

<https://doi.org/10.69639/arandu.v13i3.2482>

Desarrollo y validación psicométrica de la Escala de Falso Liderazgo (FLS-13): la discrepancia entre la imagen directiva proyectada y las conductas percibidas

*Development and Psychometric Validation of the False Leadership Scale (FLS-13):
Discrepancy Between Projected Managerial Image and Perceived Behaviors*

Alexander Rudy Ayala Huamán

alexander.ayala@unmsm.edu.pe

<https://orcid.org/0000-0001-7655-0154>

Universidad Nacional Mayor de San Marcos

Perú – Lima

Mervin Ediana De Souza Cañari

mervin.desouza@unmsm.edu.pe

<https://orcid.org/0009-0006-4728-5252>

Universidad Nacional Mayor de San Marcos

Perú – Lima

*Artículo recibido: 10 julio 2026- Aceptado para publicación: 16 agosto 2026
Conflictos de intereses: Ninguno que declarar.*

RESUMEN

El liderazgo destructivo constituye un riesgo psicosocial relevante para las organizaciones; sin embargo, existen limitaciones para evaluar la discrepancia entre la imagen directiva proyectada y las conductas percibidas por los colaboradores. El objetivo fue desarrollar la False Leadership Scale (FLS-13) y obtener evidencias de su validez y confiabilidad. Se realizó un estudio instrumental en dos fases independientes. En la fase exploratoria participaron 56 profesionales, quienes evaluaron la validez de contenido, la estructura factorial y la consistencia interna del instrumento. Posteriormente, se efectuó un análisis factorial confirmatorio (AFC) con una muestra independiente de 158 participantes. La validez de contenido alcanzó un V de Aiken global de .84. La matriz exploratoria mostró adecuada factorizabilidad ($KMO = .930$; χ^2 de Bartlett(78) = 837, $p < .001$), mientras que el análisis paralelo respaldó una estructura unifactorial, con cargas entre .724 y .917. El AFC, mediante DWLS, mostró un ajuste adecuado del modelo unifactorial reespecificado con dos covarianzas residuales conceptualmente justificadas: $\chi^2(63) = 137$, $p < .001$, CFI = .962, TLI = .953, SRMR = .032 y RMSEA = .086, IC 90% [.066, .106]. Las cargas factoriales oscilaron entre .766 y .880. Asimismo, se obtuvo evidencia favorable de validez convergente (AVE = .677), confiabilidad compuesta (CR = .965) y consistencia interna ($\alpha = .962$; $\omega = .962$). Los resultados respaldan inicialmente la estructura unidimensional de la FLS-13, aunque se requieren nuevas investigaciones para examinar su estabilidad, invariancia y diferenciación respecto de constructos próximos.

Palabras clave: liderazgo destructivo, liderazgo tóxico, liderazgo oscuro, falso liderazgo, organizaciones

ABSTRACT

Destructive leadership constitutes a relevant psychosocial risk for organizations; however, limitations remain in assessing the discrepancy between the managerial image projected by those in leadership positions and the behaviors perceived by their collaborators. The objective was to develop the False Leadership Scale (FLS-13) and obtain evidence of its validity and reliability. An instrumental study was conducted in two independent phases. In the exploratory phase, 56 professionals participated and evaluated the instrument's content validity, factorial structure, and internal consistency. Subsequently, confirmatory factor analysis (CFA) was conducted with an independent sample of 158 participants. Content validity reached an overall Aiken's V of .84. The exploratory matrix showed adequate factorability ($KMO = .930$; Bartlett's $\chi^2(78) = 837$, $p < .001$), while parallel analysis supported a unidimensional structure, with factor loadings ranging from .724 to .917. The CFA, using diagonally weighted least squares (DWLS), showed adequate fit for the re-specified unidimensional model with two conceptually justified residual covariances: $\chi^2(63) = 137$, $p < .001$, CFI = .962, TLI = .953, SRMR = .032, and RMSEA = .086, 90% CI [.066, .106]. Factor loadings ranged from .766 to .880. Favorable evidence was also obtained for convergent validity (AVE = .677), composite reliability (CR = .965), and internal consistency ($\alpha = .962$; $\omega = .962$). The results provide initial support for the unidimensional structure of the FLS-13, although further research is needed to examine its stability, measurement invariance, and differentiation from conceptually related constructs.

Keywords: destructive leadership, toxic leadership, dark leadership, false leadership, organizations

Todo el contenido de la Revista Científica Internacional Arandu UTIC publicado en este sitio está disponible bajo licencia Creative Commons Attribution 4.0 International. 

INTRODUCCIÓN

El liderazgo constituye un fenómeno central en el estudio de las organizaciones por su función en los procesos de influencia y en la consecución de objetivos colectivos. Su comprensión evolucionó desde perspectivas centradas en las características del líder hacia enfoques que incorporan la interacción entre líderes, seguidores y contexto. Northouse (2022) concibe el liderazgo como un proceso mediante el cual un individuo influye en un grupo para alcanzar un objetivo común, perspectiva que enfatiza su naturaleza relacional. Burns (1978) distinguió formas de liderazgo sustentadas en diferentes relaciones de influencia, mientras que Bass (1985) amplió esta perspectiva mediante el desarrollo del liderazgo transformacional. Posteriormente, Avolio et al. (2009) revisaron los avances y desafíos acumulados en la investigación del liderazgo, Lord et al. (2017) enfatizaron su carácter dinámico y Carton (2022) planteó la necesidad de avanzar hacia una mayor precisión en la explicación de los comportamientos de liderazgo.

El predominio de perspectivas orientadas al liderazgo favorable impulsó modelos como el transformacional (Bass & Avolio, 1994), el auténtico, sustentado en autoconocimiento y transparencia relacional (Avolio & Gardner, 2005), y el ético, vinculado con la conducta normativamente apropiada y su promoción entre los seguidores (Brown & Treviño, 2006). Desde el liderazgo de servicio, Gandolfi y Stone (2018) destacaron la orientación hacia las personas y los objetivos compartidos. En conjunto, estos enfoques ampliaron el análisis del liderazgo más allá de los resultados institucionales, incorporando las relaciones y experiencias de quienes reciben su influencia.

Esta ampliación resulta relevante porque los comportamientos de liderazgo se relacionan no solo con el desempeño, sino también con el bienestar de los trabajadores. Inceoglu et al. (2018) identificaron mecanismos sociales-cognitivos, motivacionales, afectivos, relacionales y de identificación mediante los cuales el comportamiento del líder puede incidir en el bienestar. Asimismo, Schneider et al. (2013) señalaron el papel del liderazgo en la comprensión de la cultura y el funcionamiento organizacional. Por ello, el análisis contemporáneo del liderazgo requiere considerar tanto la conducta directiva como la forma en que esta es experimentada por los colaboradores.

Sin embargo, el desarrollo de la evidencia empírica reveló que los problemas organizacionales no podían explicarse únicamente por la ausencia de liderazgo positivo. Determinadas prácticas directivas pueden deteriorar tanto el funcionamiento institucional como el bienestar de los trabajadores, impulsando el estudio del denominado lado oscuro del liderazgo (Hogan & Kaiser, 2005; Lord et al., 2017; Carton, 2022). En este marco surgieron constructos como la *petty tyranny* (Ashforth, 1994), la supervisión abusiva (Tepper, 2000) y el liderazgo destructivo, entendido como comportamientos sistemáticos que perjudican a la organización y/o a sus miembros (Einarsen et al., 2007). Asimismo, el modelo del Triángulo Tóxico amplió esta

perspectiva al explicar los resultados destructivos mediante la interacción entre líderes, seguidores susceptibles y contextos organizacionales propicios (Padilla et al., 2007).

Dentro de esta corriente, Schmidt (2008) desarrolló la Toxic Leadership Scale (TLS), identificando cinco dimensiones del liderazgo tóxico: supervisión abusiva, liderazgo autoritario, narcisismo, autopromoción e imprevisibilidad. De manera complementaria, Paulhus y Williams (2002) conceptualizaron la Dark Triad de la personalidad a partir del narcisismo, el maquiavelismo y la psicopatía, ampliando el estudio de rasgos socialmente aversivos vinculados al comportamiento humano. Estos desarrollos evidencian que el lado oscuro del liderazgo comprende manifestaciones diferenciadas pero relacionadas y abren la necesidad de examinar formas disfuncionales que podrían no estar plenamente representadas por los constructos existentes. En efecto, el liderazgo destructivo se ha relacionado con actitudes laborales negativas y afectaciones del bienestar (Schyns & Schilling, 2013), mientras que los comportamientos de liderazgo pueden incidir en la salud ocupacional mediante mecanismos motivacionales, afectivos, cognitivos y relacionales (Inceoglu et al., 2018).

