที่ 42	1.00	อำนาจจำแนกสูง
ข้อที่ 9	1.00	อำนาจจำแนกสูง	ข้อที่ 43	0.97	อำนาจจำแนกสูง
ด้านที่ 2 (11 ข้อ)			ข้อที่ 44	0.94	อำนาจจำแนกสูง
ข้อที่ 10	0.92	อำนาจจำแนกสูง	ข้อที่ 45	0.92	อำนาจจำแนกสูง
ข้อที่ 11	0.89	อำนาจจำแนกสูง	ข้อที่ 46	0.86	อำนาจจำแนกสูง
ข้อที่ 12	1.00	อำนาจจำแนกสูง	ข้อที่ 47	0.92	อำนาจจำแนกสูง
ข้อที่ 13	1.00	อำนาจจำแนกสูง	ข้อที่ 48	0.97	อำนาจจำแนกสูง
ข้อที่ 14	1.00	อำนาจจำแนกสูง	ข้อที่ 49	0.94	อำนาจจำแนกสูง



ตารางที่ 3 (ต่อ)

ด้าน/ข้อที่	B-Index	แปลผล	ด้าน/ข้อที่	B-Index	แปลผล
ข้อที่ 15	0.97	อำนาจจำแนกสูง	ข้อที่ 50	0.86	อำนาจจำแนกสูง
ข้อที่ 16	0.86	อำนาจจำแนกสูง	ข้อที่ 51	0.50	อำนาจจำแนก ค่อนข้างสูง
ข้อที่ 17	0.97	อำนาจจำแนกสูง	ข้อที่ 52	0.22	อำนาจจำแนก ปานกลาง
ข้อที่ 18	1.00	อำนาจจำแนกสูง	ข้อที่ 53	0.78	อำนาจจำแนกสูง
ข้อที่ 19	0.86	อำนาจจำแนกสูง	ข้อที่ 54	0.94	อำนาจจำแนกสูง
ข้อที่ 20	1.00	อำนาจจำแนกสูง	ข้อที่ 55	0.83	อำนาจจำแนกสูง
ด้านที่ 3 (18 ข้อ)			ข้อที่ 56	0.72	อำนาจจำแนกสูง
ข้อที่ 21	1.00	อำนาจจำแนกสูง	ข้อที่ 57	0.50	อำนาจจำแนก ค่อนข้างสูง
ข้อที่ 22	0.97	อำนาจจำแนกสูง	ด้านที่ 5 (13 ข้อ)		
ข้อที่ 23	0.92	อำนาจจำแนกสูง	ข้อที่ 58	0.94	อำนาจจำแนกสูง
ข้อที่ 24	0.94	อำนาจจำแนกสูง	ข้อที่ 59	0.97	อำนาจจำแนกสูง
ข้อที่ 25	0.97	อำนาจจำแนกสูง	ข้อที่ 60	0.78	อำนาจจำแนกสูง
ข้อที่ 26	0.97	อำนาจจำแนกสูง	ข้อที่ 61	1.00	อำนาจจำแนกสูง
ข้อที่ 27	0.92	อำนาจจำแนกสูง	ข้อที่ 62	1.00	อำนาจจำแนกสูง
ข้อที่ 28	1.00	อำนาจจำแนกสูง	ข้อที่ 63	0.97	อำนาจจำแนกสูง
ด้าน/ข้อที่	B-Index	แปลผล	ด้าน/ข้อที่	B-Index	แปลผล
ข้อที่ 29	0.94	อำนาจจำแนกสูง	ข้อที่ 64	0.97	อำนาจจำแนกสูง
ข้อที่ 30	0.78	อำนาจจำแนกสูง	ข้อที่ 65	1.00	อำนาจจำแนกสูง
ข้อที่ 31	0.44	อำนาจจำแนก ค่อนข้างสูง	ข้อที่ 66	0.89	อำนาจจำแนกสูง
ข้อที่ 32	0.92	อำนาจจำแนกสูง	ข้อที่ 67	0.81	อำนาจจำแนกสูง
ข้อที่ 33	0.89	อำนาจจำแนกสูง	ข้อที่ 68	1.00	อำนาจจำแนกสูง
ข้อที่ 34	0.97	อำนาจจำแนกสูง	ข้อที่ 69	0.94	อำนาจจำแนกสูง
ข้อที่ 35	0.94	อำนาจจำแนกสูง	ข้อที่ 70	1.00	อำนาจจำแนกสูง

