

Fiabilidad y validez de la escala de competencias docentes de educación media superior

Reliability and validity of the upper secondary education teaching competencies scale

Dr. Emerson LÓPEZ-LÓPEZ. Profesor investigador. Licenciatura y posgrado en Educación, Universidad Linda Vista, México (emerson.lopez@ulv.edu.mx).

Dr. Luis Gibran JUÁREZ-HERNÁNDEZ. Profesor investigador. Facultad de Medicina y Ciencias Biomédicas, Universidad Autónoma de Chihuahua y Centro Universitario CIFE, México (gibbjuarez@gmail.com).

Resumen:

Se analizaron las propiedades psicométricas de validez de contenido, constructo y confiabilidad de un instrumento para evaluar las competencias docentes en la educación media superior. Para la validez de dicho contenido, el instrumento se sometió al juicio de 21 jueces y se calculó el coeficiente de validez de contenido V de Aiken. La confiabilidad fue evaluada con el Omega de McDonald. Para el análisis de la validez de constructo, el instrumento se aplicó a 3726 estudiantes de la educación media superior. En el análisis de validez del constructo, se empleó un proceso de validación cruzada que involucró el análisis factorial exploratorio (AFE) y el análisis factorial confirmatorio (AFC). El instrumento integró de manera simplificada las ocho competencias establecidas por el Gobierno de México y en total se conformó de 20 ítems. Los ítems fueron validados en contenido (V de Aiken > 0.75, VI > 0.50). La confiabilidad fue óptima (Omega de McDonald: 0.959 IC al 95 %: 0.957 ± 0.961). Por su parte, en el AFE se verificó la correspondencia del instrumento con el modelo teórico, ya que únicamente denotó un factor que explicó 60 % de varianza, en el cual se encontraron representados 18 de los 20 ítems. La evaluación de este modelo mediante el análisis factorial confirmatorio reveló un ajuste óptimo (razón χ^2/gl : 1.89; GFI: 0.995; RMSAE: 0.050; RMR: 0.028; CFI: 0.992; TLI: 0.994). El instrumento «ECDEMS» posee adecuadas propiedades psicométricas.

Palabras clave: análisis factorial, docencia, educación basada en las competencias, estudiante, evaluación, instrumento de medida.

Abstract:

The psychometric properties of content validity were analyzed, construct and reliability of an instrument to assess teaching competencies in upper secondary education. For content validity, the instrument was submitted to the judgment of 21 judges consequently the Aiken

Fecha de recepción del original: 01/05/2025.

Fecha de aprobación: 01/06/2025.

Cómo citar este artículo: López-López, E. y Juárez-Hernández, L. G. (2025). Fiabilidad y validez de la escala de competencias docentes de educación media superior [Reliability and validity of the upper secondary education teaching competencies scale]. *Revista Española de Pedagogía*, 83(292), 687-708 <https://doi.org/10.9781/rep.2025.384>

content validity coefficient V was calculated. For the construct validity analysis, the instrument was implemented to 3726 students of upper secondary education. Reliability was evaluated with McDonald's Omega. In the construct validity analysis, a cross-validation process was used that involved exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). The instrument integrated in a simplified manner the eight competencies established by the Government of Mexico and in total consisted of 20 items. The items were validated in content (V of Aiken > 0.75 , $LV > 0.50$). Reliability was optimal (McDonald's Omega: 0.959 95% CI: 0.957 ± 0.961). Moreover, the AFE verified the correspondence of the instrument with the theoretical model, considering that it denoted a factor that explained 60% of variance, in which 18 of the 20 items were represented. The evaluation of this model by confirmatory factor analysis revealed an optimal fit (χ^2/df ratio: 1.89 ; GFI: 0.995 ; RMSAE: 0.050 ; RMR: 0.028 ; CFI: 0.992 ; TLI: 0.994). The instrument «ECDEMS» demonstrates adequate psychometric properties.

Keywords: factor analysis, teaching, competency-based education, student, assessment, measurement instrument.

1. Introducción

El término competencias usado primero por McClelland (1973) se definió como una capacidad que incluye la inteligencia, las aptitudes, la motivación, los rasgos de la personalidad y los comportamientos, que hacen eficiente el desempeño laboral humano y el éxito en la vida. Posteriormente, Boyatzis (1983) elaboró un enfoque sistémico-empírico para evaluar y desarrollar el talento directivo en las organizaciones, y, más adelante, Boyatzis (2006) las categorizó en cognitivas, emocionales y conductuales, al comparar los atributos de empleados de alto rendimiento con los de rendimiento promedio. Aunque existen avances terminológicos, el concepto sigue siendo multifacético, complejo, debatible, sin consenso único.

Entre otras, existen dos tendencias del término dependiendo del contexto en el que se utiliza. La primera se orienta hacia la evaluación de la tarea laboral; y la segunda, en los comportamientos, actitudes y desempeño de las personas (Wong, 2020). Pero en sentido amplio, las competencias se consideran capacidades, atributos, conocimientos, habilidades, auto-conceptos, valores, rasgos personales, motivaciones, actitudes, o factores entrenables, observables y medibles que permiten desarrollar un trabajo complejo, específico y de manera sobresaliente (Arribas *et al.*, 2024; Zumstein-Shaha y Grace, 2023; WHO, 2022).

Algo similar ocurre con el concepto de competencias docentes, ya que poseen una amplia variedad de definiciones, de las que se pueden resumir como saberes, habilidades, capacidades, valores, actitudes y experiencias didáctico-pedagógicas, las cuales permiten a los profesores crear entornos de aprendizaje eficaces, eficientes, excelentes y productivos; en contextos específicos y demandas complejas, buscando desarrollar el potencial y la formación del estudiante, acorde a los objetivos educacionales establecidos por la institución (Burbano *et al.*, 2024; Otto, 2024; Singh, 2024; Uysal *et al.*, 2021; Zhou *et al.*, 2023).

Ante la polisemia del término competencias docentes, en este artículo se utiliza la definición del Gobierno mexicano establecida para la Educación Media Superior (EMS), expresada como «las que formulan las cualidades individuales, de carácter ético, académico, profesional y social que debe reunir el docente» (DOF, 2008c, p. 2). Estas están caracterizadas por ocho competencias personalizadas y están contextualizadas dentro del Sistema Nacional de Bachillerato: formación profesional continua, dominio de saberes para lograr aprendizajes significativos, planificación con enfoque por competencias y contextualizados, practicar el proceso enseñanza-aprendizaje efectivo, creativo e innovador; evaluar de manera formativa, construir aprendizajes autónomos y colaborativos, promover el desarrollo sano e integral y participar en la mejora continua de la escuela y en la gestión institucional (pp. 2-3), las cuales

se alinean con las competencias genéricas (DOF, 2008b, pp. 2-4), disciplinares (DOF, 2008b, pp. 4-7), profesionales (DOF, 2008b, pp. 7-8) y extendidas (DOF, 2008b, pp. 6-7) de los estudiantes (DOF, 2008b; DOF, 2009), así como las competencias requeridas para los directores de los planteles (DOF, 2008d, pp. 2-4).

Dada la relevancia nacional e internacional, las competencias docentes son importantes porque forman parte de la profesionalización del profesorado (Agüero-Servín, 2022), quienes son los actores clave de los sistemas educativos (Domínguez-González y Serna-Poot, 2021) y responsables de desarrollar con eficiencia el proceso enseñanza-aprendizaje (Granada *et al.*, 2024). Por el lado estudiantil, las competencias de los docentes ayudan a elevar el logro educativo de los alumnos como lo avalan algunos estudios (Di Lisio *et al.*, 2025; Firda y Khai-rat, 2023; Putra y Yanto, 2025), continuar con los estudios superiores (Hollenstein y Brühwiler, 2024), obtener mejores empleos (González y Estrella, 2023), y contribuir al desarrollo nacional-global (DOF, 2008a, p. 2).

Por otra parte, también es fundamental evaluar las competencias docentes debido a que permiten diagnosticar la capacidad del educador para facilitar los conocimientos (Brown, 2024), identificar sus fortalezas y oportunidades para mejorar la calidad de la enseñanza (Skedsmo y Huber, 2024), incidir en la mejora de su desarrollo profesional (Perrenoud, 2004; Wang y Sang, 2024), rendir cuentas a la sociedad (Bleiberg *et al.*, 2024), a los tomadores de decisiones (Hunter y Kho, 2023) y a los responsables de las políticas educativas (Castro-Castillo *et al.*, 2024).

Desde el lado estudiantil, al evaluar las competencias de los docentes se pueden obtener, entre otros, indicios del grado de desarrollo de las competencias de los estudiantes, las cuales resultan de aquellas (Nessa *et al.*, 2024), así como de los niveles de aprendizaje adquiridos (Sánchez-Tarazaga y Ferrández-Berruero, 2022). Sin embargo, es imperativo poseer instrumentos adecuados para estos propósitos.