Esta problemática adquiere especial relevancia en las organizaciones de salud, caracterizadas por elevada complejidad, presión emocional, limitaciones de recursos y permanente coordinación interdisciplinaria. En estos escenarios, las decisiones y conductas directivas pueden repercutir sobre el bienestar de los trabajadores, el funcionamiento de los equipos y los procesos asistenciales, mientras que los líderes contribuyen a configurar el clima y la cultura organizacional (Schneider et al., 2017). La evidencia internacional ha documentado, además, experiencias de intimidación, menosprecio y acoso entre profesionales sanitarios (Camm et al., 2023), así como condiciones organizacionales capaces de favorecer el silencio de los trabajadores y limitar la expresión de problemas y preocupaciones relevantes (Morrison, 2014).

Persiste, sin embargo, un vacío en la evaluación de formas de conducción, que no necesariamente se manifiestan mediante comportamientos abiertamente hostiles o abusivos. Mientras constructos como el liderazgo destructivo, la supervisión abusiva y el liderazgo tóxico cuentan con una trayectoria conceptual consolidada, permanece menos explorada *la discrepancia* entre la imagen positiva, ética o comprometida que proyecta quien ejerce funciones directivas y las conductas que sus colaboradores perciben en la interacción cotidiana. Esta brecha adquiere relevancia ante la necesidad de avanzar hacia conceptualizaciones más específicas del liderazgo (Avolio et al., 2009; Lord et al., 2017; Carton, 2022).

Construcción teórica y delimitación conceptual del Falso Liderazgo frente a constructos

La expansión de constructos asociados al lado oscuro del liderazgo ha contribuido a comprender sus diversas manifestaciones, aunque también ha puesto de relieve dificultades para su delimitación conceptual y medición. Akinyele y Chen (2025) advirtieron falta de uniformidad en la conceptualización del liderazgo tóxico y en las dimensiones consideradas por sus instrumentos. Schmidt (2008), mediante la *Toxic Leadership Scale*, lo operacionalizó en cinco

dimensiones: supervisión abusiva, liderazgo autoritario, narcisismo, autpromoción e imprevisibilidad. Desde una perspectiva más amplia, Carton (2022) cuestionó las conceptualizaciones que agrupan múltiples comportamientos en constructos agregados y planteó la necesidad de avanzar hacia una mayor especificidad conductual.

En este marco, la consolidación de constructos previos no descarta la existencia de fenómenos aún insuficientemente delimitados por las categorías disponibles. Particular interés adquiere la discrepancia percibida entre la imagen proyectada por quien ejerce funciones directivas y las conductas que sus colaboradores le atribuyen en la interacción cotidiana. Esta incongruencia imagen–conducta configura el núcleo conceptual propuesto del Falso Liderazgo y sustenta la necesidad de examinar empíricamente su estructura y diferenciación frente a constructos relacionados.

El Falso Liderazgo se inscribe en las manifestaciones disfuncionales del liderazgo, pero se distingue por situar dicha discrepancia percibida como su elemento central. La supervisión abusiva se concentra en conductas hostiles verbales y no verbales del supervisor (Tepper, 2000); el liderazgo tóxico comprende supervisión abusiva, autoritarismo, narcisismo, autpromoción e imprevisibilidad (Schmidt, 2008); y el liderazgo destructivo abarca comportamientos sistemáticos capaces de perjudicar a la organización y/o a sus miembros (Einarsen et al., 2007). En consecuencia, la incongruencia imagen–conducta *fundamenta* conceptualmente el desarrollo de una medida específica para la evaluación empírica del Falso Liderazgo.

Propuesta conceptual del Falso Liderazgo

En respuesta al vacío identificado, la presente investigación propone el Falso Liderazgo como un constructo centrado en la discrepancia percibida entre la imagen que proyecta quien ejerce funciones directivas y las conductas que sus colaboradores le atribuyen. Se define como un patrón percibido de incongruencia entre una imagen de legitimidad, ética, transparencia, cercanía, responsabilidad o compromiso y comportamientos interpretados como inconsistentes con ella, expresados mediante manipulación, favoritismo, evasión de responsabilidades, apropiación de resultados, desatención de las necesidades del personal o utilización estratégica de las relaciones de poder. Esta discrepancia imagen–conducta constituye el núcleo conceptual propuesto para su evaluación empírica, diferenciándolo conceptualmente de manifestaciones centradas en conductas perjudiciales, abusivas o disfuncionales descritas en el liderazgo destructivo (Einarsen et al., 2007), el liderazgo tóxico (Schmidt, 2008) y la supervisión abusiva (Tepper, 2000).

Tabla 1
Delimitación Conceptual

Constructo	Referente	Núcleo conceptual	Diferencia respecto del Falso Liderazgo
Supervisión abusiva	Tepper (2000)	Percepción sostenida de conductas hostiles verbales y no verbales del supervisor	La hostilidad constituye el núcleo; no requiere discrepancia entre imagen proyectada y conducta percibida
Liderazgo destructivo	Einarsen et al. (2007)	Conductas sistemáticas y repetidas que perjudican a la organización y/o a sus miembros	Constituye un dominio amplio; la incongruencia imagen–conducta no constituye necesariamente su elemento definitorio
Liderazgo tóxico	Schmidt (2008)	Supervisión abusiva, autoritarismo, narcisismo, autopromoción e imprevisibilidad	Es multidimensional y se centra en patrones disfuncionales; el Falso Liderazgo propone como núcleo la discrepancia percibida imagen–conducta
Liderazgo narcisista	Rosenthal & Pittinsky (2006)	Influencia de necesidades y rasgos narcisistas sobre el ejercicio del liderazgo	Se centra principalmente en características narcisistas del líder, no en la discrepancia percibida entre representación y comportamiento
Liderazgo despótico	De Hoogh & Den Hartog (2008)	Autoritarismo, dominación y orientación hacia intereses personales	Su núcleo se encuentra en el ejercicio autoritario y autointeresado del poder
Falso Liderazgo	Propuesta del presente estudio	Discrepancia percibida entre la imagen directiva proyectada y las conductas atribuidas al líder	La incongruencia imagen–conducta constituye la unidad conceptual central propuesta

Nota. La diferenciación presentada es de naturaleza conceptual. El presente estudio aporta evidencia inicial sobre la estructura interna de la FLS-13, pero no establece todavía su validez discriminante frente a todos los constructos relacionados. Esta deberá examinarse en investigaciones posteriores mediante la aplicación simultánea de instrumentos consolidados de liderazgo destructivo, tóxico y supervisión abusiva.

MATERIALES Y MÉTODOS

Se realizó un estudio instrumental orientado al desarrollo y evaluación de las propiedades psicométricas de la False Leadership Scale (FLS-13), diseñada para evaluar el Falso Liderazgo percibido en contextos organizacionales. La evaluación se desarrolló en dos fases independientes: una exploratoria con 56 participantes, destinada a examinar la factorizabilidad y estructura inicial mediante AFE, y una confirmatoria con 158 participantes, orientada a contrastar la estructura unidimensional mediante AFC. Este diseño responde a las recomendaciones para el desarrollo y evaluación de instrumentos de medición y obtención de evidencias de validez y confiabilidad (AERA et al., 2014; Muñiz & Fonseca-Pedrero, 2019).

Considerando la naturaleza ordinal de los 13 reactivos de la FLS-13, evaluados mediante una escala Likert de cinco categorías, el AFC se estimó mediante mínimos cuadrados ponderados diagonalmente (DWLS). El ajuste se evaluó conjuntamente mediante χ^2 , CFI, TLI, SRMR y RMSEA con su respectivo intervalo de confianza, además de examinar las cargas factoriales

estandarizadas. El uso de DWLS resulta apropiado para el análisis factorial de indicadores ordinales (Li, 2016; Kline, 2023).

Participantes

Participaron trabajadores de instituciones del sistema público de salud peruano, distribuidos en dos muestras independientes: 56 participantes en la fase exploratoria y 158 en la fase confirmatoria. Se incluyeron profesionales asistenciales con vínculo laboral vigente, permanencia mínima de seis meses y consentimiento informado. Se excluyeron cuestionarios incompletos o con patrones de respuesta inconsistentes, así como participantes en situación de suspensión perfecta o imperfecta de labores durante la recolección de datos.

Desarrollo del instrumento

La construcción de la False Leadership Scale (FLS-13) se sustentó en una revisión crítica de la literatura sobre las principales manifestaciones disfuncionales del liderazgo, incluyendo el liderazgo destructivo, el liderazgo tóxico, la supervisión abusiva, el liderazgo despótico y el liderazgo narcisista (Einarsen et al., 2007; Schmidt, 2008; Tepper, 2000; De Hoogh & Den Hartog, 2008; Rosenthal & Pittinsky, 2006). A partir del análisis y delimitación de estos constructos se identificó un espacio conceptual insuficientemente representado: la discrepancia percibida entre la imagen que proyecta quien ejerce funciones directivas y las conductas que sus colaboradores le atribuyen en la interacción cotidiana. Esta discrepancia constituyó el núcleo conceptual para la formulación del Falso Liderazgo y orientó la construcción de los reactivos de la FLS-13.

