

UNA REVISIÓN SISTEMÁTICA DE INSTRUMENTOS PSICOMÉTRICOS PARA EVALUAR EL TRABAJO EMOCIONAL EN DOCENTES

A SYSTEMATIC REVIEW OF PSYCHOMETRIC INSTRUMENTS FOR ASSESSING EMOTIONAL LABOR IN TEACHERS

 **Kathia Pamela Guadalupe Sosa-Sánchez, Mgtr.**

Universidad Juárez Autónoma de Tabasco

ppameh@hotmail.com

Villahermosa, México

 **Levit Emmanuel de los Santos Colorado, Ph. D.**

Universidad Juárez Autónoma de Tabasco

levit79@hotmail.com

Villahermosa, México

ARTÍCULO DE REVISIÓN

Recibido: 09/06/2025

Aceptado: 21/06/2025

Publicado: 30/06/2025

RESUMEN

El trabajo emocional (TE), conceptualizado como el conjunto de procesos mediante los cuales los docentes regulan, suprimen, modifican o alinean sus emociones con las normas organizacionales explícitas o implícitas para desempeñar el rol emocional esperado, constituye un ámbito aún limitado de exploración dentro de la investigación sobre las emociones en el profesorado. El objetivo de la presente revisión sistemática fue identificar y analizar los instrumentos psicométricos validados para la medición del TE en docentes, considerando sus propiedades psicométricas y adaptaciones a distintos contextos socioculturales. Se realizó una búsqueda exhaustiva de literatura publicada en inglés y español durante los últimos 25 años en las bases de datos Scopus, Web of Science, ERIC, Redalyc y PsycInfo, complementada mediante la técnica de bola de nieve, bajo estrictos criterios de inclusión y exclusión. Se identificaron 21 estudios que reflejan un interés internacional creciente en la evaluación del TE, con un predominio de investigaciones en China y Estados Unidos. La mayoría de los instrumentos revisados se estructuran en torno a las dimensiones de actuación superficial, actuación profunda y expresión genuina, y presentan niveles aceptables de validez y fiabilidad. No obstante, se observaron variaciones en la inclusión de subdimensiones y en la operacionalización de componentes específicos del constructo. Los hallazgos indican que, si bien las herramientas psicométricas disponibles resultan útiles para la evaluación del TE docente, persiste la necesidad de ampliar la conceptualización teórica, metodológica y cultural con el fin de favorecer una comprensión más integral y contextualizada de este fenómeno.

Palabras Clave: trabajo emocional, instrumentos, escala, medición, docentes

ABSTRACT

Emotional labor (EL), defined as the set of processes through which teachers regulate, suppress, modify, or align their emotions with explicit or implicit organizational norms in order to fulfill the expected emotional role, remains a relatively limited area of investigation within teacher emotion research. The purpose of this systematic review was to identify and analyze validated psychometric instruments used to assess EL in teachers, considering their psychometric properties and adaptations across different sociocultural contexts. An extensive literature search was conducted in both English and Spanish, covering the past 25 years, across the Scopus, Web of Science, ERIC, Redalyc, and PsycInfo databases, and complemented by the snowball sampling technique, following strict inclusion and exclusion criteria. A total of 21 studies were identified, reflecting growing international interest in the evaluation of EL, with a predominance of research conducted in China and the United States. Most of the reviewed instruments are organized around the dimensions of surface acting, deep acting, and genuine expression, and report acceptable levels of validity and reliability. Nevertheless, variations were observed regarding the inclusion of subdimensions and the operationalization of specific components of the construct. The findings suggest that, although the available psychometric tools are useful for assessing teachers' EL, further development of theoretical, methodological, and cultural frameworks is needed to achieve a more comprehensive and contextualized understanding of this phenomenon.

Keywords: emotional labor, instruments, scale, measurement, teachers

INTRODUCCIÓN

Los estudios sobre el TE constituyen un campo emergente dentro de la investigación de las conductas y los comportamientos laborales (Ashkanasy et al., 2002). En su formulación inicial, Hochschild (1983) definió el TE como la regulación de las emociones con el propósito de crear una expresión facial y corporal observable públicamente, la cual posee un valor de intercambio a través del salario. Aunque esta connotación se refería, habitualmente, a las interacciones de los trabajadores del sector servicios con los clientes, en la actualidad el estudio del TE se ha extendido a otros ámbitos, como el educativo. Así, su análisis suele centrarse en las relaciones que el personal docente establece con diversos actores educativos, incluidos colegas, directivos, estudiantes y padres de familia (Chen et al., 2025; De-Ruiter et al., 2021; Dickerson et al., 2024; Zheng et al., 2020) puesto que los docentes trabajan en entornos que son cada vez más

emocionalmente exigentes (Carroll et al., 2020; Donker et al., 2020; Heffernan y Bosetti, 2020; Zhang et al., 2024).

Se ha demostrado que la exposición docente al TE conlleva efectos bivariados sobre su bienestar, salud mental y desempeño laboral. Por un lado, se han identificado consecuencias negativas, tales como el agotamiento emocional, la despersonalización y el burnout, particularmente asociadas con la actuación superficial (Kariou et al., 2021; Ma et al., 2023; Yin et al., 2019). Por otro lado, también se han reportado efectos positivos, como un mayor compromiso laboral, satisfacción profesional y fortalecimiento del vínculo pedagógico, especialmente vinculados con la actuación profunda y la expresión genuina de emociones (W. Peng et al., 2023; Taheri et al., 2024; Wang et al., 2021).

Para comprender esta aparente contradicción en los hallazgos del TE en docentes, es preciso distinguir las tres dimensiones que integran el constructo. La primera de ellas, la actuación superficial (surface acting) implica regular las expresiones emocionales externas por medio de estrategias como el fingimiento, la supresión o el ocultamiento, sin que ello suponga una transformación interna auténtica en el sujeto. En contraste, la actuación profunda (deep acting) se refiere al proceso mediante el cual el individuo procura generar congruencia entre las emociones que siente y las que se espera que exprese como parte de su rol, recurriendo a diversas estrategias de alineación afectiva. Finalmente, la expresión de emociones genuinas (genuine emotions) alude a aquellas respuestas emocionales espontáneas y auténticas que surgen de manera natural en el contexto laboral (Ashforth y Humphrey, 1993).

Dado que estas dimensiones operan en interacción con múltiples determinantes, resulta fundamental examinar los factores que configuran el TE en el ámbito docente. En este sentido, el TE constituye un fenómeno complejo modulado por variables individuales y contextuales que inciden en las estrategias emocionales adoptadas. En el nivel individual, características como los rasgos de personalidad, la afectividad positiva y negativa, el agotamiento emocional y la inteligencia emocional influyen directamente en la selección de dichas estrategias. Paralelamente, a nivel organizacional, las normas emocionales explícitas e implícitas, los estilos de liderazgo y el clima laboral establecen los marcos normativos que regulan las expresiones emocionales esperadas (Wang et al., 2019). Además, el género introduce diferencias significativas, dado que la evidencia empírica ha identificado patrones diferenciados en la vivencia y expresión del trabajo emocional entre mujeres y hombres, lo cual puede tener implicaciones relevantes en términos de salud mental y desempeño profesional (Grandey y Gabriel, 2015).

PLANTEAMIENTO DEL PROBLEMA

Considerando la diversidad de factores que influyen en la configuración del TE, así como las diferencias de género en los tipos de estrategias emocionales preferidas, se demandan herramientas que sean sensibles a estas variaciones. Sin embargo, a pesar del creciente cuerpo de investigaciones empíricas, subsiste un vacío importante relacionado con los procedimientos de medición del TE en docentes, por tanto, resulta necesario contar con instrumentos psicométricos que permitan medirlo con precisión y validez.

