

UNIVERSIDAD CATÓLICA DE CUENCA

UNIVERSIDAD CATÓLICA DE CUENCA

Comunidad Educativa al Servicio del Pueblo

DIRECCIÓN DE POSGRADOS

MAESTRÍA EN PSICOLOGÍA FORENSE

EVALUACIÓN DE LA CREDIBILIDAD DEL TESTIMONIO INFANTIL EN CONTEXTOS FORENSES: REVISIÓN SISTEMATIZADA DE LOS INSTRUMENTOS UTILIZADOS

ARTÍCULO CIENTÍFICO PREVIO A LA OBTENCIÓN DEL TÍTULO DE MAGÍSTER EN PSICOLOGÍA FORENSE

AUTOR: FERNÁNDEZ YASCARIBAY KATHERINE PAMELA

DIRECTOR: DIEGO ESTEBAN FERNANDEZ OLIVO

CUENCA - ECUADOR

2026

DIOS, PATRIA, CULTURA Y DESARROLLO



UNIVERSIDAD CATÓLICA DE CUENCA

Comunidad Educativa al Servicio del Pueblo

DIRECCIÓN DE POSGRADOS

MAESTRÍA EN PSICOLOGÍA FORENSE

EVALUACIÓN DE LA CREDIBILIDAD DEL TESTIMONIO INFANTIL EN CONTEXTOS FORENSES: REVISIÓN SISTEMATIZADA DE LOS INSTRUMENTOS UTILIZADOS

ARTÍCULO CIENTÍFICO PREVIO A LA OBTENCIÓN DEL TÍTULO DE MAGÍSTER EN PSICOLOGÍA FORENSE

AUTOR: FERNÁNDEZ YASCARIBAY KATHERINE PAMELA

DIRECTOR: DIEGO ESTEBAN FERNANDEZ OLIVO

CUENCA - ECUADOR

2026

DIOS, PATRIA, CULTURA Y DESARROLLO

REVMIC Revista Multidisciplinaria Investigación Contemporánea

Artículo de revisión sistemática. Vol. 4 - Nro. 4, **pp. 180 - 203**. Octubre 2026. E-ISSN: 2960-8015

Evaluación de la credibilidad del testimonio infantil en contextos forenses: revisión sistematizada de los instrumentos utilizados

Assessing the Credibility of Child Testimony in Forensic Contexts: A Systematic Review of the Instruments Used

https://doi.org/10.58995/redlic.rmic.v4.n4.a228

Fernández Yascaribay Katherine [1], **Fernández Olivo Diego Esteban** [2]
Campoverde Antonio Xavier [3]

[1] Universidad Católica de Cuenca; kpfernandezy62@est.ucacue.edu.ec. Cuenca, Ecuador.
[2] Universidad Católica de Cuenca; esteban.fernandez@ucacue.edu.ec. Cuenca, Ecuador.
[3] Universidad Católica de Cuenca; antonio.campoverde@est.ucacue.edu.ec. Cuenca, Ecuador

¿Como citar?
Yascaribay Katherine, F., Fernández Olivo, D. E., & Campoverde, A. X. (2026). Evaluación de la credibilidad del testimonio infantil en contextos forenses: revisión sistematizada de los instrumentos utilizados. *Revista Multidisciplinaria Investigación Contemporánea, 4*(4),180-203. https://doi.org/10.58995/redlic.rmic.v4.n4.a228

Recibido: 15-04-2026
Aceptado: 05-08-2026
Publicado: 01-10-2026

Nota del editor
REDLIC se mantiene neutral con respecto a reclamos jurisdiccionales en mensajes publicados y afiliaciones institucionales.

Editorial
Red Editorial Latinoamericana de Investigación Contemporánea (REDLIC)
www.editorialredlic.com

Fuentes de financiamiento
La investigación fue realizada con recursos propios

Conflictos de interés
No presentan conflicto de intereses.

Este es un artículo de acceso abierto, bajo licencia Creative Commons (CC BY 4.0) de Atribución 4.0 Internacional

Resumen

Esta revisión sistemática sintetizó los instrumentos estructurados utilizados para evaluar la credibilidad del testimonio infantil en contextos forenses de presunto abuso sexual. Siguiendo los lineamientos PRISMA 2020, se identificaron estudios publicados entre 2016 y 2026 en Scopus, PsycINFO, PubMed/MEDLINE y SciELO. Se incluyeron once estudios empíricos y se realizó una síntesis narrativa debido a la heterogeneidad metodológica. Los hallazgos muestran que la evaluación de la credibilidad no se apoya en un único método, aunque el Análisis de Contenido Basado en Criterios y la Evaluación de la Validez de la Declaración fueron los procedimientos más reportados. También se identificaron otras herramientas, como el Narrative Coherence Coding Scheme, el Reality Monitoring y criterios judiciales aplicados en contextos específicos. En conjunto, CBCA y SVA muestran el respaldo más consistente, aunque su uso suele ser parcial y debe integrarse con un análisis cuidadoso del contexto forense.

Palabras clave: asexual infantil; psicología forense; testimonio de experto; entrevista psicológica; psicometría

Abstract

This systematic review synthesized the structured instruments used to assess the credibility of children's testimony in forensic contexts involving alleged sexual abuse. Following the PRISMA 2020 guidelines, studies published between 2016 and 2026 were identified in Scopus, PsycINFO, PubMed/MEDLINE, and SciELO. Eleven empirical studies were included, and a narrative synthesis was conducted due to methodological heterogeneity. The findings show that credibility assessment does not rely on a single method, although Criteria-Based Content Analysis and Statement Validity Assessment were the most frequently reported procedures. Other tools were also identified, such as the Narrative Coherence Coding Scheme, Reality Monitoring, and judicial criteria applied in specific contexts. Overall, CBCA and SVA show the most consistent support, although their use is often partial and must be integrated with a careful analysis of the forensic context.

Keywords: child asexual; forensic psychology; expert testimony; psychological interview; psychometrics.

1. Introducción

El abuso sexual contra niños, niñas y adolescentes constituye una forma grave de violencia y, al mismo tiempo, representa un problema complejo dentro del sistema de justicia, debido a las dificultades que existen para probar los hechos denunciados. Se estima que más de 370 millones de menores han sufrido violencia sexual antes de cumplir los 18 años, y que esta situación se presenta con mayor frecuencia en contextos socioeconómicos vulnerables (UNICEF, 2025). En el ámbito de las pericias forenses, el testimonio infantil suele ocupar un lugar central, aunque su credi-

bilidad puede verse influida por el nivel de desarrollo del niño, la forma en que se realiza la entrevista y la posible presencia de presiones externas (Cabezas-García et al., 2022).

Aunque existen métodos e instrumentos estructurados para evaluar la credibilidad, como el Análisis de Contenido Basado en Criterios o CBCA y la Evaluación de la Validez de la Declaración o SVA, la evidencia disponible todavía se encuentra dispersa y no siempre muestra resultados consistentes. Esto ocurre, en parte, porque los estudios utilizan definiciones y medidas distintas (Arce, 2017). Por esta razón, no siempre es claro qué instrumentos se aplican con mayor frecuencia, de qué manera se usan en contextos forenses y qué respaldo empírico tienen en casos específicos de abuso sexual infantil. A su vez, esto limita la toma de decisiones informadas en la práctica pericial, la formación profesional y los criterios de valoración judicial.

En este estudio, la credibilidad se entiende como el grado en que un relato representa de forma confiable un hecho, evaluado mediante procedimientos que usan criterios explícitos (Yancce, 2020). En cuanto a la evaluación, los instrumentos se consideran estructurados si contienen reglas explícitas para su aplicación o puntuación, como criterios, escalas de codificación o protocolos de evaluación, y pueden ser utilizados por otros evaluadores (Wouters et al., 2025). Las evaluaciones no estructuradas son aquellas que, en esta investigación, confían principalmente en el juicio clínico, y que, en general, no contienen criterios sistemáticos. El rendimiento, en esta idea, se refiere a la capacidad de un instrumento para informar adecuadamente sobre la veracidad de las narrativas, o sobre la falsedad de las narrativas (Niveau, 2021).