La elaboración del banco inicial de reactivos siguió recomendaciones metodológicas para el desarrollo de instrumentos psicométricos (Boateng et al., 2018; Muñiz & Fonseca-Pedrero, 2019). Los reactivos fueron formulados procurando claridad, pertinencia y correspondencia con la definición conceptual propuesta, mediante expresiones comprensibles y evitando ambigüedades, dobles negaciones y formulaciones innecesariamente complejas. La construcción de los ítems se orientó a representar manifestaciones de la discrepancia imagen-conducta percibida en el ejercicio del liderazgo, incluyendo incongruencias relacionadas con el bienestar del personal, la participación en las decisiones, la imparcialidad, la responsabilidad, el reconocimiento, los intereses organizacionales y las relaciones de poder. La versión sometida a evaluación psicométrica quedó conformada por 13 reactivos, correspondientes a la FLS-13.

Los 13 reactivos fueron evaluados mediante una escala tipo Likert de cinco categorías de respuesta, desde 1 (Totalmente en desacuerdo) hasta 5 (Totalmente de acuerdo). La puntuación se estableció en sentido directo, de modo que puntuaciones más elevadas representan una mayor percepción de Falso Liderazgo en quien ejerce funciones directivas.

Evidencias de validez de contenido

Las evidencias de validez de contenido de la FLS-13 se obtuvieron mediante el juicio de tres expertos en psicometría, metodología de la investigación y liderazgo organizacional, quienes evaluaron los reactivos según los criterios de claridad, coherencia, pertinencia y relevancia. Este procedimiento permitió examinar la correspondencia entre el contenido de los reactivos y el dominio conceptual que se pretendía representar. El grado de acuerdo se estimó mediante la V de Aiken, empleada para cuantificar la valoración de los jueces respecto de la adecuación y representatividad del contenido (Escobar-Pérez & Cuervo-Martínez, 2008; Merino-Soto & Livia-Segovia, 2009). Las observaciones formuladas por los expertos permitieron realizar ajustes de redacción antes del estudio piloto, procurando mejorar la precisión y comprensión de los reactivos sin alterar su correspondencia con la definición conceptual del Falso Liderazgo.

Procedimiento

La recolección de datos se realizó con profesionales asistenciales de instituciones públicas de salud del Perú, mediante participación voluntaria, anónima y confidencial. Previamente, los participantes fueron informados sobre los objetivos y la naturaleza académica del estudio y otorgaron su consentimiento informado. La investigación se desarrolló conforme a los principios éticos de la Declaración de Helsinki (WMA, 2013) y a las normas nacionales aplicables a la investigación científica.

Análisis estadístico

Los análisis psicométricos se realizaron en jamovi. Se calcularon estadísticos descriptivos y se examinó la factorizabilidad mediante el índice KMO y la prueba de esfericidad de Bartlett. La dimensionalidad de la FLS-13 se evaluó mediante AFE con extracción por mínimos cuadrados, rotación Oblimin y análisis paralelo para determinar el número de factores (Horn, 1965; Hayton et al., 2004). La estructura obtenida se contrastó en una muestra independiente mediante AFC con estimación DWLS, dada la naturaleza ordinal de los reactivos, evaluándose el ajuste mediante χ^2 , CFI, TLI, RMSEA y SRMR (Brown, 2015; Kline, 2023). La consistencia interna se estimó mediante α de Cronbach y ω de McDonald como indicadores complementarios de confiabilidad (McDonald, 1999; Hayes & Coutts, 2020).

Consideraciones éticas

La investigación respetó los principios de autonomía, beneficencia, no maleficencia y justicia. La participación fue completamente voluntaria y todos los participantes otorgaron su consentimiento informado antes de responder el cuestionario. Se garantizó el anonimato, la confidencialidad de la información y el uso exclusivo de los datos con fines científicos. Los resultados fueron analizados de manera agregada, evitando cualquier posibilidad de identificación individual de los participantes.

RESULTADOS Y DISCUSIÓN

Desarrollo y refinamiento inicial de la False Leadership Scale (FLS-13)

La False Leadership Scale (FLS-13) fue desarrollada mediante un proceso sistemático de construcción psicométrica, siguiendo recomendaciones para el diseño de instrumentos de medición (Boateng et al., 2018; Muñiz & Fonseca-Pedrero, 2019). La formulación de los reactivos partió de la delimitación conceptual del Falso Liderazgo frente a constructos relacionados, definiendo como núcleo la discrepancia percibida entre la imagen proyectada por quien ejerce funciones directivas y las conductas que sus colaboradores le atribuyen. Los indicadores se elaboraron procurando claridad semántica, precisión conceptual y correspondencia con el constructo.

La versión preliminar fue sometida a juicio de expertos, quienes evaluaron la claridad, coherencia, pertinencia y relevancia de los reactivos. Sus observaciones permitieron realizar ajustes antes del estudio piloto, obteniéndose una versión de 13 ítems con escala Likert de cinco categorías, desde 1 (Totalmente en desacuerdo) hasta 5 (Totalmente de acuerdo).

Evidencias de validez de contenido

La validez de contenido de la False Leadership Scale (FLS-13) fue examinada mediante juicio de expertos, procedimiento que permite valorar la correspondencia y representatividad de los reactivos respecto del constructo que se pretende medir (Escobar-Pérez & Cuervo-Martínez, 2008). Tres expertos evaluaron los 13 reactivos considerando los criterios de claridad, coherencia, pertinencia y relevancia. El grado de acuerdo entre sus valoraciones se cuantificó mediante el coeficiente V de Aiken, índice empleado para estimar la validez de contenido a partir de las puntuaciones otorgadas por los jueces a los reactivos (Aiken, 1980; Merino-Soto & Livia-Segovia, 2009).

Tabla 1

V de Aiken global según criterio de evaluación de la FLS-13

Criterio evaluado	V de Aiken
Claridad	0.87
Coherencia	0.85
Pertinencia	0.84
Relevancia	0.82
V de Aiken global	0.84

Los coeficientes fueron $V = .87$ para claridad, $V = .85$ para coherencia, $V = .84$ para pertinencia y $V = .82$ para relevancia, con una V de Aiken global = .84. En conjunto, los valores superiores a .80 evidenciaron una valoración favorable del contenido de los reactivos por parte de los jueces (Merino-Soto & Livia-Segovia, 2009).

Tabla 2*V de Aiken por reactivo y criterio de evaluación de la FLS-13*

Ítem	Claridad	Coherencia	Pertinencia	Relevancia
FL1	0.89	0.89	1.00	0.89
FL2	0.89	0.89	0.78	0.78
FL3	0.89	0.78	0.78	0.89
FL4	0.78	0.78	0.78	0.78
FL5	1.00	1.00	0.89	0.89
FL6	0.89	0.89	0.78	0.78
FL7	0.89	0.89	0.89	0.89
FL8	0.89	0.89	0.89	0.89
FL9	0.78	0.78	0.78	0.78
FL10	1.00	0.78	0.78	0.78
FL11	0.67	0.67	0.67	0.78
FL12	1.00	1.00	1.00	0.78
FL13	0.78	0.78	0.89	0.78
Total	0.87	0.85	0.84	0.82

Los coeficientes por reactivo oscilaron entre $V = .67$ y 1.00 en claridad, coherencia y pertinencia, y entre $V = .78$ y $.89$ en relevancia. FL11 presentó los valores más bajos ($V = .67$; relevancia = $.78$), por lo que fue revisado y conservado para su evaluación psicométrica posterior. Dado el reducido número de expertos, la interpretación de la V de Aiken puede complementarse mediante intervalos de confianza (Penfield & Giacobbi, 2004; Merino-Soto & Livia-Segovia, 2009).

Análisis factorial exploratorio AFE - Estudio piloto (n = 56)

Se realizó un estudio piloto con 56 profesionales asistenciales del sistema público de salud para examinar preliminarmente el comportamiento de los reactivos, la estructura interna y la consistencia de la FLS-13, etapa recomendada antes de su evaluación en una muestra independiente (Boateng et al., 2018; Muñiz & Fonseca-Pedrero, 2019).

