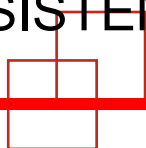




**Universidad
Europea** MADRID

FIABILIDAD DE LOS INSTRUMENTOS PARA EVALUAR LA CREDIBILIDAD DEL TESTIMONIO EN MENORES VÍCTIMAS DE ABUSO SEXUAL

REVISIÓN SISTEMÁTICA



Tutor/a: Marina Salas Asso

Coordinadora de TFM (cotutora): María Ángeles de la Cruz

Marta Gavilán Criado

Curso: 2024/2025

Índice

| | |
|--|----|
| Resumen, Abstract y palabras clave..... | 3 |
| 1. Introducción..... | 6 |
| 1.1. Justificación de la investigación | 12 |
| 1.2. Objetivo de la investigación..... | 12 |
| 2. Método..... | 14 |
| 2.1. Criterios de selección..... | 15 |
| 2.2. Búsqueda de artículos..... | 15 |
| 2.3. Proceso de selección..... | 16 |
| 2.4. Evaluación del riesgo de sesgo..... | 17 |
| 2.5. Procedimiento y análisis de datos..... | 17 |
| 3. Resultados..... | 19 |
| 3.1. Selección de estudios..... | 19 |
| 3.2. Características de los estudios..... | 20 |
| 3.3. Calidad metodológica..... | 24 |
| 3.4. Análisis de resultados..... | 26 |
| 4. Discusión..... | 32 |
| 5. Conclusiones..... | 36 |
| 6. Bibliografía..... | 37 |

Resumen:

Introducción: El Abuso Sexual Infantil (ASI) es un problema global que tiene graves consecuencias psicológicas, emocionales y sociales. En el ámbito jurídico-forense el testimonio del menor se convierte en la principal o única prueba en los procesos judiciales, lo que hace fundamental contar con metodologías, protocolos e instrumentos válidos y fiables que permitan evaluar correctamente la credibilidad del relato. Esta revisión sistemática pretende determinar la fiabilidad de los métodos utilizados y los factores que influyen en la percepción de credibilidad y en el relato de los menores.

Objetivo: Analizar la fiabilidad de los instrumentos empleados en la evaluación del testimonio de menores víctimas de abuso sexual.

Método: La revisión se ha realizado utilizando la metodología PRISMA y el Manual Cochrane, asegurando un proceso transparente y replicable. La búsqueda se efectuó en PudMed, Dialnet y Psycinfo. Tras una criba se seleccionaron 13 artículos y se evaluó el riesgo de sesgos con la herramienta del Joanna Briggs Institute (JBI).

Resultados: Los estudios muestran que las herramientas utilizadas en la evaluación del testimonio como CBCA, SVA, entrevistas forenses, CAS, FIT y CSAVC presentan múltiples limitaciones en fiabilidad, estandarización y aplicación. Además, se identifican factores externos que influyen en el relato del menor y en la percepción del evaluador o juez sobre la veracidad.

Discusión: La revisión refleja importantes desafíos en la evaluación de la credibilidad en casos de ASI. Aunque los métodos aportan información relevante, ninguno resulta fiable de forma aislada. Se han observado evaluaciones subjetivas y sesgadas, condicionadas por factores externos que limitan el relato y distorsionan la percepción del evaluador y jueces. Esto refleja la falta de consenso y estandarización y la necesidad de herramientas controladas, formación especializada y un enfoque multidimensional que integre herramientas objetivas y protección del menor.

Conclusión: La credibilidad del testimonio depende de criterios técnicos, profesionales especializados y factores externos que influyen en la evaluación y percepción del relato. Su fiabilidad requiere un enfoque multidimensional con entrevistas y herramientas estructuradas, medidas de protección y aplicación rigurosa de protocolos.

Palabras clave: Abuso sexual infantil (ASI), fiabilidad de las pruebas, testimonio, entrevista forense, niños.

Abstract:

Introduction: Child Sexual Abuse (CSA) is a global problem that has serious psychological, emotional, and social consequences. In the forensic-legal field, the child's testimony becomes the main or only evidence in judicial proceedings, which makes it essential to have valid and reliable methodologies, protocols, and instruments to correctly assess the credibility of the account. This systematic review aims to determine the reliability of the methods used and the factors that influence the perception of credibility and the child's narrative.

Objective: To analyze the reliability of the instruments used in the evaluation of the testimony of children victims of sexual abuse.

Method: The review was carried out using the PRISMA methodology and the Cochrane Handbook, ensuring a transparent and replicable process. The search was conducted in PubMed, Dialnet and Psycinfo. After screening, 13 articles were selected, and the risk of bias was assessed with the Joanna Briggs Institute (JBI) tool.

Results: The studies show that the tools used in the evaluation of testimony, such as CBCA, SVA, forensic interviews, CAS, FIT, and CSAVC, present multiple limitations in reliability, standardization, and application. In addition, external factors were identified that influence the child's narrative and the perception of the evaluator or judge regarding truthfulness.

Discussion: The review reflects major challenges in assessing credibility in CSA cases. Although the methods provide relevant information, none are reliable when applied alone. Subjective

and biased evaluations were observed, conditioned by external factors that limit the narrative and distort the perception of evaluators and judges. This reflects the lack of consensus and standardization and the need for controlled tools, specialized training, and a multidimensional approach that integrates objective tools and child protection.

Conclusion: The credibility of testimony depends on technical criteria, specialized professionals, and external factors that influence evaluation and perception. Its reliability requires a multidimensional approach with structured interviews and tools, protective measures, and rigorous application of protocols.