Ante la importancia, la complejidad y los desafíos presentados en torno a la evaluación de las competencias docentes en EMS, existen diversos instrumentos internacionales (Aydin *et al.*, 2024; Baena-Extremera *et al.*, 2015; Cabero-Almenara y Palacios-Rodríguez, 2020; Cinque Gómez-del-Pulgar y Rodríguez-Mantilla, 2020; Chee Yuet *et al.*, 2016; Cortes *et al.*, 2020; Fernández-Díaz *et al.*, 2016; Gümüş y Kulkul, 2023; Kim y Kim, 2016; Salihu, 2019; Sánchez-Tarazaga y Fernández-Berruero, 2022), así como para Latinoamérica (Hernández-Suárez *et al.*, 2021; Pérez-Contreras *et al.*, 2022) y México (Abarca-Cedeño *et al.*, 2024; Luna-Serrano y Reyes-Piñuelos, 2015; Morán *et al.*, 2015; Ramón-Santiago *et al.*, 2017); aunque predominan los instrumentos autoadministrados por el docente; y son escasos los que responden los alumnos como los de Baena-Extremera *et al.* (2015), Luna-Serrano y Reyes-Piñuelos (2015) y Ramón-Santiago *et al.* (2017).

Al observar las bases teóricas de dichos instrumentos, la mayoría: a) carece de fundamentación apropiada del constructo competencias docentes en EMS; b) no se circunscriben a un determinado enfoque, c) las dimensiones establecidas son amplias y los ítems poco representativos, d) poco se identifica el modelo teórico que subyace al instrumento, y, finalmente, e) manifiestan dimensiones y mediciones de comportamientos poco relevantes, presentando infra y sobrerepresentación de algunos dominios.

Ante estas limitaciones, conviene señalar que la construcción de un instrumento demanda la existencia de una adecuada justificación, una delimitación precisa de la variable de estudio, una correcta definición operativa, semántica y sintáctica de las dimensiones relevantes y de los comportamientos pertinentes de la misma; y una interrelación entre dichos elementos e ítems (Muñiz y Fonseca-Pedrero, 2019). Asimismo, es fundamental la determinación específica del contexto de estudio, la población objetivo, las circunstancias, la forma de aplicación, el uso del instrumento, además de los análisis de contenido y constructo apropiados (AERA *et al.*, 2014).

Por lo declarado, se decidió realizar el presente estudio con los siguientes propósitos: 1) Diseñar una escala para evaluar las competencias docentes de la educación media superior, 2) Determinar la validez de contenido del instrumento que evalúe las competencias individuales de los docentes desde la perspectiva de los estudiantes, 3) Definir la confiabilidad general del instrumento mediante el Omega de McDonald y 4) Establecer la validez de constructo de la herramienta mediante el análisis factorial exploratorio y confirmatorio.

2. Metodología

2.1. Tipo de estudio

Se efectuó un estudio instrumental, el cual, de acuerdo con Montero y León (2007), consiste en el desarrollo de pruebas y aparatos, incluyendo su diseño y adaptación, así como el estudio de sus propiedades psicométricas.

2.2. Instrumento

Desde la mirada de las competencias individuales, se asumió como definición operacional del constructo competencias docentes la siguiente: «Las que formulan las cualidades individuales, de carácter ético, académico, profesional y social que debe reunir el docente de EMS» (DOF, 2008c, p. 1). Estas incluyen las áreas del ser (carácter ético), saber (académico), hacer (profesional) y convivir (social), junto con sus respectivas competencias y atributos que establece, superpone e interrelaciona el DOF (2008c, pp. 2-3). Tales cualidades también están alineadas con las competencias estudiantiles genéricas, disciplinares y extendidas (DOF, 2008b; DOF, 2009), así como con las competencias requeridas a los directivos de los planteles (DOF, 2008d).

Sin embargo, para elaborar la Escala de Competencias Docentes de Educación Media Superior (ECDEMS) se reflexionó sobre cada competencia que establece el DOF (DOF, 2008c), las cuales son: 1) Formación profesional continua, 2) Dominio de saberes para lograr aprendizajes significativos, 3) Planificación con enfoque por competencias y contextualizados, 4) Practicar el proceso enseñanza-aprendizaje efectivo, creativo e innovador, 5) Evaluar de manera formativa, 6) Construir aprendizajes autónomos y colaborativos, 7) Promover el desarrollo sano e integral y, 8) Participar en la mejora continua de la escuela y en la gestión institucional (pp. 2-3), y sus respectivos atributos. Estos atributos se plasmaron en los ítems que integran el instrumento, y quedó delimitado por 20 ítems redactados en sentido positivo para ser respondidos con escala Likert, cuyos valores van del 1 = Nunca, a 5 = Siempre (Tabla 1).

TABLA 1. Estructura de la ECDEMS.

Ítems
1) Cuando mi profesor(a) comete un error sobre lo que nos está enseñando, se esfuerza por mejorar en las siguientes clases.
2) Noto que mi profesor(a) muestra siempre una disposición para seguir aprendiendo y mejorar sus clases.
3) Noto que mi profesor(a) relaciona la materia con las diferentes asignaturas que estudio.
4) Mi profesor(a) nos pregunta sobre los aprendizajes adquiridos en sus clases.
5) Mi profesor(a) deja trabajos escolares basados en proyectos.
6) Cuando mi profesor(a) nos enseña, enfatiza el aprendizaje práctico.
7) Observo que mi profesor(a) planifica y organiza bien las clases que imparte.
8) Mi profesor(a) es creativo(a) para enseñar.
9) A mi profesor(a) le gusta innovar al enseñarnos.
10) Se nos hace comprensible la forma en que nos evalúa.

- 11) Mi profesor(a) nos hace observaciones sobre cómo mejorar una tarea o actividad antes de anotarnos la calificación definitiva.
- 12) Noto que mi profesor(a) al evaluarme, le interesa que mejore mi aprendizaje.
- 13) Considero que a mi profesor(a) solo le interesa calificarme para ponerme una nota.
- 14) Por su forma de enseñar me anima a seguir aprendiendo por mí mismo.
- 15) Busca formas en que aprendamos con la colaboración de nuestros compañeros.
- 16) Es respetuoso con nuestra diversidad de creencias, valores e ideas en el salón de clases.
- 17) Con su ejemplo nos inspira a practicar un estilo de vida sano.
- 18) Nos hace participar en la creación de normas de convivencia en el aula.
- 19) Observo que mi profesor (a) está comprometido con la mejora continua de la escuela.
- 20) Percibo que mi profesor (a) apoya a la administración de este plantel.

Fuente: Elaboración propia.

2.3. Análisis de validez de contenido

Para el análisis de validez de contenido, se realizó un juicio de expertos al instrumento, el cual fue conformado por 21 expertos en el tema (Tabla 2). Para su selección se consideraron criterios como la experiencia en el área de investigación, el grado académico, así como experiencia en el diseño y validación de instrumentos (CIFE, 2018b; Juárez-Hernández *et al.*, 2017; Juárez-Hernández y Tobón, 2018).

TABLA 2. Características de los expertos

Datos	Características
Género (%)	23 % hombres y 77 % mujeres
Último nivel de estudios	8 % posdoctorado, 38 % doctorado y 54 % maestría
Roles	23 % coordinadores de posgrado y administrativos, 77 % profesores investigadores en educación superior
Áreas de experiencia profesional	100 % en docencia e investigación universitaria

Promedio de años de experiencia profesional (media $\pm$ desviación estándar)	19.84 ($\pm$ 9.68)
Promedio de años de experiencia docente investigativa universitaria (media $\pm$ desviación estándar)	13.61 ($\pm$ 7.82)
Promedio de artículos publicados en el área (media $\pm$ desviación estándar)	21.69 ($\pm$ 29.45)
Promedio de libros publicados en el área (media $\pm$ desviación estándar)	4.23 ($\pm$ 7.86)
Promedio de capítulos de libro publicados en el área (media $\pm$ desviación estándar)	8.69 ($\pm$ 13.74)
Experiencia en el análisis, revisión, diseño y/o validación de instrumentos de investigación	100 %
Institución donde laboraban al momento de la evaluación del instrumento	23 % universidades privadas, 77 % uni- versidades públicas

Fuente: Elaboración propia.

El juicio de expertos tuvo un enfoque cuali-cuantitativo (Juárez-Hernández y Tobón, 2018) y se empleó la escala de jueces expertos (CIFE, 2018b), la cual integra una evaluación cualitativa en donde los jueces pueden sugerir mejoras precisas en los ítems. Para la evaluación cuantitativa, el instrumento incluye una escala Likert para evaluar la pertinencia y redacción de los ítems (Tabla 3), siendo evaluados con el coeficiente de validez de contenido V de Aiken y su intervalo de confianza al 95 % (Penfield y Giacobbi, 2004). Específicamente, se consideró un valor mínimo de aceptación superior de 0.75 para el coeficiente y de 0.50 para el valor inferior del intervalo (Bulger y Housner, 2007).

TABLA 3. Escala de juicio de expertos.