Las evaluaciones forenses de testigos infantiles en casos sospechosos de abuso sexual infantil que se realizan a nivel internacional dependen en gran medida de instrumentos estructurados, y están dirigidas a la evaluación de la indicación de pistas relacionadas con narrativas verdaderas y fabricadas, así como indicadores relacionados con la memoria verbal y no verbal del niño. En términos generales, los estudios en el campo sugieren que varios criterios del CBCA muestran un alto grado de efectividad, con el tamaño del efecto siendo mediano a grande. El modelo *Reality Monitoring* también permite identificar la diferencia entre recuerdos reales y construidos con un efecto moderado. Sin embargo, estos hallazgos no pueden transferirse a todos los contextos legales de manera irrestricta y, en muchos casos, se recomienda que, para lograr un grado razonable de certeza en la evaluación, se integren metodologías (Gancedo et al., 2021; Sporer y Masip, 2025).

Sin embargo, el aspecto metodológico sigue siendo un punto clave; la solidez de los hallazgos depende de cómo se define y verifica la llamada "verdad de base", y también de cómo se estandariza la aplicación de los protocolos. Para poner eso en práctica, en el caso del CBCA, se han encontrado pocos estudios rigurosos, y en casos reales es difícil hacer un juicio claro sobre lo que ocurrió, lo que afecta la generalización. En cuanto al SVA, la evidencia revisada sugiere que su aplicación tiende a ser parcial, ya que en algunos casos se aplican criterios de contenido que no integran otras partes del protocolo. Además, hay discontinuidad en las publicaciones, lo que podría reflejar un desarrollo

discontinuo del propio campo. Por estas razones, esto disminuye la comparabilidad entre estudios (Sporer y Masip, 2025; Wouters et al., 2024).

Dentro del contexto latinoamericano e iberoamericano, este diálogo se amplía aún más porque se discute el tema de la técnica de credibilidad/factores psicosociales dentro de un entorno penal junto con las demandas del castigo. Se han descrito asociaciones de variables socio-familiares con juicios de credibilidad, lo que inyecta variabilidad, impactando potencialmente las interpretaciones de los expertos. La calidad del uso del SVA también puede ser críticamente cuestionada sin descripciones de procedimientos o criterios mínimos en los informes, mientras que la admisibilidad del CBCA y el SVA como evidencia científica también es debatible. Esto se debe en parte a la ausencia de validaciones empíricas con la población local. El debate, por lo tanto, no es solo técnico, sino también legal y contextual (Cabezas-García et al., 2022; Guzmán-Díaz, 2024; Palomares-Rodríguez et al., 2024).

Basado en datos de Ecuador, la evidencia existente sugiere prácticas expertas heterogéneas en la evaluación del testimonio infantil en casos de presunto abuso sexual. No siguiendo un enfoque único para todos, los expertos utilizan una variedad de protocolos de entrevista y recolección de narrativas, como la Entrevista Cognitiva Mejorada, el Protocolo de Michigan, NICHD, RATAC, la Guía de Entrevista Investigativa GEV y propuestas dirigidas a elicitación de señales cognitivas (Fisher y Geiselman, 1992; Cabañas, 2003; Anderson et al., 2010; Muñoz et al., 2016). Colectivamente, estas estrategias contribuyen a los problemas específicos de investigación presentados para proporcionar un enfoque para estructurar la interacción con el niño, minimizando la sugestión, fomentando la recuperación narrativa de los eventos y aumentando la calidad del testimonio para informar evaluaciones de credibilidad posteriores.

Pero la literatura también menciona la falta de protocolos comprensivos, estandarizados a nivel nacional y actualizados, lo que plantea diferencias en la adquisición, registro e interpretación de declaraciones. Además, los estudios empíricos ecuatorianos son limitados y dependen extensamente de la literatura internacional, lo que también contribuye a la falta de capacitación especializada y actualización técnica, todo lo cual impacta la consistencia metodológica de la práctica forense en el país (Méndez Rodas, 2025; Ramírez Romero, 2023).

Por lo tanto, sigue siendo un requisito sintetizar, de manera fresca y holística, qué instrumentos estructurados se emplean, cómo se utilizan y qué rendimiento y limitaciones se reportan por esos instrumentos bajo condiciones similares. Esto es importante, ya que las síntesis pasadas han sido parciales, enfocadas en un solo método o de menor escala y menos específicas al contexto. Por lo tanto, el valor de esta investigación radica en consolidar instrumentos, las condiciones de uso y las disparidades entre regiones para derivar recomendaciones relevantes para el uso forense (Méndez Rodas, 2025; Ramírez Romero, 2023).

Con base en lo anterior, se plantea la siguiente pregunta de investigación: en menores de edad con presunto abuso sexual en contextos forenses, ¿qué instrumentos estructurados se utilizan para evaluar la credibilidad del testimonio infantil? Para

responderla, el objetivo general es sintetizar los instrumentos utilizados para evaluar la credibilidad del testimonio infantil en contextos forenses de presunto abuso sexual, mediante una revisión sistematizada de la literatura.

El aporte esperado es ofrecer una base organizada y verificable que permita fortalecer la práctica pericial, apoyar procesos de estandarización y formación profesional, e informar decisiones institucionales sobre protocolos de evaluación en casos de presunto abuso sexual infantil. Además, la síntesis permitirá identificar vacíos que orienten futuras investigaciones, sobre todo en contextos con menor producción empírica, con apoyo de la estrategia de selección y síntesis reportada en el apartado de Métodos bajo criterios PRISMA.

2. Metodología

Diseño

Se realizó una revisión sistemática y el reporte se elaboró siguiendo los criterios de PRISMA 2020. La síntesis fue narrativa, debido a que no se realizó metaanálisis. Esta decisión se tomó porque los estudios encontrados presentaron una alta heterogeneidad metodológica, es decir, variaron en sus diseños, en la forma de definir la credibilidad, en los procedimientos de entrevista y codificación, así como en las métricas reportadas. Por este motivo, no fue posible realizar una comparación cuantitativa adecuada. La búsqueda bibliográfica se desarrolló entre el 10 de diciembre de 2025 y el 15 de marzo de 2026, sin actualización posterior.

Criterios de elegibilidad

Los criterios de elegibilidad se organizaron mediante el marco PEO. La población estuvo conformada por niños, niñas y adolescentes menores de 18 años, presuntas víctimas de abuso sexual, cuyo testimonio fue obtenido en contextos forenses. La exposición correspondió al uso de instrumentos o métodos estructurados para evaluar la credibilidad o validez del testimonio. Los outcomes incluyeron el tipo de instrumento de evaluación de credibilidad utilizado, el método de aplicación, el desempeño reportado y las limitaciones declaradas por cada investigación.

Los estudios con personas menores de 18 años que habían estado asociadas con entrevistas forenses o dictámenes periciales fueron elegibles para ser incluidos. También se incluyeron estudios si se basaban en instrumentos estructurados con respaldo científico. Además, dada la heterogeneidad de la investigación, se aceptaron varios diseños; estudios cuantitativos experimentales y no experimentales, estudios cualitativos, métodos mixtos y análisis basados en casos judiciales. Solo se consideraron artículos revisados por pares en español, inglés y portugués publicados entre 2016 y 2026.

Se excluyeron revisiones secundarias, editoriales u otros textos sin datos empíricos, artículos sin texto completo, duplicados y estudios que involucraban poblaciones mayores de 18 años sin resultados desagregados para menores. También se excluyeron estudios que no trabajaban en contextos forenses o que no evaluaban la credibilidad a través de procedimientos estructurados.

Finalmente, en caso de que faltara información clave para comparar los resul-

tados principales, el estudio fue excluido durante la fase de texto completo. La calidad metodológica y el riesgo de sesgo se evaluaron formalmente utilizando las herramientas MMAT 2018, dado que la muestra incluía estudios con diseños metodológicos heterogéneos (Hong et al., 2018).

