ORBICOM

NETWORK OF UNESCO CHAIRS IN COMMUNICATION
RED DE LAS CÁTEDRAS UNESCO DE COMUNICACIÓN
RÉSEAU DES CHAIRES UNESCO EN COMMUNICATION

PROYECTOS COLABORATIVOS (3ª EDICIÓN)

Evaluación y análisis del uso responsable y seguro de Internet y de los dispositivos digitales para la alfabetización mediática (2026-2027)

INFORME

3ª FASE

Universidad del País Vasco
Euskal Herriko Unibertsitatea

Komunikazio eta Hezkuntza Balioak UNESCO Katedra
Cátedra UNESCO de Comunicación y Valores Educativos

unesco Cátedra

UNESCO Chaire in Communication and Educational Values

Chaire UNESCO de Communicaction et Valeurs Educatives

INFORME TÉCNICO DE EVALUACIÓN PSICOMÉTRICA Y VALIDACIÓN TRANSCULTURAL DEL CUESTIONARIO ORBICOM (COMPETENCIA DIGITAL INFANTIL, 9–12 AÑOS)

1. MARCO INSTITUCIONAL Y FUNDAMENTACIÓN METODOLÓGICA DEL PROYECTO ORBICOM

El diseño y la validación cuantitativa de un instrumento internacional estandarizado para evaluar la competencia digital y el uso seguro de las tecnologías en la infancia (9 a 12 años) constituye una prioridad estratégica en la investigación educativa contemporánea. En un entorno evolutivo caracterizado por la exposición temprana a plataformas digitales, redes sociales y dispositivos móviles, la toma de decisiones pedagógicas y el diseño de políticas públicas deben fundamentarse en evidencias empíricas irrefutables. Bajo el liderazgo de la Red Internacional de Cátedras UNESCO en Comunicación (ORBICOM), este proyecto aborda la brecha metodológica existente en la medición de la alfabetización mediática e informacional (MIL) durante la etapa de Educación Primaria (cursos 5º y 6º en España y sus equivalentes internacionales), articulando un esfuerzo multicéntrico orientado a garantizar la validez transcultural, la equivalencia lingüística y la solidez psicométrica del instrumento en diversos contextos socioculturales.

Estructura Organizativa e Hitos del Proyecto

La ejecución de la investigación responde a un esquema de gobernanza colaborativo y estandarizado en el que la Universidad del País Vasco (EHU) asume la dirección centralizada y la coordinación metodológico-estadística. La financiación del proyecto se enmarca formalmente en la *3ª Edición de Proyectos Colaborativos de ORBICOM*, concedida por la Secretaría Ejecutiva de la red en la Université du Québec à Montréal (UQAM).

El cronograma operativo del proyecto abarca desde marzo de 2026 hasta abril de 2027, estructurándose en tres fases secuenciales:

- **Fase 1: Diseño y Adaptación Terminológico-Lingüística (marzo - mayo de 2026):** Definición del cuestionario base por la EHU con asesoramiento de docentes de Primaria, revisión terminológico-gramatical por cada Cátedra socia, traducción a 5 idiomas/variantes (español, euskera, portugués, *brasileiro* y francés) e implementación de 6 versiones en línea en la plataforma www.encuestafacil.com.
- **Fase 2: Aplicación Piloto y Trabajo de Campo (mayo - julio de 2026):** Tramitación de autorizaciones institucionales y consentimientos informados, administración del instrumento en horario lectivo bajo estrictos protocolos de anonimato y custodia centralizada de datos.

- **Fase 3: Análisis Paramétricos, Validación Psicométrica y Consolidación (septiembre de 2026 - abril de 2027):** Análisis empírico de validez de contenido, análisis de consistencia interna (\alpha y \omega), Análisis Factorial Confirmatorio (AFC), estandarización de normas de aplicación e informe final ante la UQAM.