Categoría	Clasificación del nivel de pertinencia	Criterio
Pertinencia El ítem evalúa un aspecto central del propósito, dimensiones, y/o constructo teórico del instrumento.	1. No es pertinente	El ítem no contribuye a evaluar el propósito, dimensiones y/o constructo del instrumento. Puede ser eliminado completamente.
	2. Baja pertinencia	El ítem evalúa un aspecto superficial del propósito, dimensiones, y/o constructo del instrumento.
	3. Mediana pertinencia	El ítem evalúa de manera aceptable el propósito, dimensiones y/o constructo teórico del instrumento.
	4. Alta pertinencia	El ítem verdaderamente evalúa el propósito, dimensiones y/o constructo del instrumento, y está acorde con los desarrollos teóricos y metodológicos recientes del área.
Categoría	Clasificación del nivel de comprensión	Criterio
Redacción El ítem es comprensible por los usuarios potenciales y cumple con las normas gramaticales de la lengua española.	1. No es comprensible	El ítem no es comprensible por los potenciales usuarios del instrumento.
	2. Comprensión baja	El ítem debe ser mejorado en al menos la mitad de sus componentes, en aspectos de redacción y gramática.
	3. Comprensión media	El ítem requiere algunas mejoras superficiales para que quede excelente en su comprensión.
	4. Comprensión alta	El ítem es altamente comprensible por los potenciales usuarios y sigue las normas gramaticales de la lengua.

Nota: Tomado de CIFE (2018b).

2.4. Prueba piloto

Realizadas las mejoras a la escala con base a las recomendaciones de los jueces, se aplicó el instrumento a 21 estudiantes del sexto semestre de EMS. El número elegido de participantes se basó en lo descrito por Carpenter (2018), el cual refiere para esta fase que puede

estar conformado de 5 a 100 personas pertenecientes al grupo o población diana. El grupo piloto se integró por el 61.9 % mujeres y 38.1 % varones, con un promedio de edad de 17.095 años (± 0.3008). El 100 % ($n = 21$) fueron del turno matutino de un colegio privado con un nivel socioeconómico diverso. Con este pilotaje se analizó la factibilidad del instrumento (Carvajal *et al.*, 2011), que en específico se evaluó el grado de comprensión de las instrucciones e ítems mediante el cuestionario de satisfacción con el instrumento (CIFE, 2018a), y se efectuó un análisis inicial de la confiabilidad mediante el Omega de McDonald (McDonald, 1999) y sus intervalos de confianza al 95 %. Con respecto a las cuestiones éticas, a los participantes se les explicó el objetivo del instrumento, aplicando el consentimiento informado y la protección de los datos personales que señala la Cámara de Diputados del H. Congreso de la Unión (2017).

2.5. 2.5. Análisis de validez de constructo

Para efectuar el análisis, el instrumento se aplicó a una muestra por conveniencia de 3726 estudiantes de diversos subsistemas de EMS pertenecientes a varias entidades del sur-sureste mexicano de escuelas privadas y oficiales. La edad promedio fue de 17.80 (± 0.650); el 49 % mujeres y el 51 % varones. Como al grupo piloto, también se les explicó el objetivo del instrumento y se les solicitó el consentimiento informado y la protección de los datos personales que señala la Cámara de Diputados del H. Congreso de la Unión (2017).

Inicialmente se realizó un análisis de comprobación de ajuste de los ítems a la distribución normal a través del índice de Mardia (Mardia, 1970). Se analizó la correlación ítem-prueba y si algún ítem presentaba un valor superior a 0.90 o menor a 0.20 era sujeto de eliminación (Tabachnick y Fidell, 2001). Se calculó el Omega de McDonald (McDonald, 1999) para evaluar la fiabilidad del instrumento.

Posteriormente, la muestra se dividió en dos partes iguales ($N1 = 1863$; $N2 = 1863$), empleando el uso de números aleatorios a través de una hoja de cálculo electrónica. La primera submuestra se analizó con el análisis factorial exploratorio (AFE), y la segunda a través del análisis factorial confirmatorio (AFC). Estos análisis se efectuaron con el software R (R Core Team, 2020). En este orden, el AFE se realizó con la primera submuestra para explorar la estructura, y con la segunda submuestra se procedió con el AFC para confirmar esa estructura (Brown, 2006). Específicamente, con este proceso se asegura que la estructura factorial identificada no sea una característica específica de la submuestra inicial, se reduce la posibilidad de los sesgos, permite una mejor evaluación del ajuste, estabilidad y calidad del modelo (Carpenter, 2018; Kline, 2013; Lloret-Segura *et al.*, 2014; Worthington y Whittaker, 2006).

Referente al AFE, se analizó el valor de la determinante, la prueba de KMO y la prueba de Bartlett para la determinación de su pertinencia de análisis (Howard, 2016). Una vez verificados se eligió el tipo de matriz de correlación y del método de extracción, denotando que si los ítems presentaban una distribución normal, se elegirá la matriz de correlación momento-producto de Pearson y el método de máxima verosimilitud (Howard, 2016; Lloret-Segura *et al.*, 2014), y si no se cumplía el supuesto, se empleó la matriz de correlación policórica y el método de mínimos cuadrados no ponderados (ULS) (Xia y Yang, 2019). El número de factores a retener se basó en la técnica de máximo consenso entre los 23 métodos utilizados con el apoyo de la librería nFactors (Lüdtke *et al.*, 2020). Para el análisis de la matriz factorial se precisó que la significancia de las cargas factoriales por ítem fuera mayor a 0.55 (Hair *et al.*, 2010). Si se presentaban problemas de complejidad factorial, la matriz se rotó mediante el método de mayor conveniencia (Juárez-Hernández, 2018; Lloret-Segura *et al.*, 2014).

La estructura factorial obtenida mediante el AFE se analizó con la segunda submuestra mediante el AFC. El método de estimación fue el de máxima verosimilitud si se cumplía el supuesto de normalidad, y si no, se empleó el método de estimación de mínimos cuadrados no ponderados (ULS) (Yuan, 2005). La evaluación de bondad de ajuste del modelo se realizó mediante diversos indicadores e índices (χ^2/gf , GFI, CFI, TLI, RMSAE, RMR), considerando los valores umbrales indicados por Ráczová *et al.* (2021). Este análisis se realizó con el software R y librería laavan (Rosseel, 2012) y psych (Revelle, 2017).

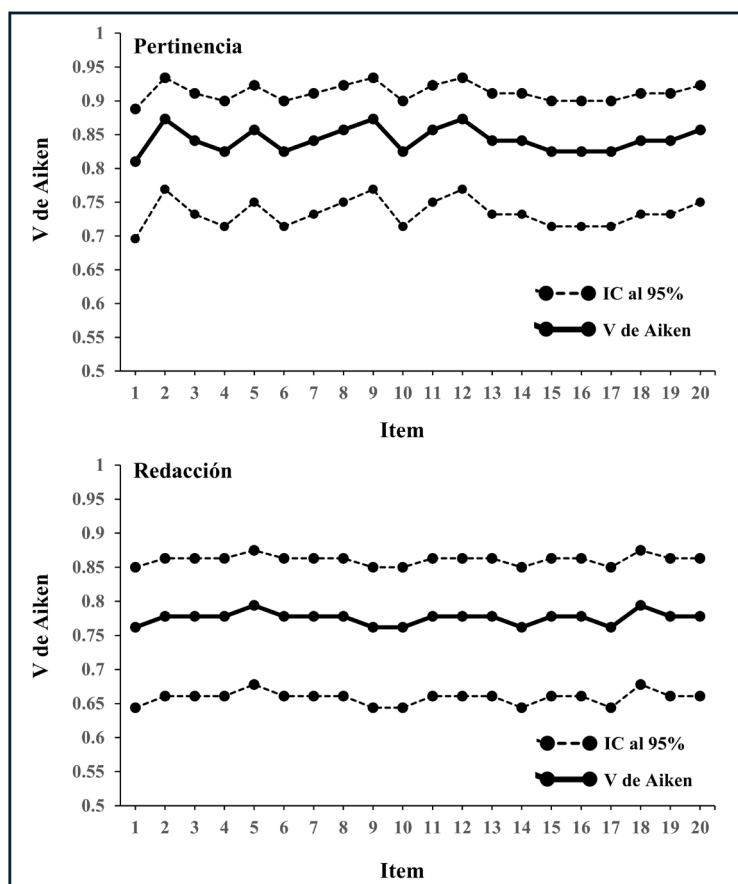
Mediante las cargas factoriales estandarizadas, se calculó la varianza media extraída (VME) y la confiabilidad compuesta considerando el valor umbral para la primera mayor a 0.50 (Fornell y Larcker, 1981) y para la segunda de 0.70 (Hair *et al.*, 2014). Se analizó la validez convergente considerando que se debe de cumplir que la VME sea mayor o igual a 0.5, que las cargas factoriales estandarizadas sean mayores a 0.50 y que la confiabilidad compuesta sea superior a 0.70 (Fornell y Larcker, 1981; Hair *et al.*, 2014). Finalmente, se evaluó la confiabilidad mediante el Omega de McDonald (McDonald, 1999) y sus intervalos de confianza al 95 %.