Los estudios se clasificaron de acuerdo con la misma categoría de la guía, según lo indicado por las preguntas de selección y los criterios para estudios cualitativos, cuantitativos descriptivos o cuantitativos no aleatorizados. Como el MMAT recomienda no cuantificar la puntuación general, estos informes descriptivos y basados en criterios se combinaron para formar una síntesis narrativa basada en la calidad, las limitaciones y los términos de utilización de los instrumentos identificados.

Fuentes de información

Se utilizaron las bases de datos Scopus, PsycINFO, PubMed y SciELO. No se buscó literatura gris. Además, se estudiaron las referencias de los artículos incluidos para encontrar estudios potencialmente calificables. No se contactó a autores o instituciones para obtener más información.

Estrategia de búsqueda

Esta búsqueda utilizó vocabulario controlado, así como términos libres, aunque la ecuación se adaptó a la sintaxis y campos disponibles en cada base de datos. Scopus utilizó una estrategia amplia en el título, resumen y palabras clave para máxima sensibilidad. Los términos psicológicos y forenses se enfatizaron en el título y resumen de PsycINFO, con uso complementario de descriptores cuando estaban disponibles. Para PubMed, los términos MeSH se combinaron con la terminología libre en el título/resumen para mantener la sensibilidad dentro de un tema interdisciplinario. En SciELO se empleó una estrategia más corta y bilingüe, orientada hacia el título y resumen, para optimizar la recuperación de literatura iberoamericana. Las ecuaciones para el conjunto reproducible por base de datos se pueden ver en la Tabla 1.

Tabla 1. Campos utilizados por base

Base de datos	Ecuación reproducible
Scopus	TITLE-ABS-KEY((child* OR adolescen* OR minor* OR niño* OR niña* OR adolescente* OR menor*) AND ("sexual abuse" OR "child sexual abuse" OR "sexual assault" OR "abuso sexual" OR "violencia sexual") AND (forensic OR judicial OR legal OR court OR "forensic interview" OR "forensic assessment" OR forense OR "entrevista forense" OR "evaluación pericial") AND (CBCA OR "criteria-based content analys*" OR SVA OR "statement validity assessment" OR "validity checklist" OR "reality monitoring" OR RM OR SCAN OR "scientific content analys*" OR NICHD OR "National Institute of Child Health and Human Development") AND (validit* OR reliab* OR "inter-rater" OR "diagnostic accuracy" OR sensitivit* OR specificit* OR AUC OR "likelihood ratio" OR LR OR validez OR confiabil* OR "precisión diagnóstica" OR sensibilidad OR especificidad)) AND PUBYEAR AFT 2015 AND PUBYEAR BEF 2027 AND DOCTYPE(ar) AND (LANGUAGE(english) OR LANGUAGE(spanish) OR LANGUAGE(portuguese))
APA PsycINFO	((TI(child* OR adolescen* OR minor* OR niño* OR niña* OR adolescente* OR menor*) OR AB(child* OR adolescen* OR minor* OR niño* OR niña* OR adolescente* OR menor*)) AND (TI("sexual abuse" OR "child sexual abuse" OR "sexual assault" OR "abuso sexual" OR "violencia sexual") OR AB("sexual abuse" OR "child sexual abuse" OR "sexual assault" OR "abuso sexual" OR "violencia sexual")) AND (TI(forensic OR judicial OR legal OR court OR "forensic interview" OR "forensic assessment" OR forense OR "entrevista forense" OR "evaluación pericial") OR AB(forensic OR judicial OR legal OR court OR "forensic interview" OR "forensic assessment" OR forense OR "entrevista forense" OR "evaluación pericial")) AND (TI(CBCA OR "criteria-based content analys*" OR SVA OR "statement validity assessment" OR "reality monitoring" OR NICHD OR NaCCS OR "narrative coherence") OR AB(CBCA OR "criteria-based content analys*" OR SVA OR "statement validity assessment" OR "reality monitoring" OR NICHD OR NaCCS OR "narrative coherence")))

PubMed	((("Child Abuse, Sexual" OR "Sex Offenses" OR "sexual abuse"[tiab] OR "child sexual abuse"[tiab] OR "sexual assault"[tiab] OR "abuso sexual"[tiab] OR "violencia sexual"[tiab]) AND ("Child" OR "Adolescent" OR child*[tiab] OR adolescen*[tiab] OR minor*[tiab] OR niño*[tiab] OR niña*[tiab] OR adolescente*[tiab] OR menor*[tiab]) AND ("Forensic Sciences" OR forensic[tiab] OR judicial[tiab] OR legal[tiab] OR court[tiab] OR "forensic interview*"[tiab] OR "forensic assessment"[tiab] OR forense[tiab] OR "entrevista forense"[tiab] OR "evaluación pericial"[tiab]) AND (CBCA[tiab] OR "criteria-based content analys*"[tiab] OR SVA[tiab] OR "statement validity assessment"[tiab] OR "validity checklist"[tiab] OR "reality monitoring"[tiab] OR RM[tiab] OR NICHD[tiab] OR NaCCS[tiab] OR "narrative coherence"[tiab]))) AND (2016:2026[pdat]) AND (english[la] OR spanish[la] OR portuguese[la])
SciELO	(child* OR niño* OR niña* OR adolescente* OR minor* OR menor*) AND ("sexual abuse" OR "child sexual abuse" OR "abuso sexual" OR "violencia sexual") AND (forensic OR forense OR judicial OR "entrevista forense" OR "evaluación pericial") AND (CBCA OR SVA OR "statement validity assessment" OR "criteria-based content analysis" OR "reality monitoring" OR NICHD)

Nota. Fuente: autores (2026)

Las referencias fueron organizadas en Zotero. La selección de estudios se realizó en dos fases. Primero, se revisaron títulos y resúmenes. Después, se evaluó el texto completo de los artículos potencialmente elegibles. En ambas fases participaron dos revisores y las diferencias se resolvieron mediante consenso. Los motivos de exclusión durante la fase de texto completo fueron registrados, y el flujo del proceso se reportó mediante el diagrama PRISMA versión 2020. El cribado fue gestionado en Excel mediante una tabla.