จากตารางที่ 3 ผลการวิเคราะห์ค่าอำนาจจำแนกของแบบทดสอบวัดแววจิตใจเป็นครู จำแนกตามองค์ประกอบ 5 ด้าน จำนวน 70 ข้อ พบว่าข้อสอบข้อที่ 52 มีค่าอำนาจจำแนกปานกลางเท่ากับ 0.22 ข้อสอบข้อที่ 31, 51 และ 57 มีค่าอำนาจจำแนกค่อนข้างสูง เท่ากับ 0.44, 0.50, 0.50 ตามลำดับ ข้อสอบข้อที่ 66 มีค่าอำนาจจำแนกสูง อยู่ระหว่าง 0.72 ถึง 1.00

3) คุณภาพของแบบทดสอบวัดแววจิตใจเป็นครูในด้านค่าความเชื่อมั่น

ผลการวิเคราะห์ค่าความเชื่อมั่นของแบบทดสอบวัดแววจิตใจเป็นครูทั้งฉบับ โดยใช้การวิเคราะห์ของโลเวท (Lovett Method) แสดงดังตารางที่ 4



ตารางที่ 4 ผลการวิเคราะห์ค่าความเชื่อมั่นของแบบทดสอบวัดแววจความเป็นครู

จำนวนข้อสอบทั้งหมด (k)	70 ข้อ
จำนวนผู้ตอบ (n)	36 คน
คะแนนรวมของผู้ตอบทุกคน (X_i)	2283 คะแนน
คะแนนเกณฑ์ (C)	42 คะแนน
ค่าความเชื่อมั่น (r_{cc})	0.99

จากตารางที่ 4 ผลการวิเคราะห์ค่าความเชื่อมั่นของแบบทดสอบวัดแววจความเป็นครูทั้งหมด โดยใช้การวิเคราะห์ของโลเวท (Lovett Method) มีค่าเท่ากับ 0.99

สรุปและอภิปรายผล

ประเด็นที่ 1 ความเที่ยงตรงเชิงเนื้อหาของแบบทดสอบ

ผลการวิจัยพบว่า แบบทดสอบวัดแววจความเป็นครูมีค่าดัชนีความสอดคล้องของผู้เชี่ยวชาญ (IOC) อยู่ระหว่าง 0.67–1.00 และสามารถนำไปใช้ได้ทุกข้อ แสดงให้เห็นว่าแบบทดสอบมีความเที่ยงตรงเชิงเนื้อหาในระดับสูง ผลดังกล่าวสอดคล้องกับแนวคิดเรื่องความเที่ยงตรง (Validity) ซึ่งมองว่าเครื่องมือวัดที่ดีต้องสามารถวัดคุณลักษณะได้ตรงตามจุดมุ่งหมายที่กำหนด (AERA, APA, & NCME, 2014) โดยเฉพาะความเที่ยงตรงเชิงเนื้อหา ซึ่งเน้นความสอดคล้องระหว่างข้อคำถามกับนิยามเชิงทฤษฎีและขอบเขตของโครงสร้างที่ต้องการวัด (Linn & Gronlund, 2000) การใช้ผู้ทรงคุณวุฒิพิจารณาความสอดคล้องของข้อคำถามจึงเป็นกระบวนการสำคัญในการรวบรวมหลักฐานสนับสนุนความตรงของเครื่องมือ นอกจากนี้ Messick (1995) เสนอว่าความเที่ยงตรงควรพิจารณาในฐานะ “หลักฐานสนับสนุนการตีความคะแนน” (Validity as Unified Concept) ซึ่งรวมถึงความสอดคล้องทางเนื้อหา โครงสร้าง และผลกระทบจากการใช้เครื่องมือ ดังนั้น การที่แบบทดสอบฉบับนี้ผ่านการพิจารณาจากผู้เชี่ยวชาญอย่างเป็นระบบ จึงเป็นหลักฐานเชิงประจักษ์ที่สนับสนุนความเหมาะสมของการตีความคะแนนแววจความเป็นครู

ประเด็นที่ 2 ค่าอำนาจจำแนกแบบอิงเกณฑ์ของข้อคำถาม

จากผลการวิจัยที่พบว่า ข้อสอบข้อที่ 52 มีค่าอำนาจจำแนกเท่ากับ 0.22 เมื่อพิจารณาผลการตอบของกลุ่มตัวอย่าง พบว่า ตอบกระจายทุกตัวเลือก จึงทำให้ผลการวิเคราะห์ค่าอำนาจจำแนกอยู่ในระดับปานกลาง (0.22) ตัวอย่างข้อสอบข้อที่ 52

คำถาม : ระหว่างกิจกรรมกลุ่ม นักเรียนคนหนึ่งถูกเพื่อนไม่เลือกเข้ากลุ่ม คุณควร

- A) ให้คะแนนกลุ่มลดลง
- B) ขอให้กลุ่มแบ่งงานให้ชัด
- C) สังเกตปฏิกริยานักเรียนคนนั้น
- D) บอกให้นักเรียนในกลุ่มว่าควรมีน้ำใจ
- E) แยกนักเรียนคนนั้นออกไปทำงานคนเดียว