La naturaleza multidimensional y compleja del TE plantea varias preocupaciones respecto a su evaluación, operacionalización y los instrumentos de medición empleados. Como advierte Olmsted (2024), la utilización de herramientas de recolección de datos sin propiedades psicométricas sólidas compromete la validez, fiabilidad y generalización de los hallazgos, generando resultados erráticos y difíciles de replicar.

Estas limitaciones metodológicas afectan tanto el avance teórico como a la capacidad de generar recomendaciones prácticas sólidas para el acompañamiento y desarrollo profesional docente. Por ello, la adecuada evaluación del TE reviste especial relevancia para investigadores, académicos y profesionales del ámbito educativo, quienes requieren de instrumentos validados y adaptados metodológicamente que sustenten la producción de conocimiento empírico riguroso (Howard, 2023).

En ese sentido, la presente revisión busca atender este vacío mediante el mapeo y análisis de los instrumentos validados para la medición del TE en docentes, examinando en detalle sus propiedades psicométricas, su validez, fiabilidad y aplicabilidad en diferentes contextos educativos. Los resultados de este trabajo permitirán no solo clarificar el estado actual de las herramientas disponibles, sino también destacar las fortalezas y limitaciones de estos instrumentos, y proporcionar recomendaciones para futuras investigaciones que deseen profundizar en el estudio de este fenómeno complejo, contribuyendo al mismo tiempo a la selección de herramientas válidas y confiables para futuras investigaciones e intervenciones.

METODOLOGÍA

Se llevó a cabo un análisis sistemático de literatura siguiendo las directrices de la declaración PRISMA (Page et al., 2021), sustentando la investigación en las siguientes preguntas específicas:

P₁. ¿Qué características bibliométricas se identifican en los artículos analizados?

P₂. ¿Qué instrumentos para medir el TE se han utilizado?

P₃. ¿A qué poblaciones se ha aplicado el instrumento original?

P₄. ¿Cuál es el número de ítems y los tipos de escalas más comúnmente utilizados?

P₅. ¿Cuáles son las dimensiones del TE que evalúan los instrumentos analizados?

P₆. ¿Qué propiedades psicométricas se evidencian en los instrumentos y qué métodos se utilizaron para el análisis de las propiedades psicométricas?

Criterios de inclusión y exclusión

La búsqueda consistió en recopilar publicaciones que cumplieran con los criterios de inclusión y exclusión detallados en la Tabla 1.

Tabla 1

Criterios de inclusión y exclusión de la revisión sistemática

Criterios de Inclusión	Criterios de Exclusión
Artículos cuantitativos publicados en su versión final en revistas arbitradas por pares	Artículos cualitativos, de revisión y literatura gris
Evaluar el trabajo emocional a nivel global o en un solo dominio	No estar validado en profesionales de la educación
Reportar índices de fiabilidad, validez o aceptabilidad	No detallar las propiedades psicométricas del instrumento
Estudios publicados entre 2000-2025	Estar escritos en idioma distinto al inglés o al español
Pertenecer a las bases de datos Scopus, WoS, ERIC, Redalyc o PsycInfo	Publicaciones sin texto completo

Nota. Elaboración propia.

Estrategia de búsqueda

Los términos de búsqueda fueron desarrollados mediante un proceso iterativo como lo propone Terwee et al. (2009) con el uso de operadores booleanos, quedando las cadenas de búsqueda, filtros y criterios empleados a como se especifica en la Tabla 2. En concordancia con las buenas prácticas de transparencia y reproducibilidad, el historial completo de búsqueda y el set de datos iniciales exportados han sido depositados en el repositorio de acceso abierto Zenodo y se encuentra disponible en <https://doi.org/10.5281/zenodo.15713503>.

Proceso de selección de estudios

Con la finalidad de eliminar los registros duplicados, los resultados de la búsqueda fueron importados al programa Zotero (versión 7.0.15). Posteriormente dos revisores examinaron de forma independiente los títulos y resúmenes de todos los registros identificados y evaluaron los textos completos de todos los artículos potencialmente elegibles. Cualquier desacuerdo entre los dos revisores se resolvió mediante consulta con un investigador experto. Con el objeto de considerar todos los estudios potencialmente relevantes, se aplicó la técnica de bola de nieve en la búsqueda manual de las listas de referencias de los artículos incluidos para garantizar su inclusión.

Tabla 2

Cadenas de búsqueda, filtros y criterios empleados por base de datos

Base de datos	Cadena de búsqueda
Scopus	(TITLE-ABS-KEY ("emotional labor" OR "emotional labour" OR "surface acting" OR "deep acting" OR "genuine emotions") AND TITLE-ABS-KEY (education* OR kindergarten* OR school* OR university* OR teacher* OR educator* OR professor* OR lecturer* OR director* OR principal* OR viceprincipal* OR headmaster*) AND KEY (assessment* OR measure* OR evaluation* OR validation* OR metric* OR outcome* OR pilot OR instrument* OR scale* OR questionnaire* OR psychometric*)) AND PUBYEAR > 1999 AND PUBYEAR < 2026 AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (SUBJAREA , "SOC") OR LIMIT-TO (SUBJAREA , "MED")) OR LIMIT-TO (SUBJAREA , "NURS") OR LIMIT-TO (SUBJAREA , "PSYC")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English") OR LIMIT-TO (LANGUAGE , "Spanish"))
Web of Science	(TS=("emotional labor" OR "emotional labour" OR "surface acting" OR "deep acting" OR "genuine emotions") AND AK=(education* OR kindergarten* OR school* OR university* OR teacher* OR educator* OR professor* OR lecturer* OR director* OR principal* OR viceprincipal* OR headmaster*) AND TS=(assessment* OR measure* OR evaluation* OR validation* OR metric* OR outcome* OR pilot OR instrument* OR scale* OR questionnaire* OR psychometric*)) AND (PY=("2025" OR "2024" OR "2023" OR "2022" OR "2015" OR "2016" OR "2017" OR "2018" OR "2019" OR "2020" OR "2021" OR "2004" OR "2005" OR "2006" OR "2007" OR "2008" OR "2009" OR "2010" OR "2011" OR "2012" OR "2013" OR "2014" OR "2003" OR "2002" OR "2001" OR "2000") AND DT=("ARTICLE") AND LA=("ENGLISH" OR "SPANISH") AND DT=("ARTICLE") AND TASC=("EDUCATION EDUCATIONAL RESEARCH" OR "PSYCHOLOGY EDUCATIONAL" OR "PSYCHOLOGY MULTIDISCIPLINARY" OR "EDUCATION SCIENTIFIC DISCIPLINES" OR "MANAGEMENT"))
ERIC	("emotional labor" OR "emotional labour" OR "surface acting" OR "deep acting" OR "genuine emotions") AND (assessment* OR measure* OR evaluation* OR validation* OR metric* OR outcome* OR pilot OR instrument* OR scale* OR questionnaire* OR psychometric*) AND pubyearmin: 2000 Filtros adicionales: Peer reviewed only, Full text available on ERIC, Publication Type = Journal Article, Reports - Research
Redalyc	"emotional labor" AND (measurement OR instrument OR scale) Filtros adicionales: Disciplina = Psicología/Educación, Lenguaje = Español
PsycInfo	Búsqueda por campos title, abstract, full text, caption text, author keywords = ("emotional labor" OR "emotional labour" OR "surface acting" OR "deep acting" OR "genuine emotions") AND (assessment* OR measure* OR evaluation* OR validation* OR metric* OR outcome* OR pilot OR instrument* OR scale* OR questionnaire* OR psychometric*)

Extracción de datos

Una vez completado el proceso de selección, se analizaron en profundidad los 21 artículos, extrayendo datos de interés para el presente estudio, incluyendo en una tabla de Excel los siguientes elementos: (a) las características bibliométricas de los estudios como el título, resultados, autores, año de publicación, resumen, revista de publicación; (b) las características de los instrumentos como el nombre del instrumento, dimensiones, población, país, número de ítems, escala de respuestas de ítems y (c) las propiedades psicométricas del instrumento, es decir, las técnicas estadísticas empleadas para determinar la validez y confiabilidad del mismo.