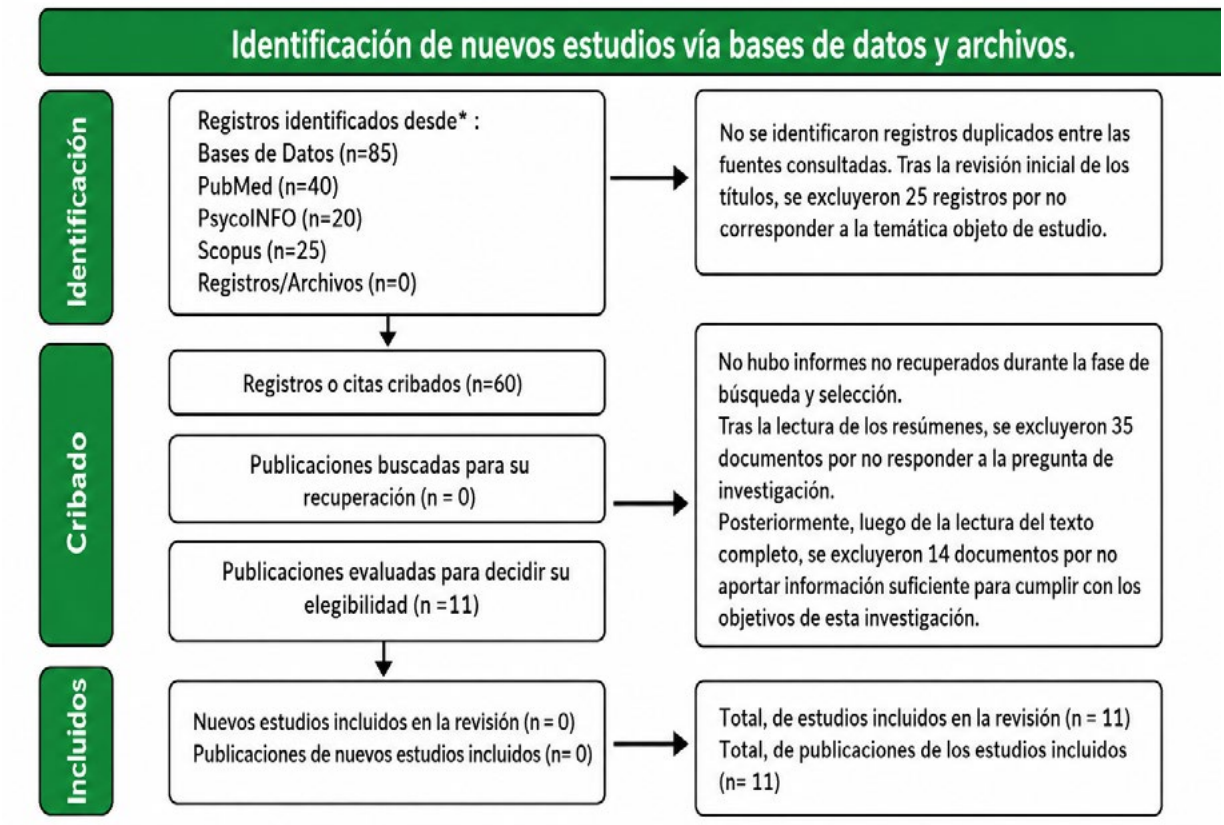
Extracción de datos

La extracción de datos se realizó en formato Excel. Se recopiló un conjunto de información de identificación del estudio (autor, año, país, objetivo, diseño y contexto). Además, cuando se presentaron los datos, se registraron las características de la muestra de los participantes, incluyendo tamaño, edad y género. También se incluyó el contexto forense: tipo de entrevista, evaluación experta y condiciones relevantes. Para los instrumentos, se extrajo información sobre su forma de aplicación, proceso de codificación, evaluadores participantes, escalas utilizadas y condiciones de uso. También se documentó la evidencia sobre el rendimiento y las limitaciones según lo reportado

Diagrama PRISMA

A continuación, en la Figura 1, se presenta el diagrama PRISMA de extracción de datos (Page et al., 2021).

Figura 1. Diagrama PRISMA



Nota. Fuente: autores (2026)

Análisis de calidad metodológica

El análisis de calidad metodológica se llevó a cabo mediante el MMAT 2018, ya que el corpus incluyó diseños heterogéneos. En general, los estudios formularon preguntas claras y aportaron datos pertinentes, lo que permitió sostener una síntesis narrativa sobre instrumentos como SVA, CBCA, NICHD, NaCCS, RCQ y BASC. No obstante, se observaron diferencias en las muestras, la procedencia de los expedientes, los contextos institucionales y el control de variables externas. Por ello, la evidencia resulta útil para comprender el uso forense de los instrumentos de credibilidad, aunque su generalización debe interpretarse con cautela. Siguiendo los criterios del MMAT, no se calculó un puntaje global, sino que se presentaron los criterios de valoración de forma detallada.

Tabla 2. Análisis de calidad metodológica con criterios MMAT

Aspecto	Resultado
Cribado MMAT	11/11 estudios con "Yes" en S1 y S2
Fortaleza principal	Objetivos claros, datos empíricos pertinentes y uso de instrumentos estructurados
Debilidad principal	Representatividad limitada y control desigual de confusores
Valoración global	Calidad metodológica moderada

Nota. Fuente: autores (2026)

Análisis de riesgo de sesgo

El análisis del riesgo de sesgo demostró que se tomaron evidentes y robustas precauciones metodológicas, particularmente de acuerdo con la naturaleza del muestreo. Muchas investigaciones se llevaron a cabo con archivos existentes, archivos judiciales o registros forenses completos, como es común en la ciencia forense, pero esto puede limitar su generalización. Además, existían riesgos potenciales de medición y confusión (las suposiciones no siempre se validaron para el cumplimiento del protocolo con el mismo detalle y para variables como la edad, la calidad de la entrevista, la gravedad del abuso o el historial psicológico). Por lo tanto, los resultados permiten la identificación de patrones accionables, pero deben ser considerados con precaución.

Tabla 3. Análisis de sesgo con criterios MMAT

Aspecto	Resultado
Sesgo más frecuente	Sesgo de selección
Sesgos secundarios	Medición, confusión y datos incompletos
Causa principal	Uso de archivos, expedientes y casos judicializados
Valoración global	Riesgo de sesgo moderado

Nota. Fuente: autores (2026)

3. Desarrollo

3.1. Caracterización de artículos incluidos

Como vemos en la Tabla 4, los estudios encontrados son diversos en su diseño, tamaño de muestra y material analizado. Estudios con grandes muestras, como el de Melkman et al. (2017) con 1,563 niños, y estudios más pequeños centrados en declaraciones, entrevistas, informes de expertos, casos judiciales o frases codificadas, cada uno con un tamaño que varía de

31 a 273 casos. En cuanto a la población, hay niños y adolescentes predominantemente en una amplia categoría de edad, entre aproximadamente 3 y 17 años, pero también se incluyen estudios que se centran en niños en edad preescolar.

Con respecto a los instrumentos utilizados, se observa un predominio de métodos estructurados para analizar la credibilidad del testimonio, sobre todo el CBCA y el SVA. Estos aparecen de forma independiente o combinados con otros recursos, como la entrevista National Institute of Child Health and Human Development (NICHD), la Lista de Validez (VCL, por sus siglas en inglés Validity Checklist), el Sistema de Codificación de Coherencia Narrativa (NaCCS, por sus siglas en inglés Narrative Coherence Coding Scheme) el Reality Criteria Questionnaire (RCQ), los dictámenes expertos de credibilidad y el análisis de criterios aplicados en decisiones judiciales. Esta variedad de herramientas muestra que la evaluación de la credibilidad infantil en el contexto forense no se basa en un solo procedimiento, sino en distintos modelos que parten de enfoques teóricos y periciales diferentes.

Tabla 4. Caracterización de artículos incluidos

Autor	Muestra	Instrumento	Resultado principal / limitaciones
(Melkman et al., 2017)	1,563 niños (3–14 años)	Evaluación clínica + entrevista NICHD	57.9% creíbles; sin datos de validez ni confiabilidad; alta dependencia del testimonio infantil.
(Niveau, 2021)	96 declaraciones (6–16 años)	SVA + CBCA + detalles sensoriales	CBCA discrimina credibilidad; criterio de incrustación contextual mejor predictor; detalles sensoriales asociados a relatos creíbles.
(Bilginer et al., 2021)	31 declaraciones	CBCA	Concordancia total aceptable pero baja en varios criterios; no debe usarse de forma aislada.
(Wouters et al., 2025)	75 informes periciales	SVA (NICHD + CBCA + VCL)	Más fiable que evaluación clínica, pero aplicación inconsistente y limitada en menores de 6 años.

(Welle et al., 2016)	60 declaraciones	CBCA	Fiabilidad intere-valuador ≈0.74; precisión 75%; edad influye en resultados.
(VanMeter et al., 2023)	80 entrevistas (7–15 años)	NaCCS	Alta fiabilidad (>0.9); coherencia mayor en narrativas iniciales; limitación: uso de transcripciones.
(Ernberg et al., 2016)	nr. (preescolares)	Criterios judiciales suecos	Difíciles de aplicar en preescolares; sin evi-dencia psicométrica.
(Roberts y Lamb, 2010)	42 declaraciones (3–16 años)	RCQ (Reality Moni-toring)	Mayor presencia de criterios de realidad en niños mayores; evi-dencia aún limitada para detectar engaño.
(Cabezas-García, et al., 2022)	99 casos de abuso sexual infantil (4–17 años)	Evaluación de credi-bilidad del testimonio / dictamen experto C-NC	Factores psicosociales se asociaron al dicta-men de credibilidad; no reporta métricas psicométricas directas del instrumento.
(García-Rico et al., 2024)	109 casos de abuso sexual infantil	Entrevistas forenses + informes profesiona-les; SVA	La psicopatología previa influyó en la credibilidad; estudio centrado en variables clínicas, no en desem-peño autónomo del SVA.
(Miragoli y Camisas-ca, 2022)	264 fallos judiciales surcoreanos; 273 casos codificados	CBCA — 19 criterios codificados en senten-cias judiciales	Cantidad de deta-lles y estado mental predijeron veredictos; analiza sentencias, no entrevistas originales.