เฉลย: B

ส่วนข้อสอบอีก 66 ข้อ มีค่าอำนาจจำแนก (B-Index) อยู่ในระดับสูง (0.72–1.00) ขณะที่ 3 ข้ออยู่ในระดับค่อนข้างสูง (0.44, 0.50 และ 0.50) แสดงว่าแบบทดสอบสามารถจำแนกผู้ที่ผ่านเกณฑ์และไม่ผ่านเกณฑ์ได้อย่างมีประสิทธิภาพ ผลดังกล่าวสอดคล้องกับแนวคิดของแบบทดสอบแบบอิงเกณฑ์ (Criterion-Referenced Testing) ซึ่งมุ่งเน้นความแม่นยำในการจำแนกสถานะของผู้สอบ (Masters–Nonmasters) มากกว่าการเปรียบเทียบเชิงลำดับ (Popham, 2014) โดย Brennan (2006) อธิบายว่า คุณภาพของข้อคำถามในแบบทดสอบแบบอิงเกณฑ์ควรพิจารณาจากความสามารถในการสนับสนุนการตัดสินใจเชิงเกณฑ์อย่างถูกต้อง ดัชนีความแตกต่างระหว่างกลุ่มผู้รอบรู้ – ไม่รอบรู้ (Difference Index หรือ B-Index) เป็นตัวชี้วัดที่สะท้อน



ประสิทธิภาพของข้อคำถามในการแยกกลุ่มตามเกณฑ์ (Nitko & Brookhart, 2011) ค่าอำนาจจำแนกที่สูงมากในงานวิจัยครั้งนี้จึงสะท้อนถึงความสอดคล้องของข้อคำถามกับโครงสร้างแนวความเป็นครู และความชัดเจนของเกณฑ์การตัดสิน (Cut Score) อย่างไรก็ตาม ข้อที่มีค่าอำนาจจำแนกต่ำกว่าข้ออื่นอาจบอกถึงความกำกวมของเนื้อหา ซึ่งควรได้รับการปรับปรุงตามหลักการวิเคราะห์ข้อสอบ (Item Analysis) เพื่อเพิ่มประสิทธิภาพในการจำแนก (Linn & Gronlund, 2000)

ประเด็นที่ 3 ความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์

ผลการวิเคราะห์ค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับโดยใช้วิธีของ Lovett ได้ค่า r_{cc} เท่ากับ 0.99 ผลการวิจัยจัดอยู่ในระดับสูงมาก ในบริบทของแบบทดสอบแบบอิงเกณฑ์ ความเชื่อมั่นไม่ได้มุ่งเพียงความคงเส้นคงวาของคะแนนรวมเท่านั้น แต่เน้นความสม่ำเสมอของการจำแนกผู้สอบตามเกณฑ์ (Classification Consistency) (Brennan, 2006) วิธีของ Lovett จึงเหมาะสมกับแบบทดสอบที่มีการกำหนดคะแนนเกณฑ์ชัดเจน ค่าความเชื่อมั่นที่สูงมากในครั้งนี้นี้ สะท้อนว่าแบบทดสอบมีความเสถียรและมีความคลาดเคลื่อนจากการวัดต่ำ ซึ่งสอดคล้องกับแนวคิดของทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) ที่ระบุว่าคะแนนที่สังเกตได้ประกอบด้วยคะแนนจริงและความคลาดเคลื่อน ($X = T + E$) และเครื่องมือที่ดีควรลดค่าความคลาดเคลื่อนให้ต่ำที่สุด (Crocker & Algina, 1986) อย่างไรก็ตาม แนวคิดของ Messick (1995) การพิจารณาคุณภาพเครื่องมือควรมองอย่างบูรณาการ ทั้งด้านความเที่ยงตรง ความเชื่อมั่น และผลกระทบจากการใช้เครื่องมือ ดังนั้นแม้เครื่องมือจะมีค่าความเชื่อมั่นสูงมาก แต่การนำไปใช้ตัดสินเชิงวิชาชีพควรใช้ข้อมูลประกอบหลายแหล่งเพื่อความรอบคอบ

ข้อเสนอแนะ

ข้อเสนอแนะในการนำไปใช้

1. ควรนำแบบทดสอบไปใช้เป็นเครื่องมือคัดกรองเบื้องต้นในการประเมินแนวความเป็นครูของนักศึกษาครูหรือผู้สมัครเข้าศึกษาในหลักสูตรครุศาสตร์/ศึกษาศาสตร์ และ/หรือใช้ควบคู่กับเครื่องมือประเมินรูปแบบอื่น เช่น การสัมภาษณ์ การสังเกตพฤติกรรม หรือแฟ้มสะสมงาน เพื่อให้การตัดสินมีความรอบด้านและลดความคลาดเคลื่อนจากการวัดเพียงรูปแบบเดียว
2. ในการนำไปใช้จริง ควรกำหนดเกณฑ์การตัดสิน (Cut Score) อย่างชัดเจนและอิงจากข้อมูลเชิงประจักษ์ รวมถึงควรมีคู่มือการใช้แบบทดสอบเพื่อสร้างความเข้าใจตรงกันของผู้ประเมิน