3. Resultados

3.1. Análisis de validez de contenido

La totalidad de ítems fueron validados en los criterios de pertinencia y redacción (V de Aiken > 0.75, VI > 0.50) (Figura 1).

FIGURA 1. Representación de los resultados del análisis cuantitativo del juicio por expertos.



Fuente: Elaboración propia.

La evaluación cualitativa brindó algunas observaciones, las cuales se refieren a cuestiones de redacción, y que se precisa fueron atendidas (Tabla 4).

TABLA 4. Observaciones de los jueces.

Ítems	Observaciones
13. Considero que a mi profesor(a) solo le interesa calificarme para ponerme una nota.	Juez 1: La pregunta debería ser en un mismo sentido a las demás porque esto dificulta la calificación.
19. Observo que mi profesor(a) está comprometido con la mejora continua de la escuela.	Juez 2: La pregunta debería ser en un mismo sentido a las demás porque esto dificulta la calificación.

Fuente: Elaboración propia.

3.2. Grupo piloto

En términos generales, los participantes indicaron un buen grado de comprensión de los ítems (66.7 %) y un buen grado de comprensión de las instrucciones (76 %). Respecto al grado de relevancia de los ítems y su pertinencia los refirieron en su mayoría como muy importantes (47.6 %). El tiempo promedio en responder al instrumento fue de 5.71 minutos. La confiabilidad en esta fase fue óptima (Omega de McDonald: 0.920 IC al 95 %: 0.853 ± 0.960).

3.3. Análisis de los ítems y confiabilidad

Se comprobó la ausencia de normalidad multivariante (Curtosis $p < 0.05$; Asimetría < 0.05). Respecto a la correlación ítem-test (CIT) se encontró que era necesario eliminar el ítem 13 (Tabla 5). A este respecto, este ítem hace una afirmación negativa e indica: «Considero que a mi profesor (a) solo le interesa calificarme para ponerme una nota». Finalmente, la confiabilidad fue óptima (Omega de McDonald: 0.959 IC al 95 %: 0.957 ± 0.961).

TABLA 5. Análisis inicial de los ítems.

Ítem	Correlación Ítem-test
1	0.741
2	0.793
3	0.535
4	0.67
5	0.382

6	0.746
7	0.792
8	0.784
9	0.789
10	0.734
11	0.722
12	0.759
13	0.044
14	0.773
15	0.743
16	0.612
17	0.752
18	0.709
19	0.777
20	0.684

Fuente: Elaboración propia.

3.4. Análisis de validez de constructo

El análisis de los supuestos de aplicación del AFE fue satisfactorio, ya que se observó que las variables se encuentran relacionadas significativamente ($p < 0.05$) y presentó una determinante cercana a cero (0.0000001735). Asimismo, las pruebas Kaiser Meyer Olkin (KMO: 0.98) y de esfericidad de Bartlett ($p < 0.001$) mostraron la pertinencia de aplicación del AFE.

Referente a la extracción de factores, se precisó la retención de un factor, el cual integró 18 de los 19 ítems (con cargas factoriales mayores a 0.50) y explicó más del 60 % de la varianza (Tabla 6). El modelo factorial resultante se correspondió con el modelo teórico, el cual indica que los ítems se refieren a las características personales de los docentes que señala el Acuerdo 447, expresadas en forma de competencias y atributos, porque se relacionan de manera directa con el ser, saber, hacer y convivir.

TABLA 6. Resultado del análisis factorial exploratorio.

Ítem	Carga factorial
9	0.87
2	0.86
7	0.85
8	0.84
14	0.83
19	0.83
12	0.82
17	0.81
1	0.8
6	0.79
15	0.79
10	0.79
11	0.77
18	0.76
20	0.73
4	0.72
16	0.69
3	0.56

Fuente: Elaboración propia.

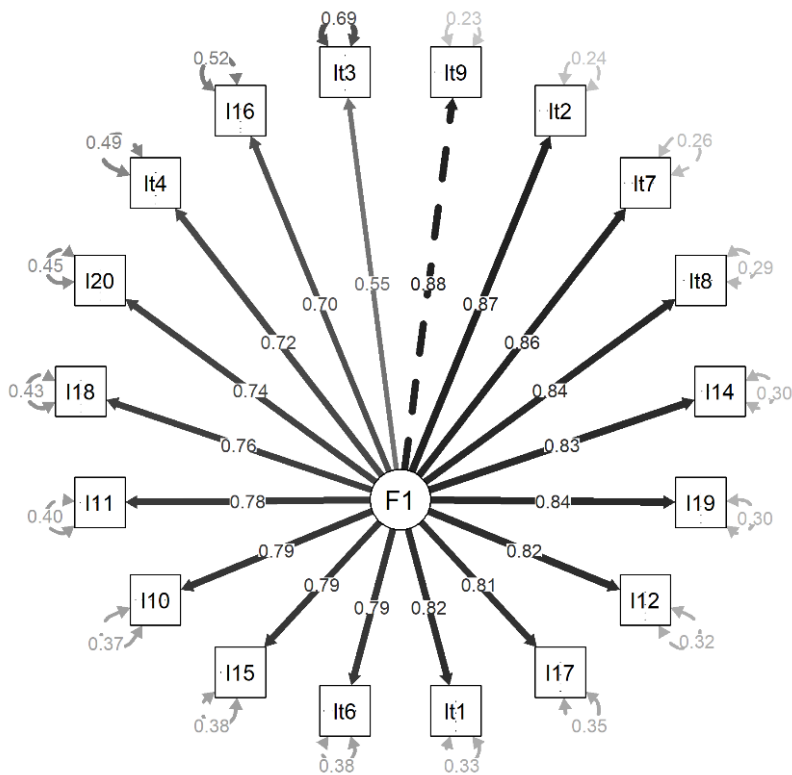
El AFC mostró un buen ajuste del modelo obtenido (Tabla 6), ya que los diferentes índices mostraron valores óptimos (Tabla 7). El valor de la varianza media extraída (VME: 0.7905) y de la confiabilidad compuesta (CC: 0.966) fueron óptimos. Cada ítem presentó cargas factoriales estandarizadas mayores a 0.50 (Figura 2).

TABLA 7. Ajuste del modelo factorial.

Índice	Valor esperado (Ráczová et al., 2021)	Valor obtenido
Razón chi-cuadrada/grados de libertad (χ^2/df)	Menor a 3	1.89
Índice de bondad de ajuste (GFI)	Mayor a 0.90	0.995
Error cuadrático medio de aproximación (RMSAE)	0.050 a 0.080	0.050 (0.048 a 0.079)
Raíz cuadrada residual (RMR)	Menor a .050	0.028
Índice comparativo de ajuste (CFI)	Mayor a 0.95	0.992
Índice de Tucker-Lewis (TLI)	Mayor a 0.90	0.994

Fuente: Elaboración propia.

FIGURA 2. Representación del modelo factorial confirmatorio.



Fuente: Elaboración propia.

4. Discusión

El estudio del constructo competencias docentes, entre otras razones, forma parte de la profesionalización del profesorado para brindar servicios educativos de calidad, ya que son actores clave en los sistemas educativos por la efectividad y eficiencia que les compete demostrar en el proceso enseñanza-aprendizaje, con el propósito de elevar el logro educativo estudiantil para que estos accedan a una mejor calidad de vida ante el mundo globalizado.

También se precisa que la evaluación de las competencias docentes sea importante en EMS, ya que provee datos valiosos para los tomadores de decisiones a nivel de políticas públicas y al interior de las instituciones. Específicamente, permiten diagnosticar la capacidad que tiene el docente para facilitar los conocimientos e identificar sus áreas fuertes y débiles como puntos de referencia para mejorar la calidad educativa y profesional; ya que, a fin de cuentas, todos los procesos educativos y actores de los mismos están de manera íntima relacionados entre sí, en aras de la formación de un estudiantado competitivo y apto para integrarse al mundo laboral, o bien para continuar con sus estudios superiores.

Ante estos hechos, aunque se demostró que existen herramientas para la evaluación de las competencias docentes, se constató que presentan limitaciones en cuanto a bases teóricas sólidas, además de que no han sido probadas en diversos contextos para afirmar que son adecuados para medir el constructo competencias docentes en EMS. Es así como surgió la necesidad de diseñar y validar de manera integral la ECDEMS, con base a los planteamientos teóricos del Marco Curricular Común del Sistema Nacional de Educación Media Superior y alineados con las competencias de los estudiantes y directivos de los planteles educativos del mismo nivel.

Acorde con estos planteamientos se da relevancia a las cualidades individuales del ser, conocer, hacer y convivir del docente como profesional, las cuales se transformaron en ítems elaborados en forma de escala por ser la herramienta que se recomienda utilizar cuando se miden atributos de una sola variable (competencias docentes), con opciones de respuestas graduadas, de entre las cuales tendrá que elegir una de ellas quien las responda, en este caso, el alumno.