Consistencia interna preliminar de la FLS-13

Tabla 3*Consistencia interna preliminar de la FLS-13 en el estudio piloto (n = 56)*

	M	DE	Alfa de Cronbach (α)	ω de McDonald
escala	2.98	1.13	0.97	0.971

En el estudio piloto ($n = 56$), la FLS-13 presentó una media global de 2.98 ($DE = 1.13$). La consistencia interna de las puntuaciones fue elevada, con un coeficiente α de Cronbach = $.970$ y un coeficiente ω de McDonald = $.971$. La proximidad entre ambos coeficientes aporta evidencia convergente sobre la confiabilidad de las puntuaciones de la escala. La estimación conjunta de α y ω resulta pertinente debido a que ambos coeficientes se sustentan en supuestos diferentes respecto de la relación entre los reactivos y el constructo latente (McDonald, 1999; Hayes & Coutts, 2020).

Adecuación de la matriz de correlaciones para el análisis factorial

Tabla 4

Adecuación de los datos para el análisis factorial exploratorio de la FLS-13 (n = 56)

Indicador	Resultado
KMO global	0.93
MSA mínimo	0.897
MSA máximo	0.957
Bartlett, $\chi^2(78)$	837
p	< .001

La factorizabilidad de los datos fue favorable, con KMO global = .930 e índices MSA entre .897 y .957. Asimismo, la prueba de esfericidad de Bartlett resultó significativa, $\chi^2(78) = 837$, $p < .001$, respaldando la pertinencia de realizar el análisis factorial (Kaiser, 1974; Hair et al., 2022).

Tabla 5

Cargas factoriales y unicidades de los reactivos de la FLS-13 en el análisis factorial exploratorio (n = 56)

Cargas de los Factores		
	Factor	
	1	Unicidad
07.- Mi jefe proyecta una imagen de responsabilidad; sin embargo, evita asumir su responsabilidad cuando ocurren problemas en el área.	0.92	0.16
05.- Mi jefe afirma priorizar los objetivos institucionales, pero sus decisiones benefician más sus intereses personales que los de la institución.	0.91	0.17
02.- Mi jefe afirma promover la participación del personal, pero toma decisiones sin considerar las opiniones del equipo.	0.91	0.18
06.- Mi jefe utiliza a personas de su confianza para influir en las decisiones del equipo.	0.90	0.19
03.- Mi jefe mantiene una imagen positiva frente a sus superiores, pero con su personal toma decisiones de manera autoritaria.	0.89	0.21
10.- Mi jefe toma decisiones sin considerar el impacto sobre los trabajadores.	0.88	0.23
08.- Mi jefe destaca el esfuerzo del equipo, pero ante sus superiores se atribuye los resultados positivos obtenidos.	0.85	0.28
09.- Mi jefe evita resolver oportunamente los conflictos laborales cuando ello puede afectar su imagen personal.	0.85	0.28
04.- Mi jefe afirma actuar con imparcialidad, pero otorga un trato preferencial a determinados trabajadores.	0.83	0.30
11.- Mi jefe presta poca atención a las necesidades de los trabajadores.	0.80	0.36
01.- Mi jefe dice preocuparse por el bienestar del personal, pero sus acciones demuestran lo contrario.	0.78	0.40
13.- Mi jefe no reconoce el esfuerzo del personal, aun cuando trabaja en condiciones difíciles.	0.78	0.40
12.- Mi jefe prioriza el cumplimiento de los objetivos y las metas, sin considerar las dificultades que enfrenta el personal.	0.72	0.48
<i>Nota.</i> El método de extracción ‘Residuo mínimo’ se usó en combinación con una rotación ‘oblimin’		

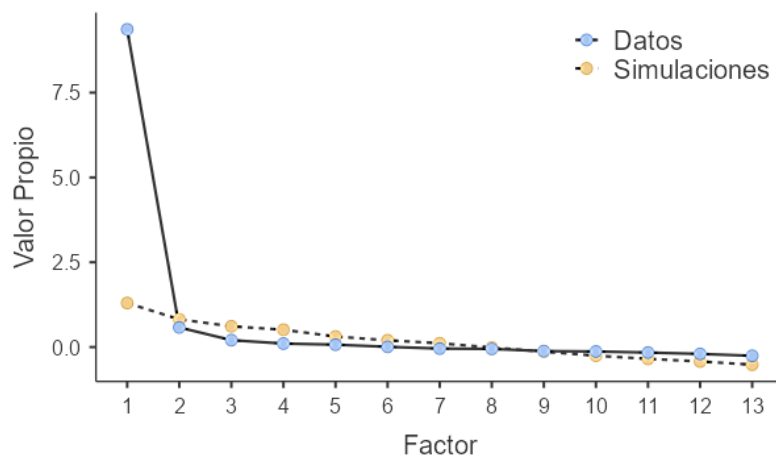
El AFE mediante mínimos residuos (MINRES) con rotación Oblimin mostró cargas factoriales entre $\lambda = .724$ y .917; las mayores correspondieron a FL7 (.917), FL5 (.910) y FL2 (.907), y la menor a FL12 (.724), evidenciando una adecuada representación de los reactivos en

el factor (Costello & Osborne, 2005; Hair et al., 2022). Las unicidades oscilaron entre .160 y .476, indicando que el factor común explicó entre 52.4% y 84.0% de la varianza de los reactivos (Fabrigar et al., 1999; Hair et al., 2022). No se identificaron indicadores que justificaran la eliminación de reactivos, por lo que se conservaron los 13 ítems. La retención factorial se complementó mediante análisis paralelo, procedimiento recomendado frente al uso exclusivo de criterios tradicionales (Hayton et al., 2004).

Determinación del número de factores mediante análisis paralelo

El número de factores a retener se determinó mediante análisis paralelo, comparando los autovalores empíricos con los obtenidos a partir de matrices de datos aleatorios. Este procedimiento constituye una alternativa más robusta que la aplicación exclusiva de criterios tradicionales de retención factorial, al permitir identificar los factores cuyos autovalores exceden los esperados por azar (Horn, 1965; Hayton et al., 2004).

Figura 1
Gráfica de Sedimentación



El análisis paralelo mostró que únicamente el primer autovalor empírico superó al simulado, respaldando la retención de un único factor y, por tanto, la estructura unidimensional propuesta para la FLS-13. Este criterio constituye un procedimiento empírico recomendado para determinar el número de factores a retener en el AFE (Horn, 1965; Hayton et al., 2004).

Validación de la FLS-13 mediante Análisis Factorial Confirmatorio (n = 158)

Se realizó una segunda fase de validación con una muestra independiente de 158 trabajadores del sistema público de salud, lo que permitió contrastar mediante AFC la estructura unidimensional identificada previamente y reducir el riesgo de sobreajuste asociado con la reutilización de la muestra exploratoria (Brown, 2015; Kline, 2023). Se evaluaron las cargas factoriales estandarizadas, los índices de ajuste, la consistencia interna y la validez convergente de la FLS-13.

Características de la muestra de validación (n = 158)

Tabla 6

Características de la muestra de validación de la FLS-13

Característica	n	%
Sexo		
Femenino	113	71.5
Masculino	45	28.5
Sector		
EsSalud	106	67.1
MINSA	21	13.3
Otras instituciones	31	19.6
Ocupación		
Enfermería	86	54.4
Técnico de enfermería	34	21.5
Médico	18	11.4
Auxiliar	13	8.2
Personal administrativo	7	4.4
Región		
Junín	97	61.4
Moquegua	31	19.6
Pasco	19	12
Huánuco	11	7
Régimen laboral		
Decreto Legislativo 728	62	39.2
CAS	60	38
Locación de servicios	24	15.2
Decreto Legislativo 276	12	7.6

La muestra confirmatoria estuvo conformada por 158 participantes, con edades entre 20 y 70 años ($M = 41.85$; $DE = 12.98$). Predominaron las mujeres (71.5%), los participantes de EsSalud (67.1%), los profesionales de enfermería (54.4%) y los procedentes de Junín (61.4%). Respecto al régimen laboral, las mayores proporciones correspondieron al Decreto Legislativo 728 (39.2%) y CAS (38.0%). Las características completas de la muestra se presentan en la Tabla 6.