Keywords: Child sexual abuse (CSA), reliability of the tests, testimony, forensic interview, childrens.

1- Introducción

El abuso sexual infantil (ASI) se define de una manera breve y concisa como cualquier tipo de interacción sexual que se produzca entre un menor de 18 años y un adulto, de manera que se den una serie de comportamientos de índole sexual, como es el caso de: conductas de exposición, insinuaciones o incluso tocamientos, llegando hasta cualquier clase de práctica con una mayor implicación sexual. La búsqueda de excitación o satisfacción del adulto es lo que suele motivar el desarrollo de este tipo de conductas, en la cual es normal que se ejerza una posición de poder y autoridad sobre el menor, por parte del adulto.

Esta descripción no es definitiva ni común, ya que realmente no existe una definición que sea universalmente aceptada. No obstante, una de las acepciones más destacadas es la que fue formulada por el *National Center on Child Abuse and Neglect* (1978, citado en López, 2002), que define el ASI como “la utilización de un menor por parte de un adulto para su propia estimulación sexual o la de terceros”.

Hay que decir que este fenómeno no se limita a la interacción adulto-menor en todo momento, ya que también puede darse entre dos o más menores, siempre que exista una diferencia significativa de edad (lo cual significa unos cinco años o más) o también cuando se emplean o se utilicen formas de coacción como es el caso de la intimidación, la manipulación o el abuso de poder, ya que todas ellas restringen la capacidad de decisión de la víctima durante esa relación (Pineda et al.,2023).

El ASI no solo se relaciona como una dinámica de tipo injusto, donde factores como la edad, el poder y la autoridad están en juego, sino que también hace que tenga un gran impacto negativo en las puesto que dejan consecuencias profundas y efectos persistentes en sus vidas.

Se han establecido toda clase de distintas investigaciones, por parte de expertos, cuyo fin ha sido mostrar que el ASI puede generar cambios neurológicos muy importantes en la víctima. Es el caso de alteraciones en la amígdala, el hipocampo y la corteza cerebral. Todo ello puede llegar a afectar a áreas fundamentales que intervienen en la memoria, el aprendizaje y la regulación emocional.

En cuanto a las consecuencias psicológicas, el ASI se relaciona con el desarrollo de: depresión, ansiedad, estrés postraumático, autolesiones, disociación de la realidad, una baja autoestima, problemas para relacionarse con los demás y muchos más que pueden llegar a darse en el menor afectado.

Otro punto importante, son los efectos que el ASI tiene en la memoria de la víctima, donde como consecuencia de la traumática relación que ha tenido puede verse afectada por recuerdos intrusivos recurrentes con episodios de amnesia disociativa, donde el trauma queda parcialmente bloqueado, pero arrastrando la marca emocional. Losada (2012) comenta que “la intensidad en la que se manifiestan estas secuelas depende de factores como la edad de inicio, la duración del abuso, la relación con el agresor y el apoyo recibido tras la revelación”.

Además de todas estas consecuencias de tipo individual, una serie de estudios han demostrado, desde hace tiempo, que el ASI se suele desarrollar en contextos de riesgo a nivel familiar, social y ambiental. Los estudios destacan la prevalencia del ASI en: familias reconstruidas por parte del padrastro, familias monoparentales donde se delega el cuidado del menor a un tercero y también en contextos con dinámicas disfuncionales y con poco espacio de privacidad en el hogar.

La interacción de estos factores aumenta la vulnerabilidad del menor, lo que evidencia que se trata de un fenómeno complejo que necesita la implementación de intervenciones preventivas integrales (García, A., 2016).

A nivel social, se puede afirmar con claridad que se trata de un fenómeno en gran medida invisibilizado y que, de hecho, está sub-registrado, en el que el silencio, el miedo, las amenazas o la desconfianza institucional consiguen que sea mucho más enrevesado que se denuncie, por lo que se perpetúa la impunidad (Girón, 2018). El ASI trasciende mucho más allá del plano individual y judicial, puesto que se transforma en un problema de salud pública y de gran impacto para toda la sociedad y a todos los niveles.

El ASI, por tanto, se reconoce y se ve como un problema mundial, no solo por expertos psicólogos de carácter individual, sino también por organismos internacionales como la Organización Mundial de la Salud (OMS, 2002) y la Oficina de Naciones Unidas contra la Droga y el Delito (UNODC, 2010) considerándose una amenaza muy alta y grave, tanto para la salud pública como para los derechos de la infancia, ya que les afecta de una manera abismal.

Tal y como afirma WHO (2017) aproximadamente 1 de cada 5 mujeres y 1 de cada 13 hombres han sufrido algún tipo de ASI. Son cifras que realmente asombran y que reflejan una prevalencia elevada y transversal, que afecta a todos los contextos sociales y culturales, aunque con una marcada tendencia al subregistro y a la invisibilización de este gran problema.

En España, los estudios estiman que entre un 10% y un 20% de la población ha sufrido algún tipo de abuso sexual antes de los 18 años, aunque la mayoría de los casos no llega ni siquiera a denunciarse. “La mayoría de los abusos sexuales infantiles en España son cometidos por personas del entorno cercano de la víctima, suelen comenzar entre los 6 y 12 años y, con frecuencia, se prolongan en el tiempo, lo que evidencia la gravedad y persistencia del problema”, explica Delgado et al. (2015).

Todo ello convierte al ASI en un problema estructural que compromete tanto la