RESULTADOS

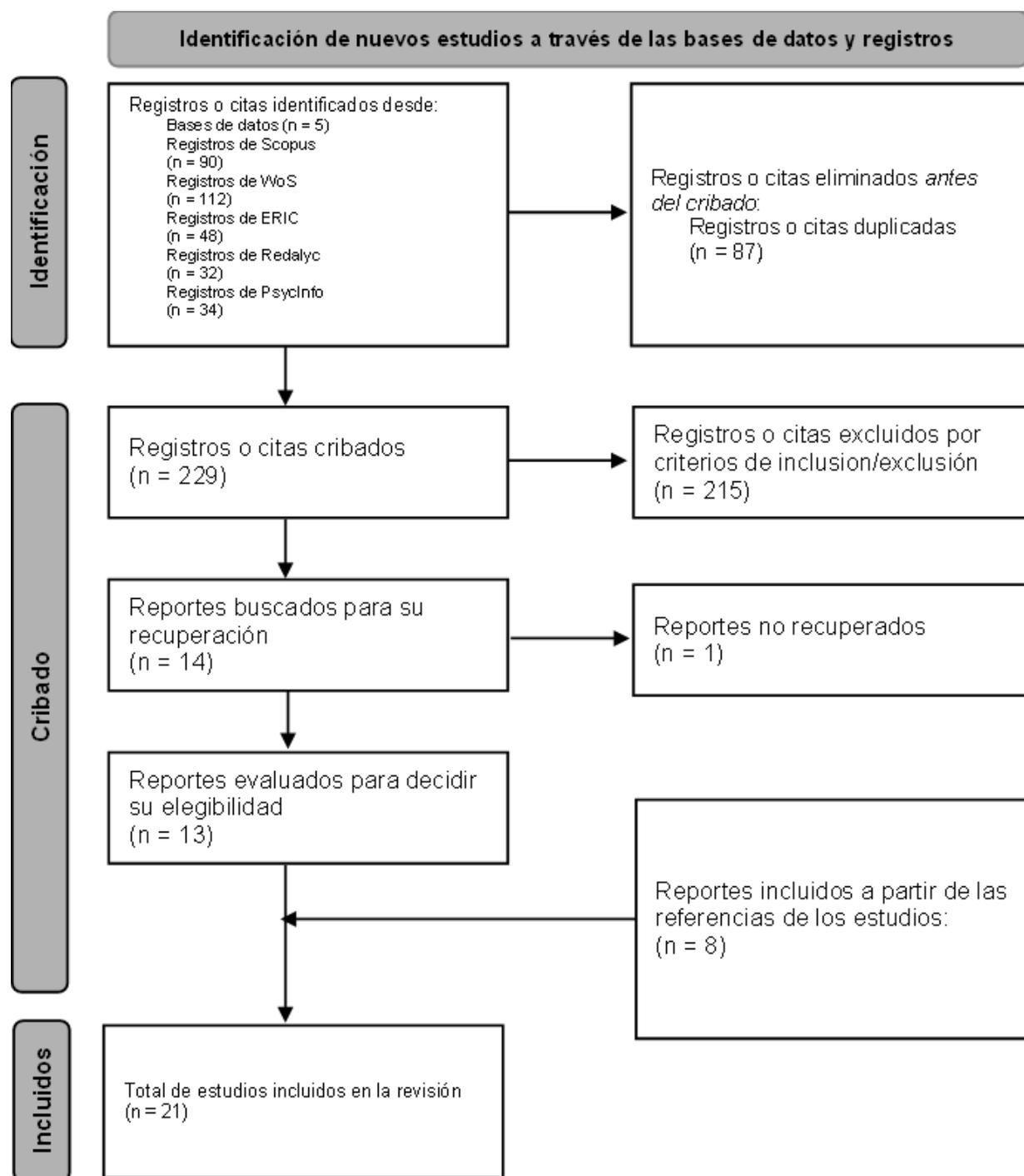
Se identificaron inicialmente 316 registros a través de la estrategia de búsqueda sistemática. Posteriormente, tras la eliminación de 87 registros duplicados, se retuvieron 229 artículos para la fase de cribado de títulos y resúmenes. Durante esta etapa, 215 artículos fueron excluidos por no cumplir con los criterios de inclusión establecidos. Como resultado, 14 artículos fueron seleccionados para la evaluación a texto completo. Sin embargo, uno de estos no pudo ser recuperado, por lo que finalmente se evaluaron 13 estudios en esta fase. Además, se incorporaron 8 registros adicionales identificados mediante búsqueda manual a través de la técnica de bola de nieve, al ser considerados relevantes y cumplir con los criterios de elegibilidad. En consecuencia, el corpus final de estudios incluidos en la presente revisión sistemática quedó conformado por 21 artículos. El proceso de identificación, selección y elegibilidad de los estudios se presenta en la Figura 1.

En primer lugar, el análisis bibliométrico (P_1) reveló un volumen moderado de publicaciones centradas en el TE dentro de los ámbitos educativo y psicológico. Aunque desde el año 2000 hasta el 2019 el número de publicaciones se mantuvo limitado a una o dos por año, se observó un incremento a partir de 2021, como se ilustra en la Figura 2.

En cuanto a la distribución geográfica de los estudios, los datos indican que China ($n = 7$) y Estados Unidos ($n = 4$) constituyen los principales focos de producción científica sobre el TE, dentro de un total de nueve países representados, denotando un liderazgo predominante de ambos países en el desarrollo de esta línea de investigación. Con respecto al idioma de publicación, se encontraron $n = 18$ publicaciones en el idioma inglés, seguido del español con $n = 3$ artículos.

Figura 1

Flujograma del proceso de revisión sistemática



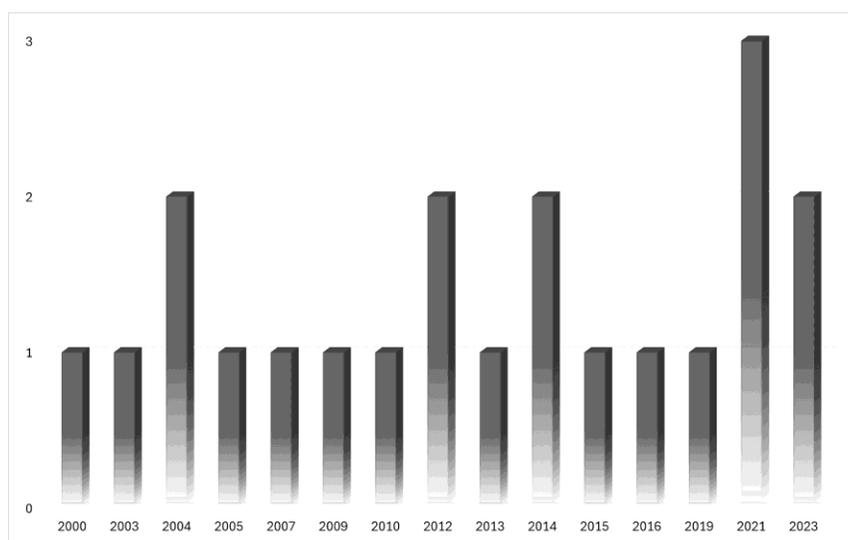
Nota. Adaptado de "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews", de Page et al., 2021, BMJ, 372(n71), p. 5 (<https://doi.org/10.1136/bmj.n71>). Licencia CC BY 4.0.