Nota. Fuente: autores (2026)

3.2. Instrumentos estructurados identificados en la revisión

Los estudios revisados muestran que la evaluación de la credibilidad del testimonio infantil en contextos forenses no se realiza a partir de un solo instrumento. Dentro del conjunto analizado, el CBCA es el que aparece con mayor frecuencia, ya sea como técnica principal o como parte del SVA. Este patrón se observa en investigaciones desarrolladas en Suiza, Turquía y otros contextos forenses europeos, donde el análisis del contenido del relato ocupa un lugar central en la valoración pericial de la credibilidad (Niveau, 2021; Bilginer et al., 2021; Welle et al., 2016; Wouters et al., 2025).

Además de CBCA, varios colaboradores describieron SVA como un enfoque de evaluación más amplio. En este sentido, SVA no se considera una técnica aislada; más bien, sirve para fusionar la entrevista forense, el análisis de contenido y los factores de validez. En esta línea, Wouters et al. (2025) lo describen en informes periciales judiciales, mientras que Niveau (2021) lo utiliza junto con la información sensorial obtenida del Monitoreo de la Realidad. Asimismo, Cabezas-García et al. (2022) lo aplicaron para clasificar casos como creíbles o no creíbles, y García-Rico et al. (2024) lo utilizaron con la ayuda de entrevistas forenses no inductivas y BASC para examinar la relación entre psicopatología y credibilidad. Esto indica que la credibilidad generalmente se determina en la práctica real de manera agregada y no solo con un criterio.

Además de estos instrumentos, también se identificaron enfoques centrados en la organización y calidad de la narrativa. En el Reino Unido, VanMeter et al. (2023) emplearon NaCCS para analizar la coherencia narrativa de entrevistas de Logro de la Mejor Evidencia o ABE. De igual manera, Miragoli y Camisasca (2022) utilizaron NaCCS para examinar la coherencia narrativa en declaraciones judiciales relacionadas con revelaciones de abuso sexual infantil. Sin embargo, Roberts y Lamb (2010) utilizan RCQ, que proviene del Monitoreo de la Realidad. Además, Ernberg et al. (2016) informa sobre los criterios judiciales suecos utilizados para evaluar testimonios de niños en edad preescolar.

También se identificó un modelo de evaluación menos estructurado, basado en el juicio clínico posterior a la entrevista. Melkman et al. (2017) describen una valoración de credibilidad apoyada en la declaración del niño, su conducta durante la entrevista y la remisión del caso, con uso del protocolo NICHD, pero sin una guía específica de análisis. Estos resultados son importantes porque permiten observar que, junto a los instrumentos estructurados, todavía existen prácticas en las que la valoración depende más del criterio profesional que de un sistema formal de codificación.

3.3. Tipo de aplicación de los instrumentos

La forma de aplicación varía según el instrumento, aunque en varios estudios se repite una base común. Tanto en el SVA como en el CBCA, la evaluación suele partir de entrevistas forenses videograbadas o transcritas, realizadas en contextos de abuso sexual infantil por personal especializado. En los trabajos de Niveau (2021), Wouters et al. (2025) y Welle et al. (2016), el protocolo NICHD se utiliza como marco de entrevista y luego el análi-

sis se concentra en el contenido del relato. Esto permite una revisión más sistemática del testimonio y favorece la comparación entre casos.

En los estudios centrados en el CBCA, el procedimiento de aplicación fue más estandarizado. Bilginer et al. (2021) aplicaron los 19 criterios a entrevistas forenses semiestructuradas y videograbadas, y cada criterio fue puntuado en una escala de 0 a 2 por cuatro evaluadores independientes. De manera similar, Welle et al. (2016) analizaron transcripciones de entrevistas policiales, y tres evaluadores entrenados calificaron cada criterio con la misma escala para obtener una puntuación total. En ambos casos, la codificación dependió de la formación de los evaluadores y de la calidad del registro de la entrevista.

El SVA fue aplicado como un procedimiento más amplio y secuencial. En el estudio de Wouters et al. (2025), primero se realizó una entrevista policial videograbada con protocolo NICHD y después psicólogos o psiquiatras forenses valoraron la credibilidad mediante CBCA y VCL. En Niveau (2021), dos expertos entrenados en SVA evaluaron las declaraciones, mientras que un investigador ciego codificó los detalles sensoriales. En Cabezas-García et al. (2022), se revisaron informes psicológicos periciales emitidos por profesionales de la Administración de Justicia; los casos fueron clasificados como creíbles o no creíbles mediante SVA y luego se analizaron variables psicológicas, sociofamiliares y del abuso. En García-Rico et al. (2024), la credibilidad fue evaluada mediante entrevistas forenses no inductivas y SVA, complementadas con informes profesionales y BASC para recopilar información sobre psicopatología previa y posterior. Esta forma de aplicación muestra que el SVA requiere varias fases de análisis y participación de personal especializado, lo que puede volver más compleja su aplicación en la práctica pericial.

Los estudios sobre coherencia narrativa y Reality Monitoring siguieron procedimientos distintos. En VanMeter et al. (2023), se trabajó con transcripciones de entrevistas ABE y se codificaron la cronología, la consistencia y el mantenimiento del tema en escalas de 0 a 5, con dos evaluadores independientes. Miragoli y Camisasca (2022) analizaron declaraciones obtenidas en procesos penales con condena; las entrevistas fueron semiestructuradas, grabadas, transcritas y codificadas mediante NaCCS por evaluadores ciegos a la edad y al PTSD. Roberts y Lamb (2010) analizaron transcripciones línea por línea, con codificadores entrenados y ciegos al estado del caso, valorando información perceptiva, espacial, temporal, afectiva y cognitiva. En cambio, Ernberg et al. (2016) y sus colaboradores describen criterios judiciales usados por fiscales y tribunales, más vinculados a la decisión procesal que a una codificación técnica formal.

3.4. Desempeño y limitaciones evidenciadas

La evidencia recopilada indica que el CBCA es uno de los instrumentos más empíricos utilizados, pero los hallazgos no son consistentes en todos estos instrumentos. Niveau (2021) observó que múltiples criterios del CBCA mejoraron la distinción entre declaraciones creíbles, y la incrustación contextual emergió como el mejor predictor. Welle et al. (2016) informaron una precisión general del 75% así como una fiabilidad interevaluador aceptable. Sin embargo, Bilginer et al. (2021)

revelaron que, aunque la puntuación general representaba un alto acuerdo, varios otros criterios mostraron baja concordancia, principalmente entre profesionales con diferente formación.

En conjunto, los hallazgos sugieren que el CBCA puede ser útil, pero no debe tomarse como un criterio definitivo para la toma de decisiones. Tanto Bilginer et al. (2021) como Welle et al. (2016) destacan que su interpretación aislada es inadecuada, que además la edad del niño, la calidad de la entrevista y la formación del evaluador pueden afectar los resultados. En el estudio de Niveau (2021), los testimonios creíbles se asociaron con detalles sensoriales, pero se sugirieron como indicadores separados y complementarios. Por lo tanto, la calidad del CBCA parece estar influenciada por su combinación con otros procedimientos y las circunstancias relacionadas con el uso del procedimiento.

El SVA se presenta como un enfoque más sólido y amplio en comparación con la evaluación clínica no estructurada en teoría. Wouters et al. (2025) lo describen como más confiable que la evaluación clínica, aunque no proporcionaron cifras directas de validez o fiabilidad en su estudio. También mencionaron la aplicación parcial o inconsistente del procedimiento, la mínima aclaración sobre el uso del VCL y las reentrevistas en algunos casos. En Cabezas-García et al. (2022), los elementos psicosociales se vincularon al juicio de credibilidad y se consideraron información complementaria, y no reemplazaron el análisis del testimonio, pero tampoco facilitaron la confirmación de la clasificación C/NC. En García-Rico et al. (2024), la falta de psicopatología previa fue indicativa de mayor credibilidad y el estado psicológico posterior fue independiente de este factor, sin embargo, al ser un estudio exploratorio, no evaluó el desempeño autónomo del SVA. Esto limita la comparación entre informes periciales y demuestra que el valor del SVA no se basa solo en el instrumento, sino también en cómo se administra dentro del ámbito judicial.