Creado el instrumento, es elemental evaluar la calidad de su medición, lo cual se realiza con el análisis de sus propiedades psicométricas (Carvajal *et al.*, 2011). A este respecto, mediante el juicio por expertos se encontró que todos los ítems presentaron una validez de contenido, lo que significa que el instrumento refleja un dominio específico del contenido que mide (Hernández-Sampieri *et al.*, 2010). Es fundamental mencionar que el juicio de expertos se efectuó con pautas de calidad, como la selección de expertos, el número de expertos, la evaluación cuali-cuantitativa y el análisis mediante un índice de relevancia (Juárez-Hernández y Tobón, 2018; Maldonado-Suárez y Santoyo-Telles, 2024), lo cual lo dotó de robustez y precisión al juicio.

Referente al pilotaje, se precisa que es una fase elemental, ya que su objetivo es efectuar un primer acercamiento de la adecuación del instrumento a la población objetivo, verificar aspectos logísticos de aplicación, administración del instrumento y de análisis inicial de confiabilidad (Carpenter, 2017; Maldonado-Suárez y Santoyo-Telles, 2024; Muñiz y Fonseca-Pedrero, 2019). En este sentido, la adecuación del instrumento en esta fase fue satisfactoria, ya que el grado de comprensión de los ítems fue aceptable y el grado de comprensión de las instrucciones fue bueno. Por su parte, la confiabilidad en esta aplicación fue óptima.

Para el análisis de validez de constructo, se efectuó un proceso de validación cruzado, el cual se considera como el óptimo (Lloret-Segura *et al.*, 2014). Se empleó el AFE para prospectar la estructura que subyace al conjunto de ítems, y el AFC para la validación de esa estructura (Lloret-Segura *et al.*, 2014). En este orden, el análisis previo de los ítems reveló que el ítem 13, redactado como «Considero que a mi profesor (a) solo le interesa calificarme para ponerme una nota», presentó una baja correlación con el instrumento. La lógica indica que es correcta que así resulte la correlación, porque aborda por una parte una incompetencia, objeto que no

es parte del estudio desarrollado, y por otra puede ser por incomprensión del que responde, o por falta de claridad en la redacción.

Acorde a los resultados del AFE, se precisa la correspondencia con el modelo teórico propuesto, es decir un solo factor, el cual explicó más del 60 % de la varianza, lo cual se considera óptimo (Velicer y Fava, 1998). En este sentido, este factor incorporó el 95 % de los ítems, lo que significa que estos representan el constructo abordado (Lagunes, 2017) y que enfatiza la significancia del proceso de validación de contenido efectuado (Hayness *et al.*, 2009). Respecto al ítem 5, que no se representó en el modelo factorial («Mi profesor (a) deja trabajos escolares basados en proyectos»), puede ser por un lado porque aborda aspectos de actividades extra clases y de trabajos inespecíficos de largo alcance, y por otro, como ya se dijo, por falta de claridad en la redacción o de incomprensión del ítem por parte del sujeto que responde.

La evaluación del modelo factorial mediante el AFC denotó la comprobación del ajuste del modelo a los datos, ya que todos los indicadores mostraron valores cercanos al óptimo. Un aspecto para señalar fue el valor de las cargas factoriales estandarizadas, la varianza media extraída y la confiabilidad compuesta, que denotan la existencia de una validez convergente (Fornell y Larcker, 1981; Hair *et al.*, 2010), lo que significa que los indicadores propuestos miden adecuadamente dicho factor (Cheung y Wang, 2017).

Al observar con la validez de contenido que el instrumento mide un dominio específico (por ejemplo, las competencias personales), que el AFE precisa la correspondencia con el modelo teórico propuesto, y que el AFC corrobora la sustentabilidad empírica del modelo, se demuestra que el instrumento atiende desde las competencias docentes, las necesidades y características propias de la naturaleza humana. Dichas en otras palabras, enfatizan las competencias analíticas, las orientadas a la acción y las sociales de Baartman (2007) mientras busca desarrollar el bienestar personal y la progresión de los estudiantes (Braun y Hooper, 2024; Cachutt-Alvarado *et al.*, 2024).

Referente a la confiabilidad, acorde con los resultados fue óptima, cabe señalar la precisión y el grado de correlación entre los 18 ítems finales del instrumento, e indica que la medición se encuentra libre de error (Carvajal *et al.*, 2011; Jabrayilov *et al.*, 2016). Ante esto, se puede concluir que la ECDEMS posee adecuadas propiedades psicométricas que dan cuenta de la calidad de su medición.

Como fortaleza, el estudio se desarrolló en el contexto del ideal que la literatura presenta al efectuar el análisis de validez de contenido, seguido del análisis de validez de constructo mediante el proceso de validación cruzada (AFE y AFC), con una muestra poblacional razonablemente amplia de un sector geográfico importante de la zona sur-sureste de México.

Una limitación importante del presente estudio es de carácter metodológico por elegir una muestra probabilística que impide generalizar los hallazgos. Otra puede ser la interpretación que cada participante dio a cada ítem, así como las condiciones socioemocionales en las que se encontraba al momento de responder. Una cuarta limitación, la forma o las circunstancias en las que los investigadores hicieron el levantamiento de la información.

5. Conclusión

Implementar escalas para evaluar como el presente instrumento tiene importantes implicaciones teórico-prácticas porque permite contrastar la medición del quehacer docente de EMS, con lo que afirma la teoría. Para los futuros docentes, les permite identificar áreas de mejora de manera personalizada; así como posibilidades de retroalimentación profesional y el desarrollo de competencias claves y específicas que busquen el éxito del logro educativo estudiantil, así como a mejorar la eficacia de los programas locales de formación profesoral.

Por otra parte, esta herramienta puede ser usada por los responsables de las políticas educativas y gestores institucionales de educación media superior para diagnosticar las competencias docentes a nivel macro y establecer las estrategias correspondientes a mediano y largo plazo.

Por lo anterior, se recomienda para futuras investigaciones replicar el estudio en otros contextos dentro y fuera del país, seleccionar la muestra mediante un muestreo probabilístico para generalizar los resultados que se obtengan, además de realizar estudios comparativos por géneros para valorar la percepción que tienen los estudiantes sobre sus profesores, de acuerdo con las regiones de donde provienen.

Contribuciones de los autores

Dr. Emerson López-López: Conceptualización, curación de datos, investigación, metodología, recogida de datos, administración del proyecto, redacción del borrador original, redacción, revisión y edición.

Dr. Luis Gibran Juárez-Hernández: Curación de datos, metodología, análisis formal, supervisión, validación, visualización, redacción del borrador original, redacción, revisión y edición.

Referencias bibliográficas

- Abarca-Cedeño, M. S., Juárez-Hernández, L. G., Márquez-Orozco, L., y Villanueva-Magaña, R. M. (2024). Validación de constructo y confiabilidad de una escala de autoevaluación de competencias docentes y contexto institucional para el fomento de la creatividad. *Ciencia Latina Revista Científica Multidisciplinar*, 8(1), 5271-5292. https://doi.org/10.37811/cl_rcm.v8i1.9860
- AERA-APA-NCME. (2014). *Standards for educational and psychological testing*. AERA.
- Agüero-Servín, M. (2022). Formación y profesionalización docente en la educación media superior en México. Revisión crítica y narrativa de la literatura científica. *Revista Electrónica en Educación y Pedagogía*, 6(10), 228-248. <https://doi.org/10.15658/rev.electron.educ.pe-dagog22.04061015>
- Arribas-Águila, D., Castaño, G., y Martínez-Arias, R. (2024). A systematic review of evidence-based general competency models: Development of a general competencies Taxonomy [Una revisión sistemática y síntesis de modelos de competencias generales basados en la evidencia: desarrollo de una taxonomía de competencias generales]. *Journal of Work and Organizational Psychology*, 40(2), 61-76. <https://doi.org/10.5093/jwop2024a6>
- Aydin, M. K., Yildirim, T., y Kus, M. (2024). Teachers' digital competences: a scale construction and validation study [Competencias digitales del profesorado: un estudio de construcción y validación de escalas]. *Frontiers in Psychology*, (15), 1-13. <https://doi.org/10.3389/fpsyg.2024.1356573>
- Baena-Extremuera, A., Granero-Gallegos, A., y Martínez-Molina, M. (2015). Validación española de la escala de evaluación de la competencia docente en educación física de secundaria. *Cuadernos de Psicología del Deporte*, 15(3), 113-122. <https://dx.doi.org/10.4321/S1578-84232015000300011>
- Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A., y Van der Vleuten, C. P. M. (2007). Evaluating assessment quality in competence-based education a qualitative comparison of two framework [Evaluación de la calidad de la evaluación en la educación basada en competencias: una comparación cualitativa de dos marcos]. *Educ. Res. Rev.*, (2), 114-129. <https://doi.org/10.1016/j.edurev.2007.06.001>
- Bleiberg, J., Brunner, E., Harbatkin, E., Kraft, M. A., y Springer, M. G. (2024). Taking teacher evaluation to scale: The effect of State reforms on achievement and attainment [Llevar la evaluación docente a escala: el efecto de las reformas estatales en el rendimiento y el logro]. *Education Policy Analysis Archives*, 32(1), 1-24. <https://doi.org/10.26300/b1ak-r251>
- Boyatzis, R. E. (1983). *The competent manager. A model for effective performance*. Wiley.
- Boyatzis, R. E. (2006). Core competencies in coaching others to overcome dysfunctional behavior [Competencias básicas en el entrenamiento de otros para superar el comportamiento disfuncional]. En V. U. Druskat, F. Sala, y G. Mount (eds.), *Linking emotional intelligence*