Muestra internacional y caracterización sociodemográfica

Durante el desarrollo del trabajo de campo se registró una captación bruta de N = 461 respuestas en la plataforma telemática. Tras la aplicación de los criterios de depuración de la base de datos (eliminación de registros incompletos, cuestionarios con patrones de respuesta monótonos o incongruencias de edad), la muestra analítica final quedó constituida por N = 429 participantes con respuestas totalmente válidas. Cabe precisar que en la sede de Grenoble (Francia) no fue posible ejecutar la recogida empírica debido a la denegación de los permisos requeridos por parte de la inspección escolar pública regional, concentrándose la recogida en las otras cinco sedes socias.

La distribución geográfica de la muestra analítica se desglosa en las siguientes tablas:

Distribución Muestral por Países y Cátedras Socias

País / Sede Institucional	Frecuencia (N)	Porcentaje Válido (%)
País Vasco (EHU - Donostia / Agurain)	100	23,31%
Brasil (UFRGS - Porto Alegre)	111	25,87%
Portugal (IPL - Lisboa)	83	19,35%
Colombia (UniAndes - Bogotá)	83	19,35%
República Dominicana (PUCMM - Santo Domingo)	52	12,12%
TOTAL MUESTRA ANALÍTICA	**429**	**100,00%**

Distribución Muestral por Género y Edad

Variable Sociodemográfica	Categoría de Respuesta	Frecuencia (N)	Porcentaje (%)
Género	Chico	207	48,25%
	Chica	207	48,25%
	No binario	2	0,47%
	Prefiero no contestar	13	3,03%
Edad	9 años	2	0,47%

	10 años	98	22,84%
	11 años	200	46,62%
	12 años	75	17,48%
	13 o más años	54	12,59%

Desde la perspectiva del muestreo en investigación educativa, la muestra participante acredita una elevada idoneidad psicométrica. Se observa una paridad de género perfecta entre chicos y chicas (48,25% respectivamente). Asimismo, el análisis por edades confirma una concentración idónea en el rango diana del proyecto (10 a 11 años representados en un 69,46% conjunto), asegurando la validez del desarrollo evolutivo para la comprensión de los ítems planteados.

Una vez caracterizada la estructura muestral y verificado el equilibrio sociodemográfico de los participantes, resulta indispensable someter al cuestionario a un análisis empírico sobre la validez de contenido entre las diferentes sedes institucionales.

2. EVALUACIÓN DE LA VALIDEZ DE CONTENIDO TRANSCULTURAL (ESTUDIO MULTICÉNTRICO DE EXPERTOS)

La evaluación de la validez de contenido desde la psicometría avanzada trasciende la mera revisión teórica *a priori*. En este proyecto se ha implantado una evaluación de validez de contenido basada en el uso o validez empírica, aprovechando la información cualitativa y cuantitativa registrada durante la prueba piloto. Someter el instrumento a la valoración de jueces expertos con posterioridad al trabajo de campo permite detectar sesgos culturales, ambigüedades terminológicas, fallos en la traducción contextual y barreras de comprensión cognitiva experimentadas por el alumnado de 9 a 12 años en entornos escolares reales.

Metodología y métricas psicométricas del juicio de expertos

El estudio multicéntrico contó con un panel evaluador integrado por 3 jueces expertos por sede, en cinco países (Colombia, Lisboa, Agurain, Brasil y República Dominicana), quienes valoraron la pertinencia, claridad y adecuación de los reactivos sobre una escala de 0 a 10. La auditoría aplicó un diseño metodológico multivariado basado en seis índices psicométricos fundamentales:

1. **Coeficiente de Correlación Intraclase (ICC):** Evalúa el grado de acuerdo absoluto interjueces mediante un modelo de efectos mixtos de dos vías (Koo & Mae, 2016). Umbral de adecuación: \ge 0,75.
2. **Razón e Índice de Validez de Contenido de Lawshe (CVR / CVI):** Cuantifica la proporción de jueces que valoran un ítem como esencial (Lawshe, 1975; recálculo de Wilson et al., 2012). Umbral aceptable: CVI \ge 0,80.