Los resultados de los estudios revisados demostraron que el NaCCS mostró un desempeño estable con altos niveles de fiabilidad interevaluador. VanMeter et al. (2023) también informaron alta fiabilidad mientras demostraban narrativas iniciales más coherentes. Los hallazgos de Miragoli y Camisasca (2022) revelaron que la edad estaba positivamente relacionada con la coherencia narrativa, mientras que el TEPT estaba negativamente relacionado con el contexto, la cronología y el tema. Pero estos estudios trataban con transcripciones o declaraciones judiciales, y su alcance estaba limitado por circunstancias como la muestra judicial, el uso de categorías dicotómicas de TEPT o el impacto potencial de aspectos sociales y el estilo de entrevista. Esto sugiere que el NaCCS es útil para la calidad narrativa de un relato, pero su área de enfoque parece ser el análisis de la coherencia frente al control de credibilidad.

Los enfoques de Monitoreo de la Realidad demuestran una eficacia más moderada, con mayores reservas para el análisis forense. En Niveau (2021) se encontraron detalles sensoriales en relatos creíbles y predijeron la credibilidad, aunque se propusieron como indicadores complementarios. Roberts y Lamb (2010) también informaron que hay más criterios de realidad en niños mayores que en niños más jóvenes, pero informaron hallazgos variables entre estudios. Por lo tanto, el RM y el RCQ parecen ofrecer alguna información

relevante para la calidad del relato, pero no son suficientes para respaldar de manera independiente el juicio experto.

Finalmente, los estudios que no emplearon instrumentos completamente desarrollados demostraron mayores limitaciones. Melkman et al. (2017) informaron una alta prevalencia de casos considerados creíbles con poca información sobre validez, fiabilidad o acuerdo entre evaluadores. Ernberg et al. (2016) afirman que los criterios suecos para el cribado judicial son difíciles de cumplir en niños en edad preescolar y dependen en gran medida de la calidad de la entrevista y el desarrollo del niño. En general, estos resultados sugieren la preferencia por instrumentos con reglas de aplicación más claras, aunque siempre bajo un cuidadoso escrutinio de las circunstancias forenses.

3.5. Discusión

Los hallazgos indican que no existe un único instrumento utilizado para evaluar la credibilidad del testimonio de los niños en casos de presunto abuso sexual. En cambio, es un campo mixto y el CBCA ocupa la posición más evidente y aparece frecuentemente como una parte indispensable del SVA en la literatura revisada. Pero esta diversidad no debe tomarse solo como un activo: también significa que todavía no hay una forma confiable de evaluar las historias de los niños.

Se puede afirmar que el CBCA es la fortaleza porque proporciona criterios claros para evaluar el contenido del testimonio. Sin embargo, ciertos instrumentos se han utilizado de manera independiente, como una forma de sustitución, para un examen más amplio. Por lo tanto, el CBCA o el SVA por sí solos no garantizan que la evaluación haya sido exhaustiva, homogénea o suficientemente completa. Esto es significativo porque el problema no es solo qué instrumento se utiliza, sino también cómo se utiliza, qué criterios se emplean y en qué circunstancias de expertos esto puede ocurrir.

La tendencia observada coincide con la revisión sistemática de Wouters et al. (2024), quienes encontraron que ningún estudio adoptó el SVA completamente y se centró principalmente en el CBCA. De manera similar, Fernandes et al. (2024) identificaron 30 procedimientos forenses diferentes y mencionaron que el protocolo NICHD tuvo el mayor uso. Tales coincidencias apuntan al hecho de que, si bien la evaluación de la credibilidad es un campo amplio y heterogéneo, no se aplica de manera consistente. Como resultado, la diversidad de metodologías también necesita ser considerada de cerca y tomada muy en serio, para abrir nuevas áreas de análisis, pero también para dificultar la comparación y desarrollar estándares similares de calidad para el trabajo forense.

Los resultados de esta revisión también sugieren que, en la práctica, los instrumentos de evaluación de credibilidad dependen mucho de la forma en que se obtiene el relato. Por este motivo, la relación entre el instrumento utilizado y el método de entrevista debe analizarse con atención. La entrevista no es un paso secundario, ya que de ella surge el material que después será evaluado. Si durante la entrevista se utilizan preguntas sugestivas, se limita la evocación libre o no se considera adecuadamente la edad y el desarrollo del niño, el instrumento puede terminar analizando un relato que ya fue influido desde el inicio. En ese sentido, la calidad de la entrevista funciona como una condición previa para que el instrumento

tenga mayor utilidad. Un instrumento estructurado puede ordenar el análisis, pero no puede corregir por completo los problemas producidos durante la obtención inicial del testimonio.

Lo anterior permite comprender por qué el CBCA y el SVA aparecen con mayor frecuencia en estudios europeos, mientras que otros enfoques, como la coherencia narrativa o el Reality Monitoring, ocupan un papel más complementario (Niveau, 2021; Bilginer et al., 2021; Welle et al., 2016). En términos aplicados, esto indica que la psicología forense no debería considerar como equivalentes el protocolo de entrevista y el instrumento de valoración, porque cumplen funciones diferentes dentro del proceso pericial. Es decir, el protocolo ayuda a obtener el relato en mejores condiciones, mientras que el instrumento permite valorar ciertos indicadores de credibilidad. Por eso, no deben confundirse sus funciones. Si se confunden, se podría asumir de manera errónea que el instrumento es capaz de corregir los problemas derivados de una entrevista mal realizada. Esta idea coincide con lo reportado por Fernandes et al. (2024) y Melinder et al. (2021).

Por su parte, Fernandes et al. (2024) concluyeron que, pese a la diversidad de procedimientos, las entrevistas suelen compartir una estructura semejante, compuesta por introducción, rapport, práctica narrativa, evocación libre, preguntas de profundización y cierre. Al mismo tiempo, Wouters et al. (2024) observaron que la mayoría de estudios sobre SVA trabajaron con transcripciones y que el método no se utilizó de manera completa, lo que limita la comparabilidad entre investigaciones. Así, que las entrevistas tengan una estructura parecida no significa que la evaluación de credibilidad se realice de forma uniforme. Aunque varios estudios comparten fases similares, esto no garantiza que los relatos obtenidos tengan la misma calidad ni que los instrumentos puedan aplicarse con igual nivel de precisión.

Dicho esto, la existencia de una estructura similar no significa que las aplicaciones del mismo enfoque sean iguales en todos los estudios. Los hallazgos incluyen diferencias en el número de evaluadores, en el uso de escalas, en el tipo de registro analizado y en la adaptación basada en la edad. Estas diferencias son más que detalles técnicos; pueden alterar cómo se desarrolla la misma historia. Por ejemplo, la experiencia del evaluador, su formación o la información contextual pueden influir en la puntuación de un criterio CBCA. El instrumento en sí mismo no tiene menos influencia en su fiabilidad que los criterios internos, al igual que en las condiciones humanas e institucionales en las que se administra el instrumento. Así, la estandarización no podría considerarse como la aplicación de un instrumento mecánico, sino más bien como un mecanismo de aplicación claro, bien documentado y coherente, consistente con la naturaleza del caso.

Esto está de acuerdo con lo que sugirieron Melinder et al. (2021), quienes también encontraron evidencia de que un protocolo más acorde con la interacción con el niño estaba asociado con respuestas más ricas y fomentaba preguntas más directas y profundas. Además, Danby et al. (2022) muestran que las herramientas de revisión estructurada contribuyen a la evaluación de entrevistas, pero pueden ser rígidas y menos flexibles en el futuro, ya que se requerían mejoras más personalizadas. Colectivamente, estas contri-

buciones demuestran que la técnica por sí sola no es suficiente. El procedimiento requeriría formación, supervisión y criterios claramente establecidos para ajustarse sin sacrificar el rigor.