Del mismo modo, la mayor concentración de estudios empíricos se registró en las revistas *Frontiers in Psychology*, *Journal of Vocational Behavior* y *Psicothema* con dos publicaciones cada una, mientras que el resto de la literatura se encontraba dispersa con un máximo de un artículo por revista. En lo que respecta a los instrumentos de medida (P_2), la herramienta más utilizada y preferida en los estudios cuantitativos es la escala ($n = 18$), seguida del cuestionario ($n = 3$) con datos recopilados de manera auto informada.

Al analizar las poblaciones en las cuales fueron originalmente validados los instrumentos (P_3), se observó que el tamaño muestral osciló entre los 37 y 1,975 participantes. La mayoría de los instrumentos fueron aplicados en contextos laborales ajenos al sector educativo.

Figura 2

Distribución temporal del número de publicaciones



Nota. Elaboración propia.

Específicamente, las validaciones iniciales se llevaron a cabo en profesionales dedicados a la atención directa al público, tales como vendedores, masajistas, cajeros, personal de enfermería, así como en los sectores hotelero, restaurantero y de salud.

Sin embargo, cabe señalar que algunas de estas escalas, como la *Emotional Labour Scale (ELS)*, la *Emotional Labor Strategies Scale (ELSS)*, la *Discrete Emotions Emotional Labor Scale (DEELS)* y la *Frankfurt Emotion Work Scale (FEWS)*, han demostrado ser válidas y replicables en contextos educativos.

De forma paralela, se identificaron instrumentos diseñados específicamente para el sector educativo. Entre ellos destacan la *Teacher Emotional Labor Scale (TELS)*, aplicada en docentes de nivel secundaria; el *Emotional Labor Questionnaire for Primary and Secondary School Teachers (ELQ-PST)* y la *Teacher Emotional Labour Strategy Scale (TELSS)*, ambos validados en profesores de primaria y secundaria. Asimismo, se hallaron escalas adaptadas a contextos educativos particulares, como la *Teachers' Emotional Labor Strategy in Classrooms (TELSC)*, orientada a la enseñanza presencial dentro del aula, y la *Online Teaching Emotional Labor Strategies Scale (OTELSS)*, desarrollada para el ámbito virtual.

Además, se identificó un conjunto de instrumentos dirigidos a subpoblaciones docentes específicas. Por un lado, los maestros de inglés como lengua extranjera (EFL) cuentan con dos escalas propias: la *Chinese EFL Teacher Emotional Labor Scale (C-EFL-TELS)* y la *Beginning EFL Teachers' Emotional Labor Strategy Scale (B-EFL-TELSS)*. Por otro lado, se halló la *Emotional Labor Scale among Turkish school administrators (ELS-SA)*, validada para personal directivo de instituciones educativas.

En relación con la estructura de los instrumentos (P₄), se identificó que el número de ítems varió de 6 a 27 (M = 15.71; Me = 16), con un formato de respuesta, predominantemente basado en la escala Likert de cinco puntos (n = 12), seguida por la de siete puntos (n = 3); en el resto de los casos, no se especificó el número de opciones de respuesta.

En cuanto a los aspectos logísticos de aplicación, se observó que la gran mayoría de los estudios no reportaron la duración estimada para completar los instrumentos (n = 19). De los pocos estudios que proporcionaron esta información, el tiempo de administración fluctuó entre 12 y 20 minutos.

Finalmente, en términos del diseño temporal de los estudios, se constató que todos los trabajos analizados emplearon diseños transversales, con la única excepción del *Daily Emotional Labor Scale (DELS)* desarrollado por De-Ruiter et al. (2021), que corresponde a un diseño longitudinal.

Los 21 instrumentos analizados (ver Apéndice) en relación con su dimensionalidad (P₅) presentan, en su mayoría, una estructura interna sustentada en las tres dimensiones clásicas propuestas en la literatura: actuación superficial, actuación profunda y expresión genuina de las emociones. No obstante, algunos autores han ampliado este modelo incorporando subdimensiones adicionales que permiten una evaluación más matizada del fenómeno (Figura 3). Por ejemplo, se han incluido la disonancia emocional, la amplificación expresiva y la supresión

como subdimensiones de la actuación superficial, así como el despliegue atencional y la reevaluación dentro de la actuación profunda (Andela et al., 2015).

Del mismo modo, ciertos instrumentos han integrado en su medición aspectos relacionados con la temporalidad y la naturaleza del TE, considerando dimensiones como la frecuencia, la intensidad, la variedad y la duración de los episodios emocionales (Brotheridge y Lee, 2003). Otras herramientas, como la propuesta por Çukur (2009) añade la regulación automática de las emociones y la desviación emocional basándose en la Teoría de los Eventos Afectivos de Weiss y Cropanzano (1996), según la cual las estrategias de TE son una reacción de los empleados ante los eventos emocionales que experimentan en su lugar de trabajo.

Figura 3

Dimensiones agregadas en la instrumentación del trabajo emocional



Nota. Elaboración propia.

Otros aspectos agregados son el esfuerzo emocional (Kruml y Geddes, 2000) y la consonancia positiva y negativa (Li y Liu, 2021), este último aspecto hace referencia a que la consonancia positiva se presenta cuando las emociones expresadas son congruentes con las

emociones auténticamente experimentadas por el individuo. En contraste, la consonancia negativa describe situaciones en las cuales existe una correspondencia entre las emociones sentidas y expresadas; sin embargo, dichas emociones resultan incompatibles con las normas emocionales prescritas por el rol laboral. En tal sentido, y desde una óptica organizacional Moreno-Jiménez et al. (2010) propone medir las normas organizacionales sobre las emociones, los requerimientos emocionales del puesto y la sensibilidad exigida. Todo ello pone de manifiesto tanto la diversidad de aproximaciones conceptuales como la multidimensionalidad del constructo.

Para abordar la P₆, resulta indispensable establecer que la confiabilidad se refiere al grado en que las puntuaciones obtenidas mediante un instrumento están libres de errores de medición, mientras que la validez se entiende como la capacidad del instrumento para evaluar de forma precisa el constructo de interés (Hogan, 2015). A partir de esta premisa, los resultados obtenidos permiten identificar patrones comunes, así como discrepancias relevantes en los procedimientos psicométricos empleados.

En términos generales, todos los estudios revisados reportaron coeficientes alfa de Cronbach (α) superiores al umbral aceptado de 0.70, lo que, en principio, sugiere una adecuada consistencia interna de los instrumentos. Sin embargo, un análisis más detallado revela ciertas limitaciones en dimensiones específicas. Concretamente, los trabajos de Dewaele y Wu (2021) y Yang et al. (2019) informaron coeficientes por debajo de este estándar en las dimensiones de actuación profunda ($\alpha = 0.54$ y $\alpha = 0.30$, respectivamente) y terminación emocional ($\alpha = 0.56$). Este hallazgo indica que, si bien los instrumentos presentan globalmente niveles aceptables de confiabilidad, algunas subescalas podrían presentar problemas de consistencia interna, lo que obliga a interpretar sus resultados con la debida cautela.