3. **V de Aiken:** Evalúa la relevancia del reactivo en función del rango de la escala (Aiken, 1980, 1985; Penfield & Giacobbi, 2004). Umbral de aceptación: V \ge 0,70 (exigente V \ge 0,80).
4. **Coeficiente de Validez de Contenido de Hernández-Nieto (CVC):** Mide la validez del ítem corrigiendo el error por azar atribuible al tamaño del panel de expertos (Hernández-Nieto, 2002). Umbral: CVC \ge 0,80.
5. **Índice de Validez de Face (IVF):** Proporción de ítems evaluados como aceptables en apariencia por la mayoría de los jueces (Rubio et al., 2003). Umbral: IVF \ge 80,0\%.
6. **Análisis de Decisión Multicriterio (ADM):** Integración matemática de indicadores para fundamentar la retención o eliminación de reactivos (Claeys et al., 2013).

Análisis Comparativo Global por Sede

Índice Psicométrico	Colombia	Lisboa	Agurain	Brasil	República Dominicana	Umbral de Referencia
Media Global de Valoraciones	7,81	8,97	9,10	8,68	8,69	—
ICC (Concordancia Interjueces)	0,392	0,789	0,940	0,610	0,723	\ge 0,75 (Buena)
V de Aiken (Global)	0,759	0,897	0,910	0,843	0,869	\ge 0,70 / 0,80
CVI / IVC (Lawshe Global)	0,457	0,857	0,963	1,000†	0,693	\ge 0,80
CVC (Hernández-Nieto Global)	0,726	0,854	0,888	0,843	0,832	\ge 0,80
IVF (Rubio et al.)	97,1%	91,4%	94,4%	61,1%‡	100,0%	\ge 80,0\%
ADM (Claeys et al.)	—	8,97	9,10	0,840	0,878	\ge 0,70 / 7,0

† El CVI de Brasil (1,000) fue calculado de manera restringida sobre los 22 ítems retenidos con CVR = 1,00, habiéndose excluido del cómputo global 14 ítems debido al patrón de respuesta bimodal del Experto 1. Por consiguiente, este valor no es directamente comparable con los CVI no restringidos de Agurain (0,963) o Lisboa (0,857). ‡ El IVF de Brasil (61,1\%) se encuentra artificialmente deprimido por las puntuaciones severas del Experto 1 (quien asignó puntuaciones de 1 a 14 reactivos), mientras que los Expertos 2 y 3 valoraron la totalidad de la prueba con puntuaciones \ge 7,0.

El diagnóstico comparativo evidencia una pronunciada heterogeneidad entre las sedes. Mientras que **Agurain** (ICC = 0,940; CVC = 0,888) y **Lisboa** (ICC = 0,789; CVC = 0,854) presentan índices de concordancia y validez excelentes, la sede de **Colombia** arroja una concordancia interjueces deficiente (ICC = 0,392; CVC = 0,726), penalizada por la discrepancia de criterios en los bloques de publicación e Inteligencia Artificial. En **Brasil**, la discordancia del Experto 1 deprimió la consistencia del panel (ICC = 0,610), aunque los coeficientes V de Aiken (0,843) y CVC (0,843) confirman la suficiencia del instrumento tras el filtro cualitativo.

Desglose de Validez de Contenido por Dimensiones Temáticas

Para precisar el comportamiento analítico del instrumento, las métricas V de Aiken y CVC de Hernández-Nieto fueron desglosadas por dimensión en las cinco sedes evaluadas:

Coeficiente V de Aiken por Dimensión Temática

Dimensión Temática	Colombia	Lisboa	Agurain	Brasil	República Dominicana
1. Hábitos de Uso	0,776	0,919	0,914	0,747	0,800
2. Gestión de Datos	0,838	0,933	0,833	0,835	0,852
3. Contenidos	0,748	0,862	0,924	0,776	0,862
4. Relaciones	0,850	0,911	0,922	0,945	0,878
5. Publicaciones	0,734	0,900	0,933	0,945	0,917
6. Inteligencia Artificial	0,675	0,833	0,907	0,882	0,920