ข้อเสนอแนะในการวิจัยต่อไป

1. ควรศึกษาความเที่ยงตรงเชิงโครงสร้าง (Construct Validity) โดยใช้การวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) หรือเชิงยืนยัน (CFA) เพื่อยืนยันโครงสร้าง 5 ด้านของแนวความเป็นครู
2. ควรศึกษาความตรงเชิงเกณฑ์สัมพันธ์ (Criterion-Related Validity) เช่น ความสัมพันธ์ระหว่างคะแนนแนวความเป็นครูกับผลการฝึกประสบการณ์วิชาชีพครู

เอกสารอ้างอิง

- คุรุสภา. (2562). มาตรฐานวิชาชีพครู. กรุงเทพฯ: คุรุสภา.
- บุญชม ศรีสะอาด. (2560 ก). การวัดและประเมินผลการศึกษา. กรุงเทพฯ: สุวีริยาสาส์น.
- บุญชม ศรีสะอาด. (2560 ข). การวิจัยเบื้องต้น. กรุงเทพฯ: สุวีริยาสาส์น.
- ล้วน สายยศ และอังคณา สายยศ. (2560). เทคนิคการวัดผลทางการศึกษา. กรุงเทพฯ: สุวีริยาสาส์น.
- ศูนย์บริการวิชาการและเผยแพร่นวัตกรรมทางการศึกษา และภาควิชาวิจัยและพัฒนาการศึกษา. (2565). โปรแกรมวิเคราะห์คุณภาพเครื่องมือที่ใช้ในงานวิจัย (Research Tools Analysis Program: RTAP).
- มหาสารคาม: คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม. จาก <http://edurtap.msu.ac.th/>



rtapapp/

- สมหวัง พิธิยานุวัฒน์. (2558). การวัดผลการศึกษา: แนวคิด ทฤษฎี และการปฏิบัติ. กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สิริมงคล สุวรรณผา. (2566). การพัฒนาตัวชี้วัดคุณลักษณะความเป็นครูมืออาชีพของบัณฑิตใหม่ คณะศึกษาศาสตร์ มหาวิทยาลัยรามคำแหง. วารสารสหวิทยาการวิจัยและวิชาการ, 3(5), 1033–1054.
- อนุวัติ คุ้มแก้ว. (2563). การศึกษาคูณลักษณะของครูไทยในศตวรรษที่ 21. สักทอง: วารสารมนุษยศาสตร์และสังคมศาสตร์ (สทมส.), 26(4), 43–55.
- เอกภูมิ จันทร์ขันธ์. (2557). การพัฒนาทัศนคติเกี่ยวกับคุณลักษณะความเป็นครูของนิสิตครุวิทยาการศึกษาศาสตร์. วารสารมหาวิทยาลัยศิลปากร ฉบับภาษาไทย, 34(1), 161–179.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. Washington, DC: AERA.
- Brennan, R. L. (1972). A generalized upper-lower item discrimination index. Educational and Psychological Measurement, 32, 289–303.
- Brennan, R. L. (2006). Educational measurement (4th ed.). Westport, CT: Praeger.
- Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. New York, NY: Holt, Rinehart and Winston.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297–334.
- Ebel, R. L., & Frisbie, D. A. (1991). Essentials of educational measurement (5th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Glaser, R. (1963). Instructional technology and the measurement of learning outcomes. American Psychologist, 18, 519–521.
- Hambleton, R. K., & Novick, M. R. (1973). Toward an integration of theory and method for criterion-referenced tests. Journal of Educational Measurement, 10, 159–170.
- Hambleton, R. K., & Pitoniak, M. J. (2006). Setting performance standards. In R. L. Brennan (Ed.), Educational measurement (4th ed.). Westport, CT: Praeger.
- Linn, R. L., & Gronlund, N. E. (2000). Measurement and assessment in teaching (8th ed.). Upper Saddle River, NJ: Prentice Hall.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances. American Psychologist, 50(9), 741–749.
- Nitko, A. J., & Brookhart, S. M. (2011). Educational assessment of students (6th ed.). Boston, MA: Pearson.
- Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). New York, NY: McGraw-Hill.
- Popham, W. J. (2014). Classroom assessment: What teachers need to know (7th ed.). Boston, MA: Pearson.