En lo que concierne a los procedimientos de validez, cabe destacar que la validación de contenido fue abordada sistemáticamente en la mayoría de los estudios mediante análisis estadísticos de los ítems. El Análisis Factorial Confirmatorio (AFC) emergió como la técnica predominante para examinar la estructura interna de los instrumentos ($n = 6$), seguido por el Análisis Factorial Exploratorio (AFE) ($n = 4$). Además, resulta relevante que diez investigaciones aplicaron ambos procedimientos de manera complementaria, lo cual fortalece la robustez de sus hallazgos estructurales. Únicamente un estudio verificó la estructura interna mediante correlación intraclase, lo que denota escasa diversidad metodológica en este aspecto.

De manera análoga, los análisis de validez externa se centraron principalmente en la validez discriminante ($n = 3$) y convergente ($n = 3$), aspectos esenciales para garantizar que los instrumentos diferencien adecuadamente entre constructos relacionados pero distintos, y que

correlacionen con otros indicadores conceptualmente afines. No obstante, es particularmente destacable la escasa atención otorgada a la validez de criterio, la cual fue abordada únicamente en dos estudios. Esta limitación metodológica restringe el conocimiento disponible respecto al desempeño predictivo o concurrente de los instrumentos en escenarios prácticos y de intervención.

Con el fin de sintetizar la calidad de la evidencia disponible en relación con las propiedades psicométricas de los instrumentos analizados, se procedió a la construcción de la Tabla 3, la cual ofrece una visión sintetizada de la heterogeneidad de los estudios revisados. Como se detalla, se evaluaron de forma sistemática once propiedades de medición ampliamente aceptadas en la literatura metodológica: validez de contenido, validez estructural, consistencia interna, validez transcultural, medición de invariancia, fiabilidad, medición del error, validez de criterio, pruebas de hipótesis para la validez del constructo y tasa de respuesta.

Tabla 3

Calidad de la evidencia de las propiedades de medición de los instrumentos

Aspecto/ Instrumento	(Andela et. al, 2015)	(Blau et al., 2010)	(Brotheridge y Lee, 2003)	(Çukur, 2009)	(Dewaele y Wu, 2021)	(Diefendorff et al., 2005)	(Glomb y Tews, 2004)	(Kiral, 2016)	(Kruml y Geddes, 2000)	(Li y Liu, 2021)	(Liu y Zhang, 2014)	(Ma et al., 2023)	(Moreno-Jiménez et al., 2004)	(Näring et al., 2007)	(Peng et al., 2023)	(Quiñones-García et al., 2013)	(Ruiter et al., 2021)	(Salessi y Omar, 2014)	(Yang et al., 2019)	(Yin, 2012)	(Zapf, 1999)
Relevancia	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Exhaustividad	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Comprensibilidad	+	-	±	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+
Validez estructural	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Consistencia interna	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Validez transcultural	?	?	-	?	?	+	+	?	+	?	?	?	?	?	?	?	?	?	?	?	+
Medición de invariancia	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Fiabilidad	+	+	+	+	±	+	+	+	+	+	+	+	+	+	+	+	+	+	±	+	±
Medición del error	?	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Validez del criterio	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Pruebas de hipótesis para la validez del constructo	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Tasa de respuesta	+	-	+	-	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	-

Nota. "+" = evidencia alta; "-" = evidencia suficiente; "±" = resultados inconsistentes o insuficientes; "?" = sin evidencia. Fuente: Elaboración propia.

En términos generales, los hallazgos ponen de manifiesto la existencia de variabilidad en el grado de evaluación y reporte de dichas propiedades. Aunque la mayoría de los instrumentos examinados demuestran evidencia favorable en cuanto a la validez de contenido, consistencia interna, validez estructural y validez del constructo, se identifican vacíos significativos en otros dominios metodológicos críticos. Tal es el caso de la validez transcultural y la medición de invariancia, las cuales fueron escasamente abordadas. Limitando la generalización intercultural de los instrumentos revisados. De manera similar, la medición del error y los análisis de fiabilidad test-retest o división por mitades fueron incorporados de forma marginal en los estudios analizados pues solo dos investigaciones complementaron el análisis de consistencia interna con estos métodos, lo cual sugiere una limitada evaluación de la estabilidad temporal de las mediciones.

Adicionalmente, resulta pertinente destacar que solo una fracción de los estudios aportó información sistemática sobre la tasa de respuesta, un aspecto esencial para valorar la aceptabilidad y aplicabilidad de los instrumentos en contextos de investigación aplicada.

DISCUSIÓN

El presente estudio tuvo como objetivo central identificar, sistematizar y evaluar los instrumentos psicométricos desarrollados entre los años 2000 y 2025 para la medición del TE en el profesorado. En consecuencia, los resultados obtenidos permiten señalar avances importantes en el desarrollo de instrumentos válidos y confiables, pero también revelan notables diferencias en la estandarización del TE.

En consonancia con los criterios establecidos por Sierra (2001) y Schmitt (1996), los instrumentos analizados exhiben coeficientes de confiabilidad interna ($\alpha = 0.70-0.88$) que se sitúan dentro de los rangos considerados aceptables e incluso óptimos, lo que respalda su solidez psicométrica fundamental. Desde la perspectiva de la consistencia interna, estos resultados ofrecen un sustento metodológico suficiente para la evaluación del TE en docentes. Sin embargo, ciertas escalas, como las propuestas por Dewaele y Wu (2021) y Yang et al. (2019), reportan niveles de confiabilidad reducidos en algunas subdimensiones específicas, lo cual evidencia la necesidad de realizar análisis adicionales que profundicen en la estabilidad interna de estos componentes.

Se pone de relieve que esta solidez no se extiende de manera uniforme a otras propiedades psicométricas. De manera preocupante, se observa una atención insuficiente a aspectos esenciales como la validez de criterio, la validez discriminante y la invariancia factorial. Esta

heterogeneidad metodológica restringe la capacidad de los instrumentos para capturar con precisión las complejidades contextuales del TE en escenarios educativos diversos. Así mismo, el limitado uso de procedimientos de fiabilidad temporal, como el test-retest o la división por mitades, representa otra limitación que compromete la estabilidad longitudinal de las medidas.

Por otra parte, el análisis revela que la mayoría de los estudios revisados adopta escalas tipo Likert como principal formato de medición. Este patrón puede atribuirse a la economía de su aplicación y a la posibilidad de captar de manera parsimoniosa constructos latentes (El-Den et al., 2020). Sin embargo, resulta pertinente señalar los sesgos asociados a los instrumentos auto-reportados, como la deseabilidad social en la que los participantes responden de una manera socialmente aceptable en lugar de proporcionar sus propias opiniones (Grimm, 2010); el sesgo de tendencia central que se produce debido a que los sujetos evitan las respuestas extremas optando por el punto medio para evitar las respuestas atípicas y finalmente a un sesgo de consentimiento en el que las personas muestran un acuerdo con las afirmaciones tal como se presentan para evitar controversia (Pimentel, 2019), los cuales podrían distorsionar la validez de los datos obtenidos. Bajo esta perspectiva, futuras investigaciones podrían beneficiarse de la diversificación metodológica, incorporando enfoques mixtos que complementen los autorreportes con técnicas de observación directa o mediciones fisiológicas.