Coeficiente CVC de Hernández-Nieto por Dimensión Temática

Dimensión Temática	Colombia	Lisboa	Agurain	Brasil	República Dominicana
1. Hábitos de Uso	0,734	0,868	0,877	0,725	0,764
2. Gestión de Datos	0,773	0,896	0,796	0,849	0,816
3. Contenidos	0,711	0,825	0,887	0,777	0,840
4. Relaciones	0,813	0,874	0,885	0,913	0,846
5. Publicaciones	0,697	0,863	0,896	0,913	0,880
6. Inteligencia Artificial	0,638	0,797	0,870	0,882	0,875

El desglose transversal confirma que la dimensión **Relaciones** representa el núcleo más estable del cuestionario (CVC \ge 0,813 en todas las sedes). En contraste, la dimensión **Inteligencia Artificial** registra la mayor variabilidad cross-cultural: mientras que en

República Dominicana (V = 0,920) y Agurain (V = 0,907) alcanza valores excelentes, en Colombia sufre una caída drástica (V = 0,675; CVC = 0,638), acusando la falta de consenso sobre la madurez cognitiva de los menores frente a la IA.

3. EVALUACIÓN DE LA CONSISTENCIA INTERNA Y FIABILIDAD DE LAS ESCALAS

En la evaluación psicométrica de cuestionarios tipo Likert, la fiabilidad no debe estimarse en exclusiva mediante el coeficiente Alfa de Cronbach, debido a que este indicador asume estricta tau-equivalencia (cargas factoriales idénticas entre reactivos). Por ello, esta investigación calcula de forma coordinada el coeficiente Omega de McDonald (\omega), un estimador no paramétrico altamente robusto ante desviaciones de la tau-equivalencia y heterogeneidad en las varianzas de los ítems.

Análisis de Fiabilidad por Dimensiones Temáticas

Se analizaron las respuestas de la muestra internacional (N = 429) sobre las subescalas del instrumento estandarizado, obteniéndose los siguientes parámetros de consistencia interna:

Subescala / Dimensión Latente	Código de Variable	N.º Ítems	Alfa de Cronbach (\alpha)	Omega de McDonald (\omega)	Evaluación y Calificación Psicométrica
Hábitos de Uso	P13USOSE	7	0,735	0,738	Satisfactoria / Aceptable
Gestión de Datos	P14DATOSE	7	0,853	0,854	Alta precisión de medida
Contenidos / Configuración	P15CONSE	7	0,744	0,748	Satisfactoria / Aceptable
Relaciones Digitales	P16RELSE	6	0,815	0,816	Buena homogeneidad
Publicación de Contenidos	P17PUSE	4	0,706	0,718	Aceptable (escala reducida)
Inteligencia Artificial	P19IASE	4	0,861	0,865	Excelente (DE = 0,974)

Todas las subescalas superan con holgura el umbral psicométrico de aceptabilidad (\ge 0,70). Destacan por su elevada precisión las dimensiones de **Gestión de Datos** (\alpha = 0,853; \omega = 0,854) e **Inteligencia Artificial** (\alpha = 0,861; \omega = 0,865), acreditando una excelente covariación entre sus indicadores observados.

Fiabilidad Global del Cuestionario Mejorado

El análisis integrado sobre la totalidad de la escala mejorada arrojó un **Alfa de Cronbach de 0,916** y un **Omega de McDonald de 0,918**, con una desviación típica global de la puntuación estandarizada de DE = 0,491.

Obtener coeficientes de fiabilidad superiores a 0,90 a nivel global demuestra una consistencia interna excepcional, confirmando que el conjunto articulado de reactivos mide de manera precisa e integrada un macro-constructo latente unificado sobre la competencia y seguridad digital infantil.

Asimismo, la inspección de la estadística "Alfa de Cronbach si se descarta el elemento" demostró la elevada estabilidad de la prueba. Como se deduce de la auditoría de descartes:

- Ninguna eliminación de un ítem individual logra incrementar el valor del Alfa global por encima de 0,916. Los coeficientes resultantes al retirar reactivos oscilan en un intervalo sumamente estrecho entre 0,911 (al eliminar P14Datos5Seg) y 0,916 (al eliminar P15Con2Seg).
- Incluso la retirada de reactivos con cargas factoriales moderadas (como P17Pu2Seg, que mantiene un Alfa de 0,913 tras su exclusión) corrobora que no existen ítems parásitos o que distorsionen la varianza total de la medida.