En los resultados, parece que el CBCA constituye el apoyo más directo para discriminar la credibilidad, mientras que otros métodos parecen proporcionar pistas complementarias sobre la coherencia o la naturaleza de la memoria. Pero este estímulo no debe verse como una habilidad natural para evaluar la veracidad del testimonio. El CBCA también mide los indicadores de contenido verbal, aunque estos indicadores se verán afectados por la edad y por el conocimiento del idioma, el nivel de calidad de la entrevista, el nivel de experiencia traumática, así como los métodos para establecer la verdad básica. Como tal, su utilidad será mejor cuando se utilice dentro del contexto de otros aspectos del caso y no como evidencia final.

Esta interpretación es consistente con la revisión de Sporer y Masip (2023), que señala que el CBCA se ha aplicado principalmente para denuncias de abuso sexual infantil. También está de acuerdo con el meta-análisis de Sporer et al. (2023), que informa tamaños de efecto para criterios individuales mayores que los observados en una serie de otras señales de engaño verbal, pero que son sensibles a la calidad metodológica y al proceso de verdad fundamental. Como resultado, el CBCA puede reforzar el juicio experto, pero no puede reemplazarlo ni funcionar como una ecuación de decisión cerrada. Su principal contribución, por lo tanto, es informar la historia en lugar de proporcionar una conclusión precisa sobre la naturaleza de las circunstancias presentadas.

Los resultados también destacan que el rendimiento no solo está determinado por el instrumento, sino también por las condiciones en la población y la entrevista. En niños pequeños, por ejemplo, el ajuste del procedimiento parece decisivo. Esto es importante, porque el desarrollo infantil no es solo un factor externo (por ejemplo, la presencia de los padres), sino una condición que altera la forma en que el niño recuerda, organiza y expresa la experiencia. Por eso, la ausencia de información, la menor coherencia respecto al tiempo y las respuestas mínimas no reflejan automáticamente baja credibilidad, especialmente en niños más pequeños. No basta con considerar la fortaleza del instrumento y analizar sus implicaciones; también es importante considerar las limitaciones del desarrollo infantil en la narración de historias.

Magnusson et al. (2021) encontraron ligeras diferencias entre el SI y el NICHD en información, mayor precisión con el NICHD y posibles efectos de fatiga con un mayor rapport inicial. Otra línea Szojka et al. (2020), quienes demostraron que tener múltiples entrevistas puede aumentar la completitud de la historia sin amplificar contradicciones. En consecuencia, la evaluación de la credibilidad no debe leerse rígidamente a partir de la historia del niño y debemos tener en cuenta que un testimonio menos elaborado no implica automáticamente un nivel de verdad más bajo, y puede sugerir límites evolutivos, emocionales o contextuales. Así, una evaluación más rigurosa debe incorporar el instrumento, los resultados de la entrevista y la etapa de desarrollo del niño, al tiempo que se consideran las especificidades del caso antes de llegar a una opinión experta.

La principal limitación en esta revisión no solo trata con la dimensión psi-

cométrica, sino también con los aspectos de implementación. El SVA incompleto, la aplicación poco clara o inexistente del VCL, la dependencia de transcripciones y la variación en la formación de los evaluadores dificultan la comparación entre estudios. Esto está de acuerdo con Wouters et al. (2024), con Danby et al. (2022), quienes advierten que una guía estructurada no es un sustituto de la retroalimentación experta.

Además, esta investigación no incluyó literatura gris y los protocolos institucionales u otros documentos técnicos relevantes pueden estar excluidos. También se consideró un corpus heterogéneo que abarca diseños, estudios, contextos judiciales, instrumentos y métricas, lo que dificultó un meta-análisis. Aunque se incluyó un análisis riguroso de la calidad metodológica y el riesgo potencial de sesgo en el MMAT 2018, los resultados están destinados a ser interpretados como una síntesis de tendencias críticas en lugar de una estimación aproximada del nivel de rendimiento de cada instrumento.

4. Conclusiones

Esta investigación tuvo como objetivo recopilar y evaluar las herramientas disponibles para valorar la validez del testimonio infantil en los sistemas forenses de presunto abuso sexual, y el objetivo se alcanzó satisfactoriamente. El análisis realizado demostró que existen múltiples formas de evaluar el testimonio infantil en este campo. Entre estas, los instrumentos son estructurados, y muchos enfatizan el análisis del testimonio del niño. Destacan entre estos instrumentos el CBCA y el SVA, además del NaCCS, herramientas de Monitoreo de la Realidad, y algunos criterios judiciales en contextos particulares.

Esta investigación también permitió comprender la manera en que se utilizan los métodos, describiendo así las condiciones del uso de los mismos. La revisión destacó que los métodos se emplean predominantemente en entrevistas forenses grabadas en video o transcritas realizadas por un profesional forense. En muchos casos, las entrevistas se llevan a cabo siguiendo los protocolos NICHD o ABE. Sin embargo, la revisión también describió que en los diversos estudios analizados, no se describió un método de análisis completamente homogéneo debido a la variabilidad en los jueces, los sistemas de puntuación, las transcripciones y la edad y desarrollo del niño entrevistado. Esto mostró que los métodos tienen múltiples diseños y que también cumplen una función plural en los contextos forenses.

Finalmente, hemos sintetizado la evidencia documentada relacionada con el funcionamiento y las deficiencias de los instrumentos respectivos para facilitar su uso durante una evaluación forense. En general, tanto el CBCA como el SVA contaron con el respaldo más sólido en cuanto a su consistencia en los estudios revisados. Además, la relevancia de ambos instrumentos disminuye cuando se usan de forma aislada, fragmentada o cuando las condiciones de la entrevista no están suficientemente controladas. Adicionalmente, NaCCS y los enfoques de Monitoreo de la Realidad ofrecen perspectivas útiles respecto a la calidad de la entrevista y la coherencia del testimonio, pero no pueden sustituir una evaluación integral del caso. Teniendo esto en cuenta, la elección de las herramientas debe considerar la edad del niño, la calidad de la entrevista y el enfoque principal de la evaluación.

5. Contribución de los autores

F-YK: Redacción borrador original
FD: Supervisión, revisión y edición
CA: Revisión y edición.

6. Agradecimientos

Este estudio se desarrolló en el marco de la Maestría en Psicología Forense de la Facultad de Posgrados de la Universidad Católica de Cuenca, a la cual los autores agradecen por el apoyo académico proporcionado a lo largo de este proceso.

7. Referencias

Anderson, J., Ellefson, J., Lashley, J., Miller, A. L., Olinger, S., Russell, A., Stauffer, J., y Weigman, J. (2010). The CornerHouse forensic interview protocol: RATAC®. *Thomas M. Cooley Journal of Practical y Clinical Law, 12*, 193–331. https://www.docsity.com/es/docs/protocolo-satac-para-entrevistas-forenses-de-cornerhous/8034378/

Arce, R. (2017). Análisis de contenido de las declaraciones de testigos: Evaluación de la validez científica y judicial de la hipótesis y la prueba forense. *Acción Psicológica, 14*(2), 171–190. https://doi.org/10.5944/ap.14.2.21347

Bilginer, Ç., Güllü, B. Ü., Orhan, G., Kalaycı, B. M., Erden, G., y Koçtürk, N. (2021). Interrater reliability of criteria based content analyses of sexually abused children's statement in Turkey. *Turkish Journal of Child and Adolescent Mental Health, 28*(1), 52–59. https://doi.org/10.4274/tjcamh.galenos.2020.97269

Cabañas, E. (2003). Traducción del protocolo de entrevista forense Michigan [Documento en línea]. Criminalística.mx. https://criminalistica.mx/descargas/documentos/pdf/PROTOCOLO%20DE%20ENTREVISTA%20FORENSE.pdf

Cabezas-García, M., de la Peña Fernández, M. E., y Andreu-Rodríguez, J. M. (2022). Factores psicosociales asociados a la credibilidad del testimonio en abuso sexual infantil. *Anales de Psicología / Annals of Psychology, 38*(2), 307–315. https://doi.org/10.6018/analesps.418161

Danby, M. C., Sharman, S. J., y Guadagno, B. (2022). Testing an evaluation tool to facilitate police officers' peer review of child interviews. *Journal of Police and Criminal Psychology, 37*(3), 711–725. https://doi.org/10.1007/s11896-022-09516-8

Ernberg, E., Tidefors, I., y Landström, S. (2016). Prosecutors' reflections on sexually abused preschoolers and their ability to stand trial. *Child Abuse y Neglect, 57*, 21–29. https://doi.org/10.1016/j.chiabu.2016.06.001

Fernandes, D., Gomes, J. P., Albuquerque, P. B., y Matos, M. (2024). Forensic interview techniques in child sexual abuse cases: A scoping review. *Trauma, Violence, y Abuse, 25*(2), 1382–1396. https://doi.org/10.1177/15248380231177317

Fisher, R. P., y Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing: The cognitive interview.* Charles C Thomas Publisher.