Análisis descriptivo de los reactivos

Tabla 7

Estadísticos descriptivos de los reactivos de la FLS-13 en la muestra de validación

Descriptivas						
			Asimetría		Curtosis	
	Media	DE	Asimetría	EE	Curtosis	EE
FL1	3.12	1.33	-0.2236	0.193	-1.05	0.384
FL2	3.00	1.31	-0.0687	0.193	-1.11	0.384
FL3	3.01	1.36	-0.0694	0.193	-1.17	0.384
FL4	2.87	1.35	0.0756	0.193	-1.21	0.384

FL5	2.66	1.37	0.1432	0.193	-1.3	0.384
FL6	2.86	1.3	-0.0203	0.193	-1.12	0.384
FL7	2.77	1.33	0.0816	0.193	-1.13	0.384
FL8	2.81	1.31	0.0286	0.193	-1.11	0.384
FL9	2.84	1.24	-0.1373	0.193	-1.02	0.384
FL10	2.85	1.28	0.0159	0.193	-1.11	0.384
FL11	2.91	1.3	-0.1385	0.193	-1.18	0.384
FL12	3.06	1.33	-0.1057	0.193	-1.13	0.384
FL13	2.89	1.38	0.0594	0.193	-1.21	0.384

Previamente al AFC, se examinó el comportamiento descriptivo de los 13 reactivos de la FLS-13 en la muestra de validación ($n = 158$). Las medias oscilaron entre 2.66 (FL5) y 3.12 (FL1), con desviaciones estándar entre 1.24 y 1.38, mientras que la asimetría varió entre -0.224 y 0.143 y la curtosis entre -1.30 y -1.02 , evidenciando variabilidad en las respuestas, baja asimetría y distribuciones moderadamente aplanadas. Considerando estas características y la naturaleza ordinal de la escala Likert de cinco categorías, el AFC se estimó mediante DWLS, procedimiento apropiado para indicadores ordinales (Brown, 2015; Kline, 2023; Li, 2016).

Análisis factorial confirmatorio de la FLS-13

Con el propósito de contrastar la estructura unidimensional identificada en la fase exploratoria, se especificó un modelo factorial confirmatorio de un solo factor latente y 13 indicadores observados, utilizando la muestra independiente de validación ($n = 158$). Debido a la naturaleza ordinal de los reactivos, el modelo fue estimado mediante mínimos cuadrados ponderados diagonalmente (DWLS).

La evaluación del ajuste global consideró de manera conjunta el estadístico χ^2 , el Comparative Fit Index (CFI), el Tucker-Lewis Index (TLI), el Root Mean Square Error of Approximation (RMSEA) y el Standardized Root Mean Square Residual (SRMR), dado que la valoración de un modelo confirmatorio no debe sustentarse en un único índice de ajuste (Brown, 2015; Kline, 2023).

Tabla 8

Índices de ajuste del modelo unifactorial inicial de la FLS-13 ($n = 158$)

Prueba Para un Ajuste Exacto					
χ^2	gl	p			
204	65	<.001			
Medidas de Ajuste					
IC 90% del RMSEA					
CFI	TLI	SRMR	RMSEA	Inferior	Superior
0.929	0.915	0.0363	0.116	0.0985	0.135

El modelo unifactorial inicial presentó un estadístico $\chi^2(65) = 204$, $p < .001$. Los índices incrementales mostraron valores de CFI = .929 y TLI = .915, mientras que el SRMR fue de .036.

Por su parte, el RMSEA alcanzó un valor de .116, con un IC del 90% [.099, .135]. En conjunto, los índices evidenciaron un comportamiento heterogéneo del ajuste: el SRMR presentó un valor favorable, mientras que el CFI y el TLI mostraron una aproximación razonable a los valores de referencia; sin embargo, el RMSEA indicó un desajuste del modelo unifactorial inicial.

Dado que la evaluación del ajuste factorial debe considerar conjuntamente múltiples indicadores y evitar decisiones sustentadas exclusivamente en puntos de corte rígidos (Brown, 2015; Kline, 2023), se procedió a examinar las fuentes específicas de desajuste mediante los índices de modificación, considerando únicamente aquellas posibles reespecificaciones que contaran con justificación conceptual.

Evaluación y reespecificación del modelo confirmatorio

Tabla 9

Covarianzas residuales

Covarianzas Residuales						
		Estimador	EE	Z	p	Estimador Estándar
FL1	FL1	0.737	0.0879	8.38	<.001	0.421
	FL2	0.249	0.0553	4.50	<.001	0.433
FL2	FL2	0.448	0.0568	7.89	<.001	0.262
FL4	FL4	0.600	0.0734	8.17	<.001	0.327
FL3	FL3	0.663	0.0800	8.29	<.001	0.367
FL5	FL5	0.423	0.0550	7.68	<.001	0.226
FL6	FL6	0.476	0.0595	7.99	<.001	0.284
FL7	FL7	0.380	0.0501	7.59	<.001	0.216
FL8	FL8	0.522	0.0644	8.10	<.001	0.307
FL9	FL9	0.586	0.0704	8.32	<.001	0.381
FL10	FL10	0.429	0.0544	7.88	<.001	0.265
	FL11	0.254	0.0520	4.89	<.001	0.485
FL11	FL11	0.642	0.0775	8.28	<.001	0.382
FL12	FL12	0.741	0.0885	8.38	<.001	0.424
FL13	FL13	0.729	0.0879	8.30	<.001	0.387

El análisis del modelo inicial condujo a incorporar dos covarianzas residuales entre FL1–FL2 y FL10–FL11, cuyos estimadores estandarizados fueron .433 ($Z = 4.50$, $p < .001$) y .485 (Z

= 4.89, $p < .001$), respectivamente. Su inclusión se sustentó en la proximidad semántica de los reactivos: FL1–FL2 comparten contenido relacionado con la discrepancia entre una intención favorable expresada por el directivo y su actuación posterior, mientras que FL10–FL11 reflejan la limitada consideración de las necesidades y consecuencias sobre los trabajadores. Estas asociaciones fueron interpretadas como varianza específica compartida no explicada completamente por el factor general, y no como evidencia de factores adicionales.

Tras la reespecificación, el ajuste mejoró de CFI = .929, TLI = .915, SRMR = .036 y RMSEA = .116 a CFI = .962, TLI = .953, SRMR = .032 y RMSEA = .086, IC 90% [.066, .106], con $\chi^2(63) = 137$, $p < .001$. Aunque el RMSEA permaneció por encima de criterios más estrictos, la valoración conjunta de los índices respaldó la plausibilidad de la estructura unidimensional de 13 reactivos, dado que el ajuste factorial debe interpretarse considerando múltiples indicadores y no un único criterio (Hu & Bentler, 1999; Brown, 2015; Kline, 2023). No obstante, la estabilidad de las covarianzas residuales identificadas deberá examinarse en muestras independientes.

Tabla 10

Comparación de los índices de ajuste del modelo inicial y reespecificado de la FLS-13

Modelo	χ^2	gl	p	CFI	TLI	SRMR	RMSEA	IC 90% RMSEA
Inicial	204	65	< .001	0.929	0.915	0.036	0.116	[.099, .135]
Reespecificado	137	63	< .001	0.962	0.953	0.032	0.086	[.066, .106]

El modelo unifactorial inicial presentó indicadores de ajuste heterogéneos, con $\chi^2(65) = 204$, $p < .001$, CFI = .929, TLI = .915, SRMR = .036 y RMSEA = .116, IC 90% [.099, .135]. Si bien el SRMR mostró un valor favorable y los índices incrementales evidenciaron una aproximación razonable al ajuste, el RMSEA indicó la presencia de discrepancias residuales que justificaron examinar posibles fuentes localizadas de desajuste. En consecuencia, se evaluó la reespecificación del modelo considerando conjuntamente criterios estadísticos y la correspondencia conceptual entre los reactivos, evitando introducir modificaciones sustentadas exclusivamente en la mejora de los índices de ajuste (Brown, 2015; Kline, 2023).

La reespecificación incorporó las covarianzas residuales entre FL1–FL2 y FL10–FL11. La primera asociación se consideró conceptualmente justificable debido a que ambos reactivos expresan una discrepancia entre una declaración favorable del directivo y una actuación posterior incongruente: preocupación por el bienestar del personal frente a acciones contrarias (FL1), y promoción de la participación frente a decisiones que omiten las opiniones del equipo (FL2). Por su parte, FL10 y FL11 comparten contenido referido a la consideración de las necesidades y consecuencias sobre los trabajadores, al evaluar respectivamente la toma de decisiones sin considerar su impacto y la escasa atención prestada a las necesidades del personal. En este sentido, las covarianzas residuales fueron interpretadas como varianza específica compartida entre reactivos de contenido próximo, no explicada completamente por el factor general.

Como se observa en la Tabla 9, el modelo reespecificado mostró una mejora sustancial respecto del modelo inicial, $\chi^2(63) = 137$, $p < .001$, CFI = .962, TLI = .953, SRMR = .032 y RMSEA = .086, IC 90% [.066, .106]. Particularmente, el CFI y el TLI superaron .95 y el SRMR permaneció en un nivel favorable, mientras que el RMSEA disminuyó de .116 a .086. Aunque este último indicador permaneció por encima de criterios más estrictos de ajuste, la valoración conjunta de los índices mostró una mejora consistente del modelo reespecificado.

Estos resultados respaldan la plausibilidad de una estructura unidimensional de la FLS-13, manteniendo los 13 reactivos y reconociendo la presencia de varianza residual compartida entre dos pares de indicadores conceptualmente relacionados.