Demostrada la fiabilidad y la homogeneidad interna de las subescalas, el siguiente paso metodológico consiste en verificar empíricamente si la estructura factorial teórica postulada se sostiene mediante un Análisis Factorial Confirmatorio (AFC).

4. VALIDEZ DE CONSTRUCTO: ANÁLISIS FACTORIAL CONFIRMATORIO (AFC) Y SECUENCIA DE DEPURACIÓN

Desde la perspectiva de la psicometría avanzada, el Análisis Factorial Confirmatorio (AFC) constituye la prueba restrictiva fundamental para contrastar la estructura factorial teórica del cuestionario. Cuando los modelos confirmatorios iniciales presentan desajustes en los índices globales de bondad de ajuste, la metodología exige aplicar una depuración iterativa guiada por criterios teóricos y psicométricos, evitando el sobreajuste estadístico y priorizando la validez de la escala.

Evaluación del Modelo Factorial Inicial

El modelo confirmatorio inicial probó la estructura teórica de 6 factores latentes interrelacionados incorporando la totalidad de los indicadores del cuestionario. Los índices globales de ajuste arrojaron los siguientes valores:

- **Prueba de Ajuste Exacto:** \chi^2 = 978; gl = 545; p < .001; Razón \chi^2/gl = 1,79.
- **Índices de Ajuste Incremental:** CFI = 0,893; TLI = 0,883.
- **Índices de Ajuste Absoluto:** RMSEA = 0,0424 (IC 90%: 0,0381 - 0,0466); SRMR = 0,0636.

La lectura metodológica de este modelo inicial muestra un comportamiento asimétrico: mientras que las métricas de error absoluto (RMSEA \le 0,05 y SRMR \le 0,08) resultaron excelentes, los índices de ajuste comparativo (CFI y TLI) no alcanzaron el umbral convencional de 0,90. Este diagnóstico confirmó la presencia de indicadores ruidosos con cargas factoriales débiles que requerían depuración.

Proceso Iterativo de Depuración de Ítems y Modelos Intermedios

Paso 1: Eliminación del Reactivo PUB2 (P17Pu2Seg)

La auditoría de las cargas estandarizadas del modelo inicial reveló que la subescala de Publicaciones presentaba las estimaciones más frágiles, destacando el ítem P17Pu2Seg con una carga de solo 0,578. Su retirada del cuestionario produjo una reestructuración inmediata de los índices:

- \chi^2 = 899; gl = 512; \chi^2/gl = 1,76.
- **CFI = 0,903**; **TLI = 0,893**. La eliminación de P17Pu2Seg permitió al modelo superar por primera vez el umbral crítico de CFI \ge 0,90, eliminando una fuente considerable de varianza de error.
- RMSEA = 0,0413; SRMR = 0,0597.

Paso 2: Depuración Secundaria de Ítems con Cargas Moderadas

Con el objetivo de consolidar la validez del instrumento y elevar el índice TLI por encima de 0,90, se procedió a la eliminación secuencial de tres indicadores que exhibían cargas inferiores a 0,550 o elevados índices de modificación:

1. P13Seg4 (Carga estandarizada: 0,536).
2. P13Seg2 (Carga estandarizada: 0,581).
3. P15Con5Seg (Carga estandarizada: 0,545).

Evaluación y Cargas Factoriales del Modelo Final Depurado

La eliminación fundamentada de estos cuatro reactivos condujo al **Modelo Final Depurado**, cuyos parámetros satisfacen los estándares internacionales de excelencia psicométrica.