En relación con la dimensionalidad, se observa unanimidad en torno al modelo tripartito de actuación superficial, actuación profunda y expresión genuina. Sin embargo, la incorporación de subdimensiones como, por ejemplo, la disonancia emocional, la amplificación expresiva o el despliegue atencional, evidencia un proceso de refinamiento teórico que persigue capturar matices situacionales y regulativos cada vez más específicos. Este avance, dificulta la comparación directa entre estudios y, por ende, la acumulación de conocimiento. Es recomendable que la comunidad científica consensue un marco integrador que precise definiciones y fronteras conceptuales, evitando solapamientos y lagunas en la operacionalización del TE.

Para quienes deseen instrumentar alguna de las escalas o cuestionarios reportados en esta investigación, es vital que consideren el tiempo de administración para juzgar la viabilidad de su aplicación, pues fue recurrente la falta de informe en la literatura psicométrica revisada. Si bien, la mayoría de los autores se inclina por escalas Likert de 5 puntos, siendo una elección que facilita la interpretación para los participantes, esto podría reducir la sensibilidad de respuesta frente a configuraciones de 7 u 8 anclajes comúnmente recomendadas para captar la variabilidad emocional fina (Koo y Yang, 2025).

Desde el enfoque del diseño investigativo, la marcada predominancia de diseños transversales apunta a la necesidad de incrementar los estudios longitudinales en la investigación, los cuales posibiliten trazar trayectorias y detectar efectos acumulativos del TE sobre el bienestar y el desempeño docente. Del mismo modo, se advierte una asimetría metodológica pues resulta imperativo que las futuras líneas de investigación amplíen el espectro de la validez más allá de la consistencia interna e incorporen evidencias derivadas de su relación con variables externas pertinentes, como el *burnout*, el *engagement* o la autoeficacia.

Otro hallazgo importante en el desarrollo de instrumentos de medición y que debiese tratarse con cautela es la adaptación cultural, pues la concentración geográfica de los estudios en los países de China y Estados Unidos, evidencia una ausencia de estudios validados en otros países. Esta distribución desigual genera potenciales sesgos culturales que limitan la extrapolación de los instrumentos a escenarios educativos con características socioculturales distintas. Para los países orientales este es un aspecto clave puesto que las escalas de Li y Liu (2021) y de Liu y Zhang (2014) incorporan ajustes y procedimientos de validación específicamente diseñados para el contexto educativo y cultural chino, lo que garantiza su pertinencia y solidez en dicho entorno. Al respecto, existe una brecha relevante en la literatura actual puesto que no se identificaron estudios que desarrollaran marcos teóricos adaptados específicamente al ámbito escolar latinoamericano.

Al integrar los hallazgos psicométricos, es posible identificar un núcleo de instrumentos que, por la solidez de su evidencia empírica, se posicionan como referentes más robustos para la medición del TE en docentes, la *Teacher Emotional Labor Scale (TELS)*, la *Teacher Emotional Labour Strategy Scale (TELSS)* y el *Emotional Labor Questionnaire for Primary and Secondary School Teachers (ELQ-PST)* destacan no solo por haber sido desarrollados específicamente para el contexto educativo, sino también por incorporar procedimientos rigurosos de validación, tales como análisis factorial confirmatorio, análisis de validez convergente y discriminante, así como consistencia interna elevada ($\alpha > 0.70$).

Además, estos instrumentos han sido aplicados en muestras docentes representativas, lo que refuerza su pertinencia contextual y su sensibilidad para captar las particularidades emocionales del ejercicio profesional docente. Por su parte, escalas originalmente diseñadas para otros ámbitos laborales, como la *Emotional Labour Scale (ELS)*, la *Emotional Labor Strategies Scale (ELSS)*, la *Discrete Emotions Emotional Labor Scale (DEELS)* y la *Frankfurt Emotion Work Scale (FEWS)*, han mostrado capacidad de adaptación y replicabilidad aceptable en población docente, siempre que se acompañen de procesos adecuados de validación transcultural y ajuste

semántico. En conjunto, estos instrumentos configuran el repertorio actualmente más respaldado desde el punto de vista psicométrico, proporcionando herramientas fiables para futuras investigaciones e intervenciones sobre el TE en el ámbito educativo.

Limitaciones y futuras direcciones

Es pertinente reconocer ciertas limitaciones inherentes al presente estudio. En primer lugar, la exclusión de la literatura gris podría haber provocado la omisión de información relevante, dado su valor informativo (Yoshida et al., 2024). Además, la exclusión de métodos cualitativos pudo haber dejado fuera evidencias contextuales que contribuyen a una comprensión más profunda del TE.

En conjunto, estos hallazgos permiten identificar una serie de desafíos para el campo: (1) la falta de integración conceptual entre los modelos teóricos y los contextos educativos específicos; (2) la necesidad de ampliar la evaluación psicométrica más allá de la confiabilidad interna; y (3) la urgencia de validar instrumentos en entornos diversos, particularmente en América Latina. Superar estas limitaciones permitirá avanzar hacia una comprensión más integral del TE docente y sus implicaciones en el bienestar profesional y la calidad educativa.

CONCLUSIONES

La revisión sistemática realizada permitió identificar y analizar un conjunto relevante de instrumentos cuantitativos validados para la evaluación del TE en docentes, publicados en las últimas dos décadas y media en contextos educativos diversos. Los resultados evidencian un creciente interés académico en la medición rigurosa de esta variable emocional, así como una variedad de escalas que contemplan múltiples dimensiones del constructo.

En relación con otras revisiones sistemáticas disponibles como la desarrollada por Yang y Chen (2021), los resultados del presente estudio coinciden en destacar la centralidad de la actuación superficial y la actuación profunda como dimensiones prioritarias. Sin embargo, también revelan la persistencia de una visión reduccionista del fenómeno que omite otras formas de regulación emocional relevantes en el contexto escolar. Esta limitación guarda relación con lo señalado por Grandey y Melloy (2017), quien advierte sobre la superposición imperfecta entre los modelos de TE y Regulación Emocional.

Mientras que el modelo de TE se centra en estrategias desplegadas para cumplir con reglas emocionales en un contexto laboral y con fines interpersonales (Brotheridge y Lee, 2003), la Regulación Emocional como lo conceptualiza Gross (1998) abarca un repertorio más amplio de estrategias cognitivas, conductuales y fisiológicas orientadas a modificar la experiencia o

expresión emocional, muchas de las cuales no han sido consideradas en los instrumentos revisados.

En este sentido, la limitada presencia de escalas que evalúan estrategias como la disonancia emocional, la amplificación expresiva, el despliegue atencional, la reevaluación, la supresión expresiva, la regulación automática de las emociones, la desviación emocional, el esfuerzo emotivo, la terminación de las emociones y la consonancia emocional (Andela et al., 2015; Çukur, 2009; Kruml y Geddes, 2000; Li y Liu, 2021; Näring et al., 2007; Quiñones-García et al., 2013) sugiere un desfase entre los avances conceptuales en el estudio del TE y los enfoques instrumentales predominantes en el ámbito educativo. Por tanto, es necesario ampliar los marcos teóricos y metodológicos de evaluación del TE docente, incorporando estrategias de regulación emocional más diversas y pertinentes, así como explorar enfoques alternativos como el análisis de perfiles latentes (Grandey y Gabriel, 2015), que podrían capturar de manera más precisa las dinámicas emocionales complejas que enfrenta el magisterio.

Desde una perspectiva crítica, este estudio también confirma vacíos detectados en trabajos previos, como la investigación de Ma et al. (2023), en la que se señala la escasa representación de la voz docente en el diseño de instrumentos. A este respecto, resulta preocupante que pocos estudios consideren la aceptabilidad y aplicabilidad de los instrumentos desde la experiencia de los propios usuarios, lo que limita su utilidad práctica en contextos escolares reales.