Cargas factoriales y evidencias de validez convergente

Los 13 reactivos presentaron cargas factoriales estandarizadas entre $\lambda = .766$ y .880, sin indicadores que justificaran su eliminación. Asimismo, se obtuvo AVE = .677 y CR = .965, valores que aportan evidencia favorable de validez convergente y confiabilidad compuesta de la FLS-13 (Fornell & Larcker, 1981; Hair et al., 2022).

Tabla 11

Cargas factoriales estandarizadas de la FLS-13 en el modelo confirmatorio reespecificado

Factor	Indicador	Estimador	EE	Estimador Estándar
Factor 1	FL1	1.029	0.0889	0.778
	FL2	1.133	0.0829	0.866
	FL3	1.061	0.0896	0.790
	FL4	1.106	0.0888	0.817
	FL5	1.200	0.0860	0.877
	FL6	1.088	0.0836	0.840
	FL7	1.167	0.0833	0.880
	FL8	1.078	0.0849	0.827
	FL9	0.969	0.0831	0.781
	FL10	1.107	0.0805	0.870
	FL11	1.044	0.0857	0.805
	FL12	1.013	0.0895	0.766
	FL13	1.086	0.0916	0.791

Nota. λ = carga factorial estandarizada; AVE = varianza media extraída; CR = confiabilidad compuesta. Los valores de AVE y CR fueron calculados a partir de las cargas factoriales estandarizadas del modelo confirmatorio reespecificado. AVE = .677; CR = .965.

Consistencia interna de la FLS-13 en la muestra confirmatoria

Tabla 12

Consistencia interna de la FLS-13 en la muestra confirmatoria (n = 158)

Escala	N.º de ítems	α de Cronbach	ω de McDonald
FLS-13	13	.962	.962

En la muestra confirmatoria ($n = 158$), la FLS-13 evidenció una elevada consistencia interna, con valores de α de Cronbach = .962 y ω de McDonald = .962. La coincidencia entre ambos coeficientes proporciona evidencia convergente sobre la precisión de las puntuaciones obtenidas mediante los 13 reactivos.

La estimación conjunta de α y ω permite complementar la evaluación de la confiabilidad, dado que ambos coeficientes se sustentan en supuestos diferentes respecto de la relación entre los indicadores y el constructo latente (McDonald, 1999; Hayes & Coutts, 2020). En conjunto, estos resultados respaldan la adecuada consistencia interna de la FLS-13 en la muestra independiente utilizada para la fase confirmatoria.

Modelo final de la False Leadership Scale (FLS-13)

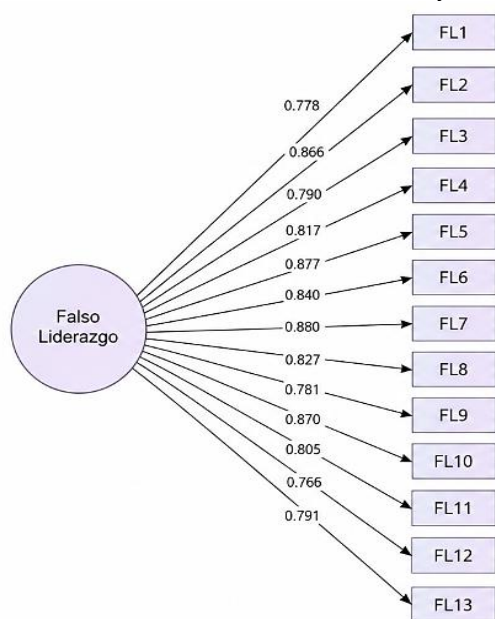
A partir de las evidencias obtenidas en las fases exploratoria y confirmatoria, se estableció como solución final de la False Leadership Scale (FLS-13) una estructura unidimensional integrada por 13 reactivos. En la muestra confirmatoria ($n = 158$), todos los indicadores mostraron cargas factoriales estandarizadas elevadas ($\lambda = .766-.880$), por lo que no fue necesario eliminar reactivos durante la especificación del modelo final.

El modelo confirmatorio reespecificado incorporó las covarianzas residuales entre FL1–FL2 y FL10–FL11, sustentadas en la proximidad conceptual y el contenido compartido de dichos pares de reactivos. Esta especificación mostró una mejora respecto del modelo inicial, alcanzando $\chi^2(63) = 137$, $p < .001$, CFI = .962, TLI = .953, SRMR = .032 y RMSEA = .086, IC 90% [.066, .106]. En consecuencia, la reespecificación permitió conservar la totalidad de los reactivos y mantener la estructura de un único factor latente.

De manera complementaria, la varianza media extraída (AVE = .677) y la confiabilidad compuesta (CR = .965) aportaron evidencia favorable de validez convergente y consistencia del constructo. Asimismo, los coeficientes α de Cronbach = .962 y ω de McDonald = .962 obtenidos en la muestra confirmatoria respaldaron la consistencia interna de las puntuaciones. En conjunto, estos resultados proporcionan evidencia favorable sobre la estructura interna, validez convergente y confiabilidad de las puntuaciones de la FLS-13 en la muestra estudiada, sustentando la solución unidimensional propuesta.

Figura 2

Modelo unidimensional del Constructo y la FLS-13 mediante análisis factorial confirmatorio



Resumen de las propiedades psicométricas

En conjunto, los resultados obtenidos proporcionaron evidencia favorable sobre las propiedades psicométricas de la False Leadership Scale (FLS-13). La evaluación mediante juicio de expertos mostró niveles adecuados de validez de contenido. Posteriormente, el estudio piloto ($n = 56$) evidenció una adecuada factorizabilidad de la matriz de correlaciones y respaldó una estructura unidimensional de 13 reactivos mediante el análisis paralelo y el análisis factorial exploratorio. Asimismo, la escala presentó elevada consistencia interna en esta fase.

En una muestra independiente de validación ($n = 158$), el análisis factorial confirmatorio corroboró la estructura unidimensional propuesta. Tras una reespecificación conceptualmente justificada mediante la incorporación de las covarianzas residuales FL1–FL2 y FL10–FL11, el modelo presentó una mejora sustancial de los índices de ajuste. Las cargas factoriales estandarizadas del modelo final fueron elevadas ($\lambda = .766-.880$), mientras que el AVE = .677 y la CR = .965 aportaron evidencia favorable de validez convergente y confiabilidad compuesta. Adicionalmente, la consistencia interna en la muestra confirmatoria fue elevada ($\alpha = .962$; $\omega = .962$). En conjunto, estos hallazgos respaldan la estructura interna y la confiabilidad de las puntuaciones de la FLS-13 en la población estudiada.

Tabla 13

Resumen psicométrico del estudio ($n = 158$)

Propiedad psicométrica	Evidencia obtenida	Resultado
Validez de contenido	V de Aiken global	0.84
Claridad	V de Aiken	0.87
Coherencia	V de Aiken	0.85
Pertinencia	V de Aiken	0.84
Relevancia	V de Aiken	0.82

Fase exploratoria (n = 56)		
Adecuación muestral	KMO global	0.93
Factorizabilidad	Bartlett	$\chi^2(78) = 837, p < .001$
Dimensionalidad	Análisis paralelo	1 factor
Estructura factorial	AFE	Unidimensional
Cargas factoriales	AFE	.724–.917
Unicidades	AFE	.160–.476
Consistencia interna	α / ω	.970 / .971
Fase confirmatoria (n = 158)		
Modelo final	AFC reespecificado	$\chi^2(63) = 137, p < .001$
Ajuste incremental	CFI / TLI	.962 / .953
Ajuste residual	SRMR	0.032
Ajuste aproximado	RMSEA (IC 90%)	.086 [.066, .106]
Cargas factoriales	AFC final	.766–.880
Validez convergente	AVE	0.677
Confiabilidad compuesta	CR	0.965
Consistencia interna	α / ω	.962 / .962

DISCUSIÓN

El presente estudio tuvo como propósito desarrollar y evaluar las propiedades psicométricas de la False Leadership Scale (FLS-13) como medida del Falso Liderazgo percibido. En conjunto, la evidencia obtenida mediante juicio de expertos, análisis paralelo, AFE y AFC respalda inicialmente una estructura unidimensional de 13 reactivos. La convergencia entre procedimientos exploratorios y confirmatorios resulta relevante en el desarrollo de instrumentos, dado que la evaluación de la dimensionalidad constituye un componente fundamental para determinar si la estructura empírica corresponde con la conceptualización del atributo que se pretende medir (AERA et al., 2014; Boateng et al., 2018).