- and performance at work: Current research evidence with individuals and groups, 81-95. Lawrence Erlbaum Associates Publishers.
- Braun, S. S., y Hooper, A. L. (2024). Social and emotional competencies predict pre-service teachers' occupational health and personal well-being [Las competencias sociales y emocionales predicen la salud ocupacional y el bienestar personal de los docentes en formación]. *Teaching and Teacher Education*, (147), 1-11. <https://doi.org/10.1016/j.tate.2024.104654>
- Brown, G. T. L. (2024). Editorial: Insights in assessment, testing, and applied measurement: 2022 [Editorial: Perspectivas sobre evaluación, pruebas y medición aplicada: 2022]. *Frontiers in Education*, (9), 1-3. <https://doi.org/10.3389/feduc.2024.1488012>
- Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research* [Análisis factorial confirmatorio para investigación aplicada]. The Guilford Press.
- Bulger, S. M., y Housner, L. D. (2007). Modified Delphi investigation of exercise science in physical education teacher education [Investigación Delphi modificada de la ciencia del ejercicio en la formación de profesores de educación física]. *Journal of Teaching in Physical Education*, 26(1), 57-80. <https://doi.org/10.1123/jtpe.26.1.57>
- Burbano, A., Ledesma, A. J., y Ordoñez, D. A. (2024). Competencias docentes de ingeniería. Revisión bibliográfica de literatura. *Revista Mexicana de Investigación Educativa*, 29(101), 181-203. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-66662024000200181
- Cabero-Almenara, J., y Palacios-Rodríguez, A. (2020). Marco europeo de competencia digital docente «DigCompEdu». Traducción y adaptación del cuestionario «DigCompEdu». *ED-METIC Revista de Educación Mediática y TIC* 9(1), 213-234. <https://doi.org/10.21071/edmetic.v9i1.12462>
- Cachutt-Alvarado, C., Méndez-Gómez-Humaran, I., y Velasco-Álvarez, J. (2024). Personal competencies for work-family integration and its relationship with productivity and comprehensive health in salaried professionals [Competencias personales para la integración trabajo-familiar y su relación con la productividad y la salud integral en profesionales asalariados]. *Safety*, 10(1), 1-25. <https://doi.org/10.3390/safety10010028>
- Cámara de Diputados del H. Congreso de la Unión. (2017). *Ley general de protección de datos personales en posesión de sujetos obligados*. <https://www.diputados.gob.mx/LeyesBiblio/pdf/LGPDPPSO.pdf>
- Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers [Diez pasos en el desarrollo de escalas y la elaboración de informes: una guía para investigadores]. *Communication Methods and Measures*, 12(1), 25-44. <https://doi.org/10.1080/19312458.2017.1396583>
- Carvajal, A., Centeno, C., Watson, R., Martínez, M., y Sanz Rubiales, A. (2011). ¿Cómo validar un instrumento de medida de la salud. *Anales del sistema sanitario de Navarra*, 34(1), 63-72. <https://dadun.unav.edu/server/api/core/bitstreams/e7e33006-2f83-4f55-a180-ef5c-297b54b6/content>
- Castro-Castillo G, Cárdenas-Cobo, J., Soldevilla, M. T., y Vidal-Silva, C. (2024). Analyzing the teaching profile and competency-based training in online education: a case study of Ecuadorian professors [Análisis del perfil docente y la formación por competencias en la educación en línea: un estudio de caso de profesores ecuatorianos]. *Frontiers in Education*, 9. <https://doi.org/10.3389/feduc.2024.1397086>
- Chee Yuet, F. K., Yusof, H., y Syed Mohaman, S. I. (2016). Development and validation of the teacher leadership competency scale [Desarrollo y validación de la escala de competencias de liderazgo docente]. *Malaysian Journal of Learning and Instruction*, 13(2), 43-69. <https://e-journal.uum.edu.my/index.php/mjli/article/view/mjli2016.13.2.2>
- Cheung, G. W., y Wang, C. (2017). Current approaches for assessing convergent and discriminant validity with SEM: Issues and solutions [Enfoques actuales para evaluar la validez convergente y discriminante con SEM: problemas y soluciones]. *Academy of Management Proceedings*, (1). <https://doi.org/10.5465/AMBPP.2017.12706abstract>
- CIFE. (2018a). *Instrumento «Cuestionario de satisfacción con el instrumento»*. <https://www.google.com/url?q=https://docs.google.com/forms/d/e/1FAIpQLSc8-jOiWYwG64Qb-nhRyGAg1EI7ggq2aP1XiSg45pyN9XLbXNQ/viewform&sa=D&source=hangouts&ust=1590684761042000&usq=AFQjCNGAhEPiwC0U3VYcgk29198gFzHJSQ>

- CIFE. (2018b). *Instrumento «Escala Jueces Expertos»*. <https://www.google.com/url?q=https://docs.google.com/forms/d/e/1FAIpQLScWn-tYKAVPlr2EFBq-723JYZWMNabLmukc-ReaOWp0tUc0IzA/viewform&sa=D&source=hangouts&ust=1590684754446000&usg=AFQjCNFv70JsedumZjNtxrRZH8qylratQQ>
- Cinque Gómez-del-Pulgar, S. y Rodríguez-Mantilla, J. M. (2020). Las competencias instrumentales en los futuros Maestros de Educación Primaria: Autopercepción y satisfacción con la formación recibida en estudiantes de la UCM. *Profesorado. Revista de Currículum y Formación de Profesorado*, 24(3), 309-333. 10.30827/profesorado.v24i3.8158]
- Cortes, S., Pineda, H., y Geveola, I. J. (2020) Development and validation of a scale on teacher's competence in action research [Desarrollo y validación de una escala de competencias docentes en Investigación-Acción]. *International Journal of Research Studies in Education*, 9(6), 77-85. <https://doi.org/10.5861/ijrse.2020.5041>
- Di Lisio G., Mila Roa, A., Halty, A., Berástegui, A., Couso Losada, A., y Pitillas, C. (2025) Nurturing bonds that empower learning: a systematic review of the significance of teacher-student relationship in education [Cultivando vínculos que potencian el aprendizaje: una revisión sistemática de la importancia de la relación profesor-alumno en la educación]. *Front. Educ.*, 10. <https://doi.org/10.3389/feduc.2025.1522997>
- DOF. (02 diciembre 2008d). *Acuerdo número 449 por el que se establecen las competencias que definen el perfil del director en los planteles que imparten educación del tipo medio superior*. <https://www.sep.gob.mx/work/models/sep1/Resource/7aa2c-3ff-aab8-479f-ad93-db49d0a1108a/a449.pdf>
- DOF. (17 enero 2008a). *Programa sectorial de educación 2007-2012*. https://www.dof.gob.mx/nota_detalle.php?codigo=5028684&fecha=17/01/2008&print=true
- DOF. (21 octubre 2008b). *Acuerdo número 444 por el que se establecen las competencias que constituyen el marco curricular común del Sistema Nacional de Bachillerato*. https://educacionmediasuperior.sep.gob.mx/work/models/sems/Resource/10905/1/images/Acuerdo_444_marco_curricular_comun_SNB.pdf
- DOF. (29 octubre 2008c). *Acuerdo número 447 por el que se establecen las competencias docentes para quienes impartan educación media superior en la modalidad escolarizada*. https://dof.gob.mx/nota_detalle.php?codigo=5066425&fecha=29/10/2008#gsc.tab=0
- DOF. (30 abril 2009). *Acuerdo número 486 por el que se establecen las competencias disciplinares extendidas del Bachillerato General*. https://educacionmediasuperior.sep.gob.mx/work/models/sems/Resource/10905/1/images/Acuerdo_486_competencias_extendidas.pdf
- Domínguez-González, N., y Serna-Poot, D. (2021). Competencias docentes para la educación en línea. *RIDE Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 11(22), 1-27. <https://doi.org/10.23913/ride.v11i22.968>
- Fernández-Díaz, M. J., Rodríguez-Mantilla, J. M., y Fernández-Cruz, F. J. (2016). Evaluación de competencias docentes del profesorado para la detección de necesidades formativas. *Bordón*, 68(2), 85-101. <https://doi.org/10.13042/Bordon.2016.68206>
- Firda, S., y Khairat, A. (2023). The impact of teacher skills'on students' interest and learning outcomes [El impacto de las habilidades docentes en el interés y los resultados de aprendizaje de los estudiantes]. *International Journal of Education, Management, and Technology*, 1(1), 1-10. <https://doi.org/10.58578/ijemt.v1i1.1881>
- Fornell, C., y Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error [Evaluación de modelos de ecuaciones estructurales con variables no observables y error de medición]. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.2307/3151312>
- González, J., y Estrella, J. (2023). Educación del Siglo XXI, competencias, metodologías y estrategias. *Esprint Investigación*, 2(1). 5-15. <https://doi.org/10.61347/ei.v2i1.50>
- Granada, G. D., y Oco, R. M. (2024). Classroom management and teaching competencies of elementary teachers [Gestión del aula y competencias docentes de los profesores de primaria]. *International Journal of Multidisciplinary Research and Analysis*, 7(3), 1276-1283. <https://doi.org/10.47191/ijmra%2Fv7-i03-50>