En síntesis, los resultados de esta revisión no solo permiten sistematizar los principales instrumentos existentes, sino también identificar una agenda pendiente de desarrollo metodológico, contextualización cultural y validación práctica. Para los tomadores de decisiones, esto implica la necesidad de seleccionar herramientas válidas y confiables, pero también culturalmente pertinentes, éticamente justificadas y sensibles a las condiciones reales del trabajo docente. Asimismo, se recomienda impulsar investigaciones futuras que articulen la evaluación del TE con intervenciones de formación socioemocional y estrategias de cuidado institucional, reconociendo el papel central de las emociones en los procesos de enseñanza-aprendizaje.

REFERENCIAS BIBLIOGRÁFICAS

- Andela, M., Truchot, D., y Borteyrou, X. (2015). Emotional labour and burnout: Some methodological considerations and refinements. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement*, 47(4), 321–332. <https://doi.org/10.1037/cbs0000024>

- Ashforth, B., y Humphrey, R. (1993). Emotional Labor in Service Roles: The Influence of Identity. *The Academy of Management Review*, 18(1), 88–115. <https://doi.org/10.2307/258824>
- Ashkanasy, N., Härtel, C., y Daus, C. (2002). Diversity and Emotion: The New Frontiers in Organizational Behavior Research. *Journal of Management*, 28(3), 307–338. [https://doi.org/10.1016/S0149-2063\(02\)00130-7](https://doi.org/10.1016/S0149-2063(02)00130-7)
- Blau, G., Fertig, J., Surges Tatum, D., Connaughton, S., Soo Park, D., y Marshall, C. (2010). Further scale refinement for emotional labor. *Career Development International*, 15(2), 188–216. <https://doi.org/10.1108/13620431011040969>
- Brotheridge, C., y Lee, R. (2003). Development and validation of the Emotional Labour Scale. *Journal of Occupational and Organizational Psychology*, 76(3), 365–379. <https://doi.org/10.1348/096317903769647229>
- Carroll, A., Flynn, L., O'Connor, E., Forrest, K., Bower, J., Fynes-Clinton, S., York, A., y Ziaei, M. (2020). In their words: listening to teachers' perceptions about stress in the workplace and how to address it. *Asia-Pacific Journal of Teacher Education*, 49(4), 420–434. <https://doi.org/10.1080/1359866X.2020.1789914>
- Chen, Y., Sun, S., y Liu, X. (2025). The relationship between teacher emotional labor and work/family conflict: The mediating role of teacher-colleague relations. *Acta Psychologica*, 254, 1–10. <https://doi.org/10.1016/j.actpsy.2025.104832>
- Çukur, C. (2009). The development of the Teacher Emotional Labor Scale (TELS): Validity and reliability. *Educational Sciences: Theory & Practice*, 9(2), 559–574. <https://files.eric.ed.gov/fulltext/EJ847767.pdf>
- De-Ruiter, J., Poorthuis, A., y Koomen, H. (2021). Teachers' emotional labor in response to daily events with individual students: The role of teacher–student relationship quality. *Teaching and Teacher Education*, 107, 1–12. <https://doi.org/10.1016/j.tate.2021.103467>
- Dewaele, J.-M., y Wu, A. (2021). Predicting the emotional labor strategies of Chinese English Foreign Language teachers. *System*, 103, 1–14. <https://doi.org/10.1016/j.system.2021.102660>
- Dickerson, M. K., Fenech, M., y Stratigos, T. (2024). Working with families: an investigation of early childhood educators' emotional labour and wellbeing. *Early Years An International Research Journal*, 1–17. <https://doi.org/10.1080/09575146.2024.2393143>

- Diefendorff, J. M., Croyle, M. H., y Gosserand, R. H. (2005). The dimensionality and antecedents of emotional labor strategies. *Journal of Vocational Behavior*, 66(2), 339–357. <https://doi.org/10.1016/j.jvb.2004.02.001>
- Donker, M., Erisman, M., van-Gog, T., y Mainhard, T. (2020). Teachers' Emotional Exhaustion: Associations With Their Typical Use of and Implicit Attitudes Toward Emotion Regulation Strategies. *Frontiers in Psychology*, 11, 1–15. <https://doi.org/10.3389/fpsyg.2020.00867>
- El-Den, S., Schneider, C., Mirzaei, A., y Carter, S. (2020). How to measure a latent construct: Psychometric principles for the development and validation of measurement instruments. *International Journal of Pharmacy Practice*, 28(4), 326–336. <https://doi.org/10.1111/ijpp.12600>
- Glomb, T. M., y Tews, M. J. (2004). Emotional labor: A conceptualization and scale development. *Journal of Vocational Behavior*, 64(1), 1–23. [https://doi.org/10.1016/S0001-8791\(03\)00038-1](https://doi.org/10.1016/S0001-8791(03)00038-1)
- Grandey, A., y Gabriel, A. (2015). Emotional Labor at a Crossroads: Where Do We Go from Here? *Annual Review of Organizational Psychology and Organizational Behavior*, 2(1), 323–349. <https://doi.org/10.1146/annurev-orgpsych-032414-111400>
- Grandey, A., y Melloy, R. (2017). The state of the heart: Emotional labor as emotion regulation reviewed and revised. *Journal of Occupational Health Psychology*, 22(3), 407–422. <https://doi.org/10.1037/ocp0000067>
- Grimm, P. (2010). Social Desirability Bias. En *Wiley International Encyclopedia of Marketing*. Wiley. <https://doi.org/10.1002/9781444316568.wiem02057>
- Gross, J. (1998). The Emerging Field of Emotion Regulation: An Integrative Review. *Review of General Psychology*, 2(3), 271–299. <https://doi.org/10.1037/1089-2680.2.3.271>
- Heffernan, T. A., y Bosetti, L. (2020). The emotional labour and toll of managerial academia on higher education leaders. *Journal of Educational Administration and History*, 52(4), 357–372. <https://doi.org/10.1080/00220620.2020.1725741>
- Hochschild, A. (1983). *The Managed Heart: Commercialization of human feeling*. University of California Press.
- Hogan, T. (2015). *Pruebas psicológicas: Una introducción práctica* (J. Cisneros, Trad.; 2a ed.). El Manual Moderno. <http://repositorio.uasb.edu.bo/handle/20.500.14624/1201>