Gancedo, Y., Fariña, F., Seijo, D., Vilariño, M., y Arce, R. (2021). Reality monitoring: A meta-analytical review for forensic practice. *The European Journal of Psychology Applied to Legal Context, 13*(2), 99–110. https://doi.org/10.5093/ejpalc2021a10

García-Rico, M. A., Ricarte, J. J., Jimeno, M. V., y Latorre, J. M. (2024). Role of psychopathology on children credibility of sexual abuse testimony. *International Journal of Law and Psychiatry, 94,* Article 101972. https://doi.org/10.1016/j.ijlp.2024.101972

Guzmán-Díaz, C. A. (2024). La prueba CBCA *(Criterial-Based Content Analysis o Análisis de Contenido Basado en Criterios) y SVA (Statement Validity Assessment o Evaluación de la Validez de la Declaración) en el proceso penal colombiano* [Trabajo de fin de máster, Universidad de Barcelona]. Dipòsit Digital de la Universitat de Barcelona. https://hdl.handle.net/2445/219320

Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., y Vedel, I. (2018). *Mixed Methods Appraisal Tool (MMAT), version 2018:* User guide. McGill University. http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf

Magnusson, M., Joleby, M., Ernberg, E., Akehurst, L., Korkman, J., y Landström, S. (2021). Preschoolers' true and false reports: Comparing effects of the Sequential Interview and NICHD protocol. *Legal and Criminological Psychology, 26*(1), 83–102. https://doi.org/10.1111/lcrp.12185

Melinder, A., Magnusson, M., y Gilstrap, L. L. (2021). What is a child-appropriate interview? Interaction between child witnesses and police officers. *International Journal on Child Maltreatment: Research, Policy and Practice, 3,* 369–392. https://doi.org/10.1007/s42448-020-00052-8

Melkman, E. P., Hershkowitz, I., y Zur, R. (2017). Credibility assessment in child sexual abuse investigations: A descriptive analysis. *Child Abuse y Neglect, 67,* 76–85. https://doi.org/10.1016/j.chiabu.2017.01.027

Méndez Rodas, A. F. (2025). *Credibilidad de testimonio en víctimas de abuso sexual infantil para psicólogos peritos en Cuenca-Azuay en el año 2024* [Trabajo de titulación de maestría, Universidad de Cuenca]. Repositorio Institucional de la Universidad de Cuenca. https://dspace.ucuenca.edu.ec/handle/123456789/46907

Miragoli, S., y Camisasca, E. (2022). An examination of the synergy of age and PTSD on narrative coherence in child sexual abuse testimony. *Journal of Child Sexual Abuse, 31*(6), 743–757. https://doi.org/10.1080/10538712.2022.2131669

Muñoz, J. M., González-Guerrero, L., Sotoca, A., Terol, O., González, J. L., y Manzanero, A. L. (2016). La entrevista forense: Obtención del indicio cognitivo en menores presuntas víctimas de abuso sexual infantil. *Papeles del Psicólogo, 37*(3), 205–216. https://dialnet.unirioja.es/servlet/articulo?codigo=5689649

Niveau, G. (2021). Sensory information in children's statements of sexual abuse. *Forensic Sciences Research, 6*(2), 97–102. https://doi.org/10.1080/20961790.2020.1814000

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, Article n71. https://doi.org/10.1136/bmj.n71

Palomares-Rodríguez, J. M., Bustos-Berruezo, A., Calatrava-Urán, J. L., y Quevedo-Blasco, R. (2024). Contrainforme psicológico pericial de un caso de abuso sexual infantil. *Revista Iberoamericana de Psicología y Salud, 15*(1), 27–37. https://doi.org/10.23923/j.rips.2024.01.073

Ramírez Romero, M. C. (2023). *Evaluación psicológica forense en víctimas de abuso sexual infantil en Latinoamérica* [Tesis de grado, Universidad Católica de Cuenca]. Repositorio Institucional de la Universidad Católica de Cuenca. https://dspace.ucacue.edu.ec/handle/ucacue/14215

Roberts, K. P., y Lamb, M. E. (2010). Reality-monitoring characteristics in confirmed and doubtful allegations of child sexual abuse. *Applied Cognitive Psychology, 24*(8), 1049–1079. https://doi.org/10.1002/acp.1600

Sporer, S. L., y Masip, J. (2025). A systematic review of the validity of Criteria-based Content Analysis in child sexual abuse cases and other field studies. *Psychology, Crime y Law, 31*(9), 1142–1183. https://doi.org/10.1080/1068316X.2024.2335971

Sporer, S. L., Hauch, V., Masip, J., y Martschuk, N. (2025). A meta-analysis of field studies on Criteria-Based Content Analysis. *European Psychologist, 30*(4), 256–269. https://doi.org/10.1027/1016-9040/a000561

Szojka, Z. A., Nicol, A., y La Rooy, D. (2020). Narrative coherence in multiple forensic interviews with child witnesses alleging physical and sexual abuse. Applied Cognitive Psychology, 34(5), 943–960. https://doi.org/10.1002/acp.3673

UNICEF. (2024, 10 de octubre). Violencia sexual: 370 millones de niñas y mujeres. https://www.unicef.es/noticia/violencia-sexual-370-millones-de-ninas-y-mujeres

VanMeter, F., Henderson, H. M., Konovalov, H., Karni-Visel, Y., y Blasbalg, U. (2023). Children's narrative coherence in "Achieving Best Evidence" forensic interviews and courtroom testimony. *Psychology, Crime y Law, 29*(2), 203–221. https://doi.org/10.1080/1068316X.2021.2018438

Welle, I., Berclaz, M., Lacasa, M.-J., y Niveau, G. (2016). A call to improve the validity of criterion-based content analysis (CBCA): Results from a field-based study including 60 children's statements of sexual abuse. *Journal of Forensic and Legal Medicine, 43*, 111–119. https://doi.org/10.1016/j.jflm.2016.08.001

Wouters, E., Constanty, L., Urben, S., Amoussou, J. R., y Gasser, J. (2024). Use of statement validity analysis in minors alleging sexual assault: A systematic review. *Journal of Forensic Sciences, 69*(6), 1948–1958. https://doi.org/10.1111/1556-4029.15604

Wouters, E., Constanty, L., Urben, S., y Gasser, J. (2025). Expert reports for credibility assessments in case of sexual abuse allegations in the Swiss judicial context: A retrospective study of the socio-demographic, clinical and forensic characteristics of the cases and the appropriate use of the Statement Validity Analysis. *Journal of Forensic and Legal Medicine, 114,* Article 102900. https://doi.org/10.1016/j.jflm.2025.102900

Yancce, R. (2020). Credibilidad testimonial del testigo en el proceso penal. *Revista Brasileira de Direito Processual Penal, 6*(1), 453–480. https://doi.org/10.22197/rbdpp.v6i1.246

Copyright (c) 2026. Fernández Yascaribay Katherine, Fernández Olivo Diego Esteban, Campoverde Antonio Xavier.

Este es un artículo de acceso abierto, bajo licencia Creative Commons (CC BY 4.0) de Atribución 4.0 Internacional