- Gümüş, M. M., y Volkan, K. (2023) Developing a digital competence scale for teachers: validity and reliability study [Desarrollo de una escala de competencias digitales para docentes: estudio de validez y fiabilidad]. *Education and Information Technologies*, 28(1), 2747-2765. <https://doi.org/10.1007/s10639-022-11213-2>
- Hair, J. F., Black, W. C., Babin, B. J., y Anderson, R. E. (2010). *Multivariate data analysis* [Análisis de datos multivariados]. (7ma. ed.). Pearson.
- Hair, J. F., Black, W. C., Babin, B. J., y Anderson, R. E. (2014). *Multivariate data analysis* [Análisis de datos multivariados]. (7ma. ed.). Pearson.
- Haynes, S. N., Richard, D., y Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods [Validez de contenido en la evaluación psicológica: una aproximación funcional a conceptos y métodos]. *Psychological Assessment*, 7(3), 238-247. [10.1037/1040-3590.7.3.238](https://doi.org/10.1037/1040-3590.7.3.238)
- Hernández-Sampieri, R., Fernández, Collado, C., y Baptista Lucio, P. (2010). *Metodología de la Investigación*. (5ta. ed.). McGraw-Hill Interamericana.
- Hernández-Suárez, C. A., Gamboa-Suárez, A., y Avendaño-Castro, W. R. (2021). Validación de una escala para evaluar competencias investigativas en docentes de Básica y Media. *Revista Boletín Redipe*, 10(6), 393-406. <https://doi.org/10.36260/rbr.v10i6.1335>
- Hollenstein, L., y Brühwiler, C. (2024). The importance of teachers' pedagogical-psychological teaching knowledge for successful teaching and learning [La importancia de los conocimientos pedagógico-psicológicos pedagógicos de los docentes para el éxito de la enseñanza y el aprendizaje]. *Journal of Curriculum Studies*, 56(4), 480-495. <https://doi.org/10.1080/00220272.2024.2328042>
- Howard, C. M. (2016). A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? [Una revisión de las decisiones del análisis factorial exploratorio y una visión general de las prácticas actuales: ¿Qué estamos haciendo y cómo podemos mejorar?]. *International Journal of Human-Computer Interaction*, 32(1), 51-62. <https://doi.org/10.1080/10447318.2015.1087664>
- Hunter, S. B., y Kho, A. (2023). The effects of teacher evaluation policy on student achievement and teacher turnover: Leveraging teacher accountability and teacher development [Los efectos de la política de evaluación docente en el rendimiento estudiantil y la rotación de docentes: Aprovechar la responsabilidad de los docentes y el desarrollo docente]. *Journal of Education Human Resources*, (aop), e20230040. <https://doi.org/10.3138/jehr-2023-0040>
- Jabrayilov R., Emons, W. H. M., y Sijtsma, K. (2016). Comparison of classical test theory and item response theory in individual change assessment [Comparación de la Teoría Clásica de las Pruebas y la Teoría de la Respuesta al Ítem en la Evaluación del Cambio Individual]. *Applied Psychological Measurement*, 40(8), 559-572. <https://doi.org/10.1177/0146621616664046>
- Juárez-Hernández, L. G., y Tobón, S. (2018). Análisis de los elementos implícitos en la validación de contenido de un instrumento de investigación. *Espacios*, 39(53), 23-30. <http://www.revistaespacios.com/cited2017/cited2017-23.pdf>
- Juárez-Hernández, L. G. (2018). *Manual práctico de estadística básica para la investigación*. KResearch Corp. <http://dx.doi.org/10.24944/isbn.978-1-945721-24-3>
- Kim, B. H., y Kim, J. (2016) Development and validation of evaluation indicators for teaching competency in STEAM education in Korea [Desarrollo y validación de indicadores de evaluación para la competencia docente en educación STEAM en Corea]. *Eurasia Journal of Mathematics, Science and Technology Education*, 12(7), 1909-1924. <https://doi.org/10.12973/eurasia.2016.1537a>
- Kline, R. B. (2013). Beyond significance testing: Statistics reform in the behavioral sciences (2nd ed.). [Más allá de las pruebas de significancia: Reforma estadística en las ciencias del comportamiento (2.ª ed.)]. American Psychological Association. <https://doi.org/10.1037/14136-000>
- Lagunes-Córdoba, R. (2017). Recomendaciones sobre los procedimientos de construcción y validación de instrumentos y escalas de medición en la psicología de la salud. *Revista Psicología y Salud*, 27(1), 5-18. <https://psicologiaysalud.uv.mx/index.php/psicysalud/article/view/2431/4279>

- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., y Tomas-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151-1169. <https://scielo.isciii.es/pdf/ap/v30n3/metodologia.pdf>
- Lüdtke, D., Ben-Shachar, M. S., Patil, I., y Makowski, D. (2020). Extracting, computing and exploring the parameters of statistical models using R [Extracción, cálculo y exploración de los parámetros de los modelos estadísticos utilizando R]. *Journal of Open Source Software*, 5(53), 1-5. <https://doi.org/10.21105/joss.02445>
- Luna-Serrano, E., y Reyes-Piñuelas, E. P. (2015). Validación de constructo de un cuestionario de evaluación de la competencia docente. *Revista Electrónica de Investigación Educativa*, 17(3), 13-27. <https://redie.uabc.mx/redie/article/view/1090>
- Maldonado-Suárez, N., y Santoyo-Telles, F. (2024). Validez de contenido por juicio de expertos: Integración cuantitativa y cualitativa en la construcción de instrumentos de medición. *REIRE Revista d'Innovació i Recerca en Educació*, 17(2), 1-19. <https://doi.org/10.1344/reire.46238>
- Mardia, K. (1970). Measures of multivariate skewness and kurtosis with applications [Medidas de asimetría multivariante y kurtosis con aplicaciones]. *Biometrika*, 57(3), 519-530. <https://doi.org/10.1093/biomet/57.3.519>
- McClelland, D. C. (1973). Testing for competence rather than for intelligence [Evaluar la competencia en lugar de la «inteligencia»]. *The American Psychologist*, (28), 1-14. <https://doi.org/10.1037/h0034092>
- McDonald, R. P. (1999). *Test theory: A unified treatment* [Teoría de las pruebas: un tratamiento unificado]. Lawrence Erlbaum Associates Publishers.
- Montero, I., y León, O. G. (2007). Una guía para nombrar los estudios de investigación en Psicología. *International Journal of Clinical and Health Psychology*, 7(3), 847-862. https://web.archive.org/web/20180508202326id_/http://aepc.es/ijchp/GNEIP07_es.pdf
- Morán, R., Cardoso, E. O., Cerecedo, M. T., y Ortiz, J. C. (2015). Evaluación de las competencias docentes de profesores formados en instituciones de educación superior: El caso de la asignatura de tecnología en la enseñanza secundaria. *Formación Universitaria*, 8(3), 57-64. <http://dx.doi.org/10.4067/s0718-50062015000300007>
- Muñiz, J., y Fonseca-Pedrero, E. (2019). Diez pasos para la construcción de un test. *Psicothema*, 31(1), 7-16. <https://www.psicothema.com/pdf/4508.pdf>
- Nessa, K., Akter, S., Faruk, G., y Raihan, F. (2024). Teaching competence and academic achievement on Higher Secondary students in Satkhira District [Competencia docente y rendimiento académico en estudiantes de secundaria superior en el distrito de Satkhira]. *International Journal of Research and Innovation in Social Science*, 8(3), 446-456. <https://doi.org/10.47772/ijriss.2024.803030s>
- Otto, A. (2024). CLIL in pre-primary education. A qualitative analysis of teachers' view in a Spanish context [AICLE en Educación Infantil: un análisis cualitativo de las opiniones del profesorado en el contexto español]. En J. Rodas Segarra, S. Mengual-Andres y A. Cortijo Ocaña. (eds.), *Educational innovation to address complex societal challenges*, 121-136. IGI Global. 10.4018/979-8-3693-3073-9.CH009
- Penfield, R. D., y Giacobbi, P. R. (2004). Applying a score confidence interval to Aiken's item content-relevance index [Aplicación de un intervalo de confianza de puntuación al índice de relevancia del contenido de los elementos de Aiken]. *Measurement in Physical Education and Exercise Science*, 8(4), 213-225. https://doi.org/10.1207/s15327841mpee0804_3
- Pérez-Contreras, J., Bahamonde-Pérez, C., Pardo-Tamayo, C., Merino-Muñoz, P., y Aedo-Muñoz, E. (2022). Validación y fiabilidad del instrumento “Escala de percepción de competencias docente en educación física en la atención de las personas en situación de discapacidad”. *Retos*, (43), 931-937. <https://doi.org/10.47197/retos.v43i0.86477>
- Perrenoud, P. (2004). *Diez nuevas competencias para enseñar*. Editorial Grao.
- Putra, E., y Yanto, M. (2025). Classroom management: boosting student success –a meta-analysis review [Gestión del aula: Impulsando el éxito estudiantil: una revisión de metaanálisis]. *Cogent Education*, 12(21). <https://doi.org/10.1080/2331186X.2025.2458630>
- R Core Team. (2020). *R: A language and environment for statistical computing* [R: Un lenguaje y un entorno para la computación estadística]. R Foundation for Statistical Computing. R-project.org/.