- Howard, J. L. (2023). Psychometric Approaches in Self-Determination Theory. En *The Oxford Handbook of Self-Determination Theory* (pp. 438–454). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780197600047.013.15>
- Kariou, A., Koutsimani, P., Montgomery, A., y Lainidi, O. (2021). Emotional Labor and Burnout among Teachers: A Systematic Review. *International Journal of Environmental Research and Public Health*, 18(23), 1–15. <https://doi.org/10.3390/ijerph182312760>
- Kıral, E. (2016). Psychometric Properties of the Emotional Labor Scale in a Turkish Sample of School Administrators. *Eurasian Journal of Educational Research*, 16(63), 71–88. <https://files.eric.ed.gov/fulltext/EJ1112448.pdf>
- Koo, M., y Yang, S.-W. (2025). Likert-Type Scale. *Encyclopedia*, 5(18), 1–11. <https://doi.org/10.3390/encyclopedia5010018>
- Kruml, S. M., y Geddes, D. (2000). Exploring the Dimensions of Emotional Labor: The Heart of Hochschild's Work. *Management Communication Quarterly*, 14(1), 8–49. <https://doi.org/10.1177/0893318900141002>
- Li, H., y Liu, H. (2021). Beginning EFL Teachers' Emotional Labor Strategies in the Chinese Context. *Frontiers in Psychology*, 12, 1–12. <https://doi.org/10.3389/fpsyg.2021.737746>
- Liu, Y., y Zhang, D. (2014). Development of Questionnaire on Emotional Labor among Primary and Secondary School Teachers. *Journal of Education and Training Studies*, 3(1), 46–55. <https://doi.org/10.11114/jets.v3i1.551>
- Ma, P., Zhang, L., Dong, H., y Yu, J. (2023). The relationships between teachers' emotional labor and display rules, trait emotions, exhaustion, and classroom emotional climate. *Frontiers in Psychology*, 14, 1–14. <https://doi.org/10.3389/fpsyg.2023.957856>
- Moreno-Jiménez, B., Gálvez, M., Rodríguez-Carvajal, R., y Garrosa, E. (2010). Emociones y salud en el trabajo: análisis del constructo “trabajo emocional” y propuesta de evaluación. *Revista Latinoamericana de Psicología*, 42(1), 63–73. <https://www.redalyc.org/pdf/805/80515880006.pdf>
- Näring, G., Briët, M., y Brouwers, A. (2007). Validation of the Dutch Questionnaire on Emotional Labor (D-QEL) in Nurses and Teachers. En P. Richter, J. M. Peiro, y W. B. Schaufeli (Eds.), *Psychosocial Resources in Health Care Systems* (pp. 135–145). Rainer Hampp Verlag. <https://research.ou.nl/en/publications/validation-of-the-dutch-questionnaire-on-emotional-labor-d-gel-in>

- Olmsted, J. (2024). Short Report Research Reliability and Validity: Why do they matter? *The Journal of Dental Hygiene*, 98(6), 53–57. <https://jdh.adha.org/content/98/6/53>
- Ortiz, S., Navarro, C., García, E., Ramis, C., y Manassero, M. A. (2012). Validación de la versión española de la Escala de Trabajo Emocional de Frankfurt. *Psicothema*, 24(2), 337–342. <https://www.psicothema.com/pdf/4020.pdf>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372(n71). <https://doi.org/10.1136/bmj.n71>
- Peng, R., Hu, Q., y Kouider, B. (2023). Teachers' Acceptance of Online Teaching and Emotional Labor in the EFL Context. *Sustainability*, 15, 1–19. <https://doi.org/10.3390/su151813893>
- Peng, W., Liu, Y., y Peng, J.-E. (2023). Feeling and acting in classroom teaching: The relationships between teachers' emotional labor, commitment, and well-being. *System*, 116, 1–14. <https://doi.org/10.1016/j.system.2023.103093>
- Pimentel, J. (2019). Some Biases in Likert Scaling Usage and its Correction. *International Journal of Sciences: Basic and Applied Research*, 45(1), 183–191. <https://www.gssrr.org/index.php/JournalOfBasicAndApplied/article/view/9874>
- Quiñones-García, C., Rodríguez-Carvajal, R., Clarke, N., y Moreno-Jiménez, B. (2013). Desarrollo y validación trasnacional de la escala de esfuerzo emocional (EEF). *Psicothema*, 25(3), 363–369. <https://doi.org/10.7334/psicothema2012.289>
- Salessi, S., y Omar, A. (2016). Desarrollo y Validación de una Escala para Medir Actuación Emocional en el Trabajo. *Revista Iberoamericana de Diagnóstico y Evaluación*, 1(41), 66–79. <https://www.redalyc.org/journal/4596/459646901007/html/>
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, 8(4), 350–353. <https://doi.org/10.1037/1040-3590.8.4.350>
- Sierra, B. (2001). *Técnicas de Investigación Social. Teoría y Ejercicios* (Paraninfo, Ed.; 14a ed.). Thomson Learning. <https://abcproyecto.wordpress.com/wp-content/uploads/2018/11/sierra-bravo-tecnicas-de-investigacion-social.pdf>
- Sosa-Sánchez, K. P. G., y De los Santos Colorado, L. E. (2025). Cadenas de búsqueda empleadas en el artículo: "Una Revisión Sistemática de Instrumentos Psicométricos para

- Evaluar el Trabajo Emocional en Docentes" / "A Systematic Review of Psychometric Instruments for Assessing Emotional Labor in Teachers". Zenodo. <https://doi.org/10.5281/zenodo.15713504>
- Taheri, E., Shirzad, K., Afsahi, F., y Kouhsari, M. (2024). The Effect of Teacher Emotional Labor on Commitment: The Mediating Role of Teacher Psychological Capital in Iranian Primary Schools. *Educational Studies*, 60(2), 199–215. <https://doi.org/10.1080/00131946.2023.2280652>
- Terwee, C., Jansma, E., Riphagen, I., y De Vet, H. (2009). Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of Life Research*, 18(8), 1115–1123. <https://doi.org/10.1007/s11136-009-9528-5>
- Wang, H., Hall, N. C., y King, R. B. (2021). A longitudinal investigation of teachers' emotional labor, well-being, and perceived student engagement. *Educational Psychology*, 41(10), 1319 – 1336. <https://doi.org/10.1080/01443410.2021.1988060>
- Wang, H., Hall, N., y Taxer, J. (2019). Antecedents and Consequences of Teachers' Emotional Labor: a Systematic Review and Meta-analytic Investigation. *Educational Psychology Review*, 31(3), 663–698. <https://doi.org/10.1007/s10648-019-09475-3>
- Weiss, H., y Cropanzano, R. (1996). Affective Events Theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. *Research in organizational behavior: An annual series of analytical essays and critical reviews*, 18, 1–74. https://web.mit.edu/curhan/www/docs/Articles/15341_Readings/Affect/AffectiveEventsTheory_WeissCropanzano.pdf
- Yang, C., y Chen, A. (2021). Emotional labor: A comprehensive literature review. *Human Systems Management*, 40(4), 479–501. <https://doi.org/10.3233/HSM-200937>
- Yang, C., Chen, Y., y Zhao, X. (2019). Emotional Labor: Scale Development and Validation in the Chinese Context. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.02095>
- Yin, H. (2012). Adaptation and validation of the teacher emotional labour strategy scale in China. *Educational Psychology*, 32(4), 451–465. <https://doi.org/10.1080/01443410.2012.674488>

- Yin, H., Huang, S., y Chen, G. (2019). The relationships between teachers' emotional labor and their burnout and satisfaction: A meta-analytic review. *Educational Research Review*, 28, 1–18. <https://doi.org/10.1016/j.edurev.2019.100283>
- Yoshida, Y., Sitas, N., Mannetti, L., O'Farrell, P., Arroyo-Robles, G., Berbés-Blázquez, M., González-Jiménez, D., Nelson, V., Niamir, A., y Harmáčková, Z. (2024). Beyond Academia: A case for reviews of gray literature for science-policy processes and applied research. *Environmental Science & Policy*, 162, 1–11. <https://doi.org/10.1016/j.envsci.2024.103882>
- Zhang, Q., Li, W., Gao, J., Sun, B., y Lin, S. (2024). Teachers' professional identity and job burnout: The mediating roles of work engagement and psychological capital. *Psychology in the Schools*, 61(1), 123–136. <https://doi.org/10.1002/pits.23039>
- Zheng, X., Shi, X., y Liu, Y. (2020). Leading Teachers' Emotions Like Parents: Relationships Between Paternalistic Leadership, Emotional Labor and Teacher Commitment in China. *Frontiers in Psychology*, 11, 1–9. <https://doi.org/10.3389/fpsyg.2020.00519>

Información Suplementaria

[Apéndice](#)