La estructura unidimensional identificada sugiere que los comportamientos representados por los reactivos convergen en un atributo común, coherente con la conceptualización propuesta del Falso Liderazgo como discrepancia percibida entre la imagen directiva proyectada y las conductas atribuidas al líder. Este resultado adquiere particular significado al contrastarlo con los modelos previamente desarrollados sobre el lado oscuro del liderazgo. La supervisión abusiva sitúa su núcleo en la percepción sostenida de comportamientos hostiles verbales y no verbales del supervisor (Tepper, 2000). Por su parte, Einarsen et al. (2007) conceptualizaron el liderazgo destructivo como comportamientos sistemáticos y repetidos capaces de perjudicar a la organización y/o a sus miembros. Schmidt (2008), en cambio, operacionalizó el liderazgo tóxico mediante cinco dimensiones: supervisión abusiva, liderazgo autoritario, narcisismo, autopromoción e imprevisibilidad. Frente a estos modelos, la FLS-13 focaliza su medición en la incongruencia imagen–conducta, lo que responde a la necesidad planteada por Carton (2022) de

avanzar hacia una mayor especificidad en la conceptualización de los comportamientos de liderazgo.

Los resultados del AFC aportaron evidencia adicional sobre esta estructura, aunque requieren una interpretación prudente. La incorporación de covarianzas residuales entre FL1–FL2 y FL10–FL11, sustentada en la proximidad conceptual de estos pares, mejoró el ajuste del modelo hasta alcanzar CFI = .962, TLI = .953 y SRMR = .032. Sin embargo, el RMSEA = .086, IC 90% [.066, .106], permaneció por encima de criterios más estrictos. Dado que la evaluación de modelos factoriales debe realizarse considerando conjuntamente los diferentes índices y no mediante decisiones basadas en un único estadístico (Brown, 2015; Kline, 2023), estos resultados respaldan la plausibilidad de la solución unidimensional, pero justifican continuar examinando su estabilidad en muestras independientes.

La calidad del modelo de medición también estuvo respaldada por las cargas factoriales estandarizadas de los 13 reactivos ($\lambda = .766-.880$), junto con AVE = .677 y CR = .965, valores que aportan evidencia favorable de validez convergente y confiabilidad compuesta (Fornell & Larcker, 1981; Hair et al., 2022). Asimismo, la elevada consistencia interna obtenida mediante α de Cronbach y ω de McDonald respalda la coherencia de las puntuaciones de la escala, siendo recomendable considerar ambos coeficientes de manera complementaria en la evaluación de la confiabilidad (McDonald, 1999; Hayes & Coutts, 2020).

No obstante, la principal contribución del estudio debe situarse con cautela en el plano conceptual y psicométrico. La estructura unidimensional identificada no demuestra por sí misma que el Falso Liderazgo constituya un constructo empíricamente independiente del liderazgo tóxico, destructivo o de la supervisión abusiva. Esta distinción es particularmente importante ante la proliferación de constructos parcialmente superpuestos en la investigación del liderazgo y la necesidad de evitar categorías redundantes (Carton, 2022). En consecuencia, los resultados permiten sostener que la discrepancia imagen–conducta presenta una estructura interna susceptible de medición mediante la FLS-13, pero su singularidad frente a constructos próximos todavía requiere contrastación empírica.

Por ello, una etapa posterior deberá examinar simultáneamente la FLS-13 y medidas consolidadas de liderazgo tóxico, supervisión abusiva y liderazgo destructivo para establecer su validez discriminante, aspecto esencial para determinar si una nueva medida representa efectivamente un atributo diferenciable de constructos conceptualmente próximos (Boateng et al., 2018). Asimismo, su replicación en diferentes poblaciones y contextos organizacionales permitirá establecer con mayor precisión la estabilidad y alcance de la estructura propuesta. En consecuencia, la FLS-13 debe considerarse una medida con evidencia psicométrica inicial favorable, cuya consolidación dependerá de la acumulación progresiva de nuevas evidencias de validez.

Limitaciones

La FLS-13 fue validada en profesionales del sistema público de salud, por lo que su generalización a otros contextos requiere cautela y replicación en muestras más amplias y diversas. Asimismo, el diseño transversal impidió examinar la estabilidad temporal de las puntuaciones. Aunque el modelo reespecificado mostró indicadores favorables ($CFI = .962$; $TLI = .953$; $SRMR = .032$), el $RMSEA = .086$, IC 90% [.066, .106], y las covarianzas residuales FL1–FL2 y FL10–FL11 requieren contrastación en nuevas muestras independientes. Al tratarse de un constructo emergente, futuros estudios deberán examinar la validez discriminante frente a constructos próximos, la invariancia factorial, la estabilidad test–retest, las relaciones con variables externas y la capacidad predictiva de la FLS-13.

Implicancias prácticas

La FLS-13 constituye una medida breve con evidencias iniciales favorables de validez y confiabilidad para evaluar la percepción de Falso Liderazgo en profesionales del sistema público de salud. Su aplicación puede contribuir al estudio de la discrepancia entre la imagen de integridad, responsabilidad o compromiso proyectada por quienes ejercen funciones directivas y las prácticas percibidas por sus colaboradores, considerando que el liderazgo puede relacionarse con el bienestar laboral (Inceoglu et al., 2018) y con el clima organizacional (Schneider et al., 2013). Asimismo, permitirá examinar sus relaciones con variables como agotamiento, compromiso, satisfacción, confianza, seguridad psicológica y comportamiento de voz, este último especialmente relevante ante la importancia del silencio organizacional señalada por Morrison (2014).

No obstante, dado su reciente desarrollo, las puntuaciones de la FLS-13 no deberían utilizarse aisladamente para diagnósticos individuales ni decisiones administrativas, pues estos usos requieren evidencia de validez suficiente (AERA, APA, & NCME, 2014). Su aplicación en nuevas muestras y contextos permitirá ampliar la evidencia sobre su utilidad y establecer con mayor precisión su diferenciación frente a constructos próximos, aspecto especialmente relevante ante la necesidad de evitar constructos redundantes y favorecer una mayor precisión conceptual en la investigación del liderazgo (Carton, 2022).

CONCLUSIONES

La False Leadership Scale (FLS-13) presentó evidencias psicométricas iniciales favorables para la evaluación del falso liderazgo percibido en la población estudiada. La validez de contenido mediante juicio de expertos, el análisis paralelo y el análisis factorial exploratorio realizado en la muestra piloto ($n = 56$) respaldaron una estructura unidimensional de 13 reactivos, que posteriormente fue contrastada mediante análisis factorial confirmatorio en una muestra independiente ($n = 158$).

El modelo confirmatorio reespecificado, que incorporó las covarianzas residuales conceptualmente justificadas entre FL1–FL2 y FL10–FL11, mostró una mejora sustancial

respecto del modelo inicial, alcanzando CFI = .962, TLI = .953, SRMR = .032 y RMSEA = .086, IC 90% [.066, .106]. Asimismo, las cargas factoriales estandarizadas fueron elevadas ($\lambda = .766-.880$), mientras que la varianza media extraída (AVE = .677) y la confiabilidad compuesta (CR = .965) aportaron evidencia favorable de validez convergente y consistencia del modelo de medición. La elevada consistencia interna observada en la muestra confirmatoria ($\alpha = .962$; $\omega = .962$) complementó estas evidencias.

En conjunto, los resultados respaldan la interpretación de la FLS-13 como una medida esencialmente unidimensional, orientada a evaluar la discrepancia percibida entre la imagen de integridad, responsabilidad o compromiso proyectada por quienes ejercen funciones directivas y determinadas prácticas experimentadas por sus colaboradores. No obstante, al tratarse de un constructo emergente y de una primera validación, estos hallazgos deben entenderse como evidencia inicial y no como una validación definitiva del instrumento.

La FLS-13 abre una línea de investigación para examinar empíricamente la relación del falso liderazgo con variables relevantes del comportamiento organizacional, como el agotamiento laboral, el compromiso laboral, la confianza organizacional, la satisfacción laboral, la seguridad psicológica y el comportamiento de voz. Futuras investigaciones deberán replicar la estructura factorial en muestras independientes, examinar específicamente la estabilidad de las covarianzas residuales identificadas, evaluar la estabilidad temporal, la invariancia factorial y, de manera especialmente relevante, aportar evidencia de validez discriminante frente a constructos conceptualmente próximos, como el liderazgo tóxico, la supervisión abusiva y otras formas de liderazgo destructivo.

Escala de Falso Liderazgo (FLS-13) ©

Las siguientes afirmaciones describen conductas y prácticas que usted puede percibir en la persona que ejerce funciones directivas o de jefatura en su entorno laboral. Lea cada enunciado y señale su grado de acuerdo considerando su experiencia habitual de trabajo. Seleccione una sola alternativa de respuesta, de 1 a 5.