Evolución Comparativa de los Índices de Ajuste Factorial

Métrica / Índice Psicométrico	Modelo Inicial (Todos los ítems)	Modelo Intermedio (Sin P17Pu2Seg)	Modelo Final Depurado	Criterio de Aceptación / Excelencia
Chi-cuadrado (\chi^2)	978	899	**699**	Reducción progresiva
Grados de Libertad (gl)	545	512	**390**	—
Razón \chi^2/gl	1,79	1,76	**1,79**	< 3,00 (Excelente)
CFI (Comparative Fit Index)	0,893	0,903	**0,912**	\ge 0,900 (Aceptable/Bueno)
TLI (Tucker-Lewis Index)	0,883	0,893	**0,902**	\ge 0,900 (Aceptable/Bueno)
RMSEA (Root Mean Sq. Error)	0,0424	0,0413	**0,0423**	\le 0,050 (Excelente)
IC 90% del RMSEA	0,0381 – 0,0466	0,0368 – 0,0457	**0,0372 – 0,0473**	Margen estrecho
SRMR (Standardized Root Mean)	0,0636	0,0597	**0,0578**	\le 0,080 (Bueno)

Cargas Factoriales Estandarizadas del Modelo Final (Sintaxis Oficial de la Base)

- **Factor @USOSE (Uso Seguro / Hábitos):** P13Seg1 (0,549), P13Seg5 (0,673), P13Seg6 (0,618), P13Seg7 (0,603).
- **Factor @DATOSE (Gestión de Datos):** P14Datos1Seg (0,632), P14Datos2Seg (0,742), P14Datos3Seg (0,688), P14Datos4Seg (0,707), P14Datos5Seg (0,716), P14Datos6Seg (0,690), P14Datos7Seg (0,660).
- **Factor @CONSE (Configuración y Contenidos):** P15Con1Seg (0,640), P15Con2Seg (0,587), P15Con3Seg (0,676), P15Con4Seg (0,667), P15Con6Seg (0,613), P15Con7Seg (0,610).
- **Factor @RELSE (Relaciones Digitales):** P16Rel1Seg (0,646), P16Rel2Seg (0,736), P16Rel3Seg (0,681), P16Rel4Seg (0,650), P16Rel5Seg (0,724), P16Rel6Seg (0,614).
- **Factor @PUBSE (Publicación de Contenidos):** P17Pu1Seg (0,558), P17Pu3Seg (0,808), P17Pu4Seg (0,721).

- **Factor @IASE (Inteligencia Artificial):** P19IA1Seg (0,703), P19IA2Seg (0,867), P19IA3Seg (0,814), P19IA4Seg (0,705).

Todas las cargas factoriales estandarizadas del modelo final resultaron estadísticamente significativas ($p < .001$), superando holgadamente el valor de 0,50 y situándose en su mayoría por encima de 0,60.

Diagnóstico Metodológico Especial: Validez Discriminante y Codificación de IA

El examen avanzado del modelo confirmatorio identificó dos dinámicas psicométricas críticas que requieren tratamiento técnico:

1. **Debilidad Estructural del Factor Publicación:** Tras la retirada de un ítem de la dimensión Publicaciones, el P17Pu2Seg, la dimensión Publicaciones se sostiene sobre solo 3 indicadores (P17Pu1Seg, P17Pu3Seg, P17Pu4Seg). Aunque tres ítems representan el mínimo identificado para la identificabilidad en AFC, la modesta carga de P17Pu1Seg (0,558) sitúa a este factor como el más vulnerable del cuestionario.
2. **Anomalía de Codificación en Ítems de IA:** En el procesamiento mediante paquetes estadísticos (jamovi / lavaan), las variables que evalúan la frecuencia e intensidad de uso de IA (P19IA1Uso a P19IA4Uso) registraron correlaciones negativas con la escala total. No obstante, dentro del AFC, las variables enfáticas en la percepción de seguridad y competencia (P19IA1Seg a P19IA4Seg) mostraron cargas estandarizadas excelentes (0,703 a 0,867). Esta diferencia obedece a la distinción psicométrica entre *frecuencia comportamental no guiada* (donde un mayor uso sin supervisión incrementa el riesgo percibido) y *autoeficacia en el uso seguro*.
3. El análisis factorial confirma la validez de constructo del modelo hexafactorial; no obstante, para comprender el comportamiento de la prueba es necesario analizar las singularidades observadas durante la aplicación real en los centros escolares.