- Ráczová, B., Kačmár, P., y Hricová, M. (2021). A psychometric evaluation and initial validation of the Slovak version of the Goal Adjustment Scale (GAS) [Evaluación psicométrica y validación inicial de la versión eslovaca de la Escala de Ajuste de Objetivos (GAS)]. *Studia Psychologica*, 63(1), 94-109. <https://journals.savba.sk/index.php/studiapsychologica/article/view/82/53>
- Ramón-Santiago, P., Aquino-Zúñiga, S. P., y Alejandro-Jiménez, M. (2017). Evaluación de las competencias docentes de profesores de Educación Media Superior del Estado de Tabasco, México. *Sinéctica Revista Electrónica de Educación*, (47), 1-17. <https://www.scielo.org.mx/pdf/sine/n48/2007-7033-sine-48-00013.pdf>
- Revelle, W. (2017). *Psych: Procedures for personality and psychological research R package version 2.1.6* [psych: Procedimientos para la investigación psicológica y de la personalidad]. <https://cran.rproject.org/package=psych>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling [lavaan: Un paquete de R para el modelado de ecuaciones estructurales]. *Journal of Statistical Software*, 48(2), 1-36. <https://doi.org/10.18637/jss.v048.i02>
- Salihu, A. G. (2019). Assessing teachers ability on test construction and economics content validity in Nasawara State Senior Secondary School, Nigeria [Evaluación de la capacidad de los docentes en la construcción de exámenes y la validez del contenido económico en las escuelas secundarias superiores del estado de Nasarawa, Nigeria]. *International Journal of Innovative Research in Education, Technology y Social Strategies*, 6(1), 1-16. <https://internationalpolicybrief.org/wp-content/uploads/2023/10/ARTICLE1-73.pdf>
- Sánchez-Tarazaga, L., y Fernández-Berruero, R. (2022). Validación de un cuestionario de competencias docentes para el profesorado de educación secundaria. *Tendencias Pedagógicas*, (39), 95-105. <https://doi.org/10.15366/tp2022.39.008>
- Singh, A. (2024). Status of teaching competency among pupil teachers [Situación de la competencia docente entre los alumnos docentes]. *Educational Administration: Theory and Practice*, 30(1), 3988-4003. <https://doi.org/10.53555/kuey.v30i1.7709>
- Skedsmo, G., y Huber, S. G. (2024). Navigating data, evaluation, and incentives to improve teaching and instructional quality [Navegar por los datos, la evaluación y los incentivos para mejorar la calidad de la enseñanza y la instrucción]. *Educ Asse Eval Acc*, 36, 429-432. <https://doi.org/10.1007/s11092-024-09447-4>
- Tabachnick, B. G., y Fidell, L. S. (2001). *Using multivariate statistics* [Uso de la estadística multivariante]. (4ta. ed.). Allyn and Bacon
- UNESCO. (2022). *Transforming education from within: current trends in the status and development of teachers; World teachers' day 2022* [Transformar la educación desde adentro: Tendencias actuales en la situación y el desarrollo de los docentes: Día Mundial de los docentes 2022]. <https://unesdoc.unesco.org/ark:/48223/pf0000383002>
- Uysal, C., Sani-Bozkurt, S., Bozkus-Genc, G., y Gurgur, H. (2021). Empowering teachers who work in inclusive practices: E-coaching [Empoderar a los docentes que trabajan en prácticas inclusivas]. En A. Bozkurt (ed.). *Handbook of research on emerging pedagogies for the future of education: Trauma-informed, care, and pandemic pedagogy* [Manual de investigación sobre pedagogías emergentes para el future de la educación: pedagogía basada en el trauma, la atención y la pandemia], 386-400. IGI Global. 10.4018/978-1-7998-7275-7.ch020
- Velicer, W. F., y Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery [Efectos del muestreo de variables y sujetos en la recuperación de patrones de factores]. *Psychological Methods*, 3(2), 231-251. <https://doi.org/10.1037/1082-989X.3.2.231>
- Wang, H., y Sang, L. (2024). Interdisciplinary competence of primary and secondary school teachers: a systematic literature review [Competencia interdisciplinaria de profesores de primaria y secundaria: una revisión sistemática de la literatura]. *Cogent Education*, 11(1), 1-16. <https://doi.org/10.1080/2331186X.2024.2378277>
- WHO. (2022). *Global competency and outcomes framework for universal health coverage* [Marco global de competencias y resultados para la cobertura sanitaria universal]. <https://iris.who.int/bitstream/handle/10665/352711/9789240034662-eng.pdf?sequence=1>

- Wong, S. Ch. (2020). Competency definitions, development and assessment: a brief review [Definiciones, desarrollo y evaluación de competencias: una breve revisión]. *International Journal of Academic Research in Progressive Education and Development*, 9(3), 95-114. <http://dx.doi.org/10.6007/IJARPED/v9-i3/8223>
- Worthington, R. L., y Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. The counseling psychologist [Investigación de desarrollo a escala: un análisis de contenido y recomendaciones para las mejores prácticas], *The Counseling Psychologist*, 34(6), 806-838. <https://doi.org/10.1177/0011000006288127>
- Xia, Y., y Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods [RMSEA, CFI y TLI en el modelado de ecuaciones estructurales con datos categóricos ordenados: La historia que cuentan depende de los métodos de estimación]. *Behavior Research Methods*, 51(1), 409-428. <https://doi.org/10.3758/s13428-018-1055-2>
- Yuan, K. H. (2005). Fit indices versus test statistics [Índices de ajuste frente a estadísticas de prueba]. *Multivariate Behavioral Research*, 40(1), 115-148. https://doi.org/10.1207/s15327906mbr4001_5
- Zhou, Q., Li, M., Yan, Y., Chen, M., Wang, J., Yi, K., Han, X., Wang, S., y Wang, S. (2023). Teaching competency development [Desarrollo de Competencias Docentes]. En J. Cheng, W. Han, Q. Zhou y S. Wang (Eds.). *Handbook of teaching competency development in higher education [Manual de desarrollo de competencias docentes en educación superior]*, 63-126. <https://library.oapen.org/bitstream/id/05aa1d2c-93fc-4ce2-8532-b08b-b15622ee/978-981-99-6273-0.pdf>
- Zumstein-Shaha, M., y Grace, P. J. (2023). Competency frameworks, nursing perspectives, and interdisciplinary collaborations for good patient care: Delineating boundaries [Marcos de competencias, perspectivas de enfermería y colaboraciones interdisciplinarias para una buena atención al paciente: Delineando límites]. *Nursing Philosophy*, 24(1), e12402. <https://doi.org/10.1111/nup.12402>

Biografías

Dr. Emerson López-López: Doctor en Socioformación y Sociedad del Conocimiento por el Centro Universitario CIFE (emerson.lopezqulv.edu.mx). Profesor investigador de licenciatura en Educación y Contaduría Pública, así como en la maestría en Educación en la Universidad Linda Vista en Chiapas, México. Sus líneas de investigación son: Formación de investigadores, educación basada en competencias, educación financiera, currículum y didáctica.

 <https://0000-0001-6149-7856>

Dr. Luis Gibran Juárez-Hernández: Doctor en Ciencias Biológicas y de la Salud por la Universidad Autónoma Metropolitana (gibbjuarez@gmail.com). Medalla al Mérito Universitario por Estudios de Doctorado. Profesor en la Maestría en Formación e Innovación para Profesionales de la Salud de la Facultad de Medicina y Ciencias Biomédicas de la Universidad Autónoma de Chihuahua y profesor-investigador del Centro Universitario CIFE. Es integrante del Sistema Nacional de Investigadoras e Investigadores. Sus líneas de investigación son el diagnóstico ecológico, desarrollo y evaluación de instrumentos de evaluación, desarrollo sostenible.

 <https://orcid.org/0000-0003-0658-6818>