Ítem	Reactivo / Enunciado	Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo
FL1	Mi jefe dice preocuparse por el bienestar del personal, pero sus acciones demuestran lo contrario.	1	2	3	4	5
FL2	Mi jefe afirma promover la participación del personal, pero toma decisiones sin considerar las opiniones del equipo.	1	2	3	4	5
FL3	Mi jefe mantiene una imagen positiva frente a sus superiores, pero con su personal toma decisiones de manera autoritaria.	1	2	3	4	5
FL4	Mi jefe afirma actuar con imparcialidad, pero otorga un trato preferencial a determinados trabajadores.	1	2	3	4	5
FL5	Mi jefe afirma priorizar los objetivos institucionales, pero sus decisiones benefician más sus intereses personales que los de la institución.	1	2	3	4	5
FL6	Mi jefe utiliza a personas de su confianza para influir en las decisiones del equipo.	1	2	3	4	5
FL7	Mi jefe proyecta una imagen de responsabilidad; sin embargo, evita asumir su responsabilidad cuando ocurren problemas en el área.	1	2	3	4	5
FL8	Mi jefe destaca el esfuerzo del equipo, pero ante sus superiores se atribuye los resultados positivos obtenidos.	1	2	3	4	5
FL9	Mi jefe evita resolver oportunamente los conflictos laborales cuando ello puede afectar su imagen personal.	1	2	3	4	5
FL10	Mi jefe toma decisiones sin considerar el impacto sobre los trabajadores.	1	2	3	4	5
FL11	Mi jefe presta poca atención a las necesidades de los trabajadores.	1	2	3	4	5
FL12	Mi jefe prioriza el cumplimiento de los objetivos y las metas, sin considerar las dificultades que enfrenta el personal.	1	2	3	4	5
FL13	Mi jefe no reconoce el esfuerzo del personal, aun cuando trabaja en condiciones difíciles.	1	2	3	4	5

© Ayala & Desouza (2026). La Escala de Falso Liderazgo (FLS-13) puede utilizarse gratuitamente para investigaciones científicas sin fines comerciales. Su uso con fines comerciales y/o no científicos está prohibido, salvo que se cuente con la autorización previa y por escrito de los autores.

Email articulos.alexander.ayala@gmail.com / alexander.ayala@essalud.gob.pe / alexander.ayala@unmsm.edu.pe / mervin.desouza@unmsm.edu.pe

REFERENCIAS

- Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement*, 40(4), 955–959.
- Akinyele, A. I., & Chen, Z. (2025). Dark clouds of leadership: Causes and consequences of toxic leadership. *International Studies of Management & Organization*, 55(4), 476–503. <https://doi.org/10.1080/00208825.2024.2442185>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Ashforth, B. E. (1994). Petty tyranny in organizations. *Human Relations*, 47(7), 755–778.
- Avolio, B. J., & Gardner, W. L. (2005). Authentic leadership development: Getting to the root of positive forms of leadership. *The Leadership Quarterly*, 16(3), 315–338. <https://doi.org/10.1016/j.leaqua.2005.03.001>
- Bass, B. M., & Avolio, B. J. (1994). *Improving organizational effectiveness through transformational leadership*. Sage Publications.
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6, 149. <https://doi.org/10.3389/fpubh.2018.00149>
- Brown, M. E., & Treviño, L. K. (2006). Ethical leadership: A review and future directions. *The Leadership Quarterly*, 17(6), 595–616. <https://doi.org/10.1016/j.leaqua.2006.10.004>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.
- Burns, J. M. (1978). *Leadership*. Harper & Row.
- Carton, A. M. (2022). The science of leadership: A theoretical model and research agenda. *Annual Review of Organizational Psychology and Organizational Behavior*, 9, 61–93. <https://doi.org/10.1146/annurev-orgpsych-012420-091227>
- Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, 10, Article 7. <https://doi.org/10.7275/jyj1-4868>
- De Hoogh, A. H. B., & Den Hartog, D. N. (2008). *Ethical and despotic leadership, relationships with leader's social responsibility, top management team effectiveness and subordinates' optimism: A multi-method study*. *The Leadership Quarterly*, 19(3), 297–311. <https://doi.org/10.1016/j.leaqua.2008.03.002>

- Einarsen, S., Aasland, M. S., & Skogstad, A. (2007). Destructive leadership behaviour: A definition and conceptual model. *The Leadership Quarterly*, 18(3), 207–216. <https://doi.org/10.1016/j.leaqua.2007.03.002>
- Escobar-Pérez, J., & Cuervo-Martínez, Á. (2008). Validez de contenido y juicio de expertos: Una aproximación a su utilización. *Avances en Medición*, 6, 27–36.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272–299. <https://doi.org/10.1037/1082-989X.4.3.272>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.1177/002224378101800104>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2022). *Multivariate data analysis* (8th ed.). Cengage.
- Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But... *Communication Methods and Measures*, 14(1), 1–24. <https://doi.org/10.1080/19312458.2020.1718629>
- Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, 7(2), 191–205. <https://doi.org/10.1177/1094428104263675>
- Hogan, R., & Kaiser, R. B. (2005). What we know about leadership. *Review of General Psychology*, 9(2), 169–180. <https://doi.org/10.1037/1089-2680.9.2.169>
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179–185. <https://doi.org/10.1007/BF02289447>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Inceoglu, I., Thomas, G., Chu, C., Plans, D., & Gerbasi, A. (2018). Leadership behavior and employee well-being: An integrated review and a future research agenda. *The Leadership Quarterly*, 29(1), 179–202. <https://doi.org/10.1016/j.leaqua.2017.12.006>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36. <https://doi.org/10.1007/BF02291575>
- Kline, R. B. (2023). *Principles and practice of structural equation modeling* (5th ed.). Guilford Press.
- Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48, 936–949. <https://doi.org/10.3758/s13428-015-0619-7>

- Lord, R. G., Day, D. V., Zaccaro, S. J., Avolio, B. J., & Eagly, A. H. (2017). Leadership in applied psychology: Three waves of theory and research. *Journal of Applied Psychology*, 102(3), 434–451. <https://doi.org/10.1037/apl0000089>
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates.
- Merino-Soto, C., & Livia-Segovia, J. (2009). Intervalos de confianza asimétricos para el índice de la validez de contenido: Un programa Visual Basic para la V de Aiken. *Anales de Psicología*, 25(1), 169–171.
- Muñiz, J., & Fonseca-Pedrero, E. (2019). Ten steps for test development. *Psicothema*, 31(1), 7–16. <https://doi.org/10.7334/psicothema2018.291>
- Northouse, P. G. (2022). *Leadership: Theory and practice* (9th ed.). Sage.
- Padilla, A., Hogan, R., & Kaiser, R. B. (2007). The toxic triangle: Destructive leaders, susceptible followers, and conducive environments. *The Leadership Quarterly*, 18(3), 176–194. <https://doi.org/10.1016/j.leaqua.2007.03.001>
- Paulhus, D. L., & Williams, K. M. (2002). *The Dark Triad of personality: Narcissism, Machiavellianism, and psychopathy*. *Journal of Research in Personality*, 36(6), 556–563. [https://doi.org/10.1016/S0092-6566\(02\)00505-6](https://doi.org/10.1016/S0092-6566(02)00505-6)
- Penfield, R. D., & Giacobbi, P. R., Jr. (2004). Applying a score confidence interval to Aiken's item content-relevance index. *Measurement in Physical Education and Exercise Science*, 8(4), 213–225. https://doi.org/10.1207/S15327841MPEE0804_3
- Rosenthal, S. A., & Pittinsky, T. L. (2006). Narcissistic leadership. *The Leadership Quarterly*, 17(6), 617–633. <https://doi.org/10.1016/j.leaqua.2006.10.005>
- Schmidt, A. A. (2008). *Development and validation of the Toxic Leadership Scale* [Master's thesis, University of Maryland, College Park]. Digital Repository at the University of Maryland.
- Schneider, B., Ehrhart, M. G., & Macey, W. H. (2013). Organizational climate and culture. *Annual Review of Psychology*, 64, 361–388. <https://doi.org/10.1146/annurev-psych-113011-143809>
- Schneider, B., González-Romá, V., Ostroff, C., & West, M. A. (2017). Organizational climate and culture: Reflections on the history of the constructs in the *Journal of Applied Psychology*. *Journal of Applied Psychology*, 102(3), 468–482. <https://doi.org/10.1037/apl0000090>
- Schyns, B., & Schilling, J. (2013). How bad are the effects of bad leaders? A meta-analysis of destructive leadership and its outcomes. *The Leadership Quarterly*, 24(1), 138–158. <https://doi.org/10.1016/j.leaqua.2012.09.001>
- Tepper, B. J. (2000). Consequences of abusive supervision. *Academy of Management Journal*, 43(2), 178–190. <https://doi.org/10.2307/1556375>

World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA*, 310(20), 2191–2194. <https://doi.org/10.1001/jama.2013.281053> Wolor