5. CONCLUSIONES Y HOJA DE RUTA PARA LA CONSOLIDACIÓN DEL CUESTIONARIO ORBICOM

El proceso de evaluación psicométrica confirma que el Cuestionario ORBICOM sobre Competencia Digital Infantil posee una base técnica sólida, una elevada fiabilidad y una validez de constructo satisfactoria. El instrumento se encuentra en una fase de madurez adecuada para su transferencia e internacionalización, a condición de que se ejecuten las correcciones operativas, sintácticas y de redacción identificadas en este informe antes de su despliegue definitivo en las políticas públicas de la UNESCO.

Dictamen Psicométrico Consolidado

1. **Validez de Contenido Transcultural:** Aceptable a nivel global (V \ge 0,70 en las 5 sedes; CVC \ge 0,80 en 4 de 5 sedes). Requiere intervención prioritaria sobre 5 reactivos transversales que presentaron sesgos de comprensibilidad o desacuerdos entre jueces.
2. **Consistencia Interna y Fiabilidad:** Excelente. Las subescalas acreditan una alta homogeneidad de medida (\alpha = 0,706 a 0,861) y la prueba global mejorada alcanza un nivel de precisión irrefutable (\alpha = 0,916; \omega = 0,918; DE = 0,491).
3. **Validez de Constructo (AFC):** Confirmada. Tras la depuración de 4 ítems ruidosos (P17Pu2Seg, P13Seg4, P13Seg2, P15Con5Seg), el modelo hexafactorial ajusta de forma satisfactoria a los datos empíricos (\chi^2/gl = 1,79; CFI = 0,912; TLI = 0,902; RMSEA = 0,0423; SRMR = 0,0578).

Plan de Acción y Recomendaciones a efectuar en el cuestionario

Para garantizar la máxima calidad psicométrica en la versión internacional definitiva del cuestionario, se fija la siguiente hoja de ruta de obligatorio cumplimiento durante la Fase 3 del proyecto:

1. Para mejorar la validez de contenido, ampliar en **futuras investigaciones** el panel de expertos a un mínimo **de 5–7 jueces** por sede en la segunda ronda de validación, lo que estabilizará el ICC y dotará de discriminación real al CVR de Lawshe.
2. **Revisar la dimensión Inteligencia Artificial** ya que es la más heterogénea: excelente en República Dominicana (V 0.920) y Agurain (V 0.907), pero la peor valorada en Colombia (V 0.675; CVC 0.638), con problemas recurrentes en el ítem de uso de la IA para los deberes.
3. El análisis factorial confirmatorio mostró un ajuste satisfactorio de la estructura propuesta (χ^2/gl = 1.79; CFI = .912; TLI = .902; RMSEA = .042; SRMR = .058), tras la **depuración de 4 ítems ruidosos** (P17Pu2Seg, P13Seg4, P13Seg2, P15Con5Seg), evidenciando una adecuada validez de constructo. Todas las cargas factoriales resultaron significativas ($p < .001$) y, en general, superiores a .50, lo que respalda la representatividad de los indicadores respecto a sus factores latentes.

La adopción de este plan de acción garantiza un instrumento de evaluación psicométricamente irrefutable, transculturalmente equivalente y plenamente exportable para fundamentar las políticas públicas globales de alfabetización mediática en la infancia.

REFERENCIAS

Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. Educational and Psychological Measurement, 40(4), 955–959.

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. Educational and Psychological Measurement, 45(1), 131–142.

Claeys, A. S., Cauberghe, V., & Vyncke, P. (2013). Public Relations Review, 39(2), 128–134.

Hernández-Nieto, R. A. (2002). Contributions to Statistical Analysis. Universidad de Los Andes.

Koo, T. K., & Mae, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15(2), 155–163.

Lawshe, C. H. (1975). A quantitative approach to content validity. Personnel Psychology, 28(4), 563–575.

Penfield, R. D., & Giacobbi, P. R. (2004). Applying a score confidence interval to Aiken's item content-relevance index. Measurement in Physical Education and Exercise Science, 8(4), 213–225.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Research in Nursing & Health, 29(5), 489–497.

Rubio, D. M., Berg-Weger, M., Tebb, S. S., Lee, E. S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. Social Work Research, 27(2), 94–104.

Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. Measurement and Evaluation in Counseling and Development, 45(3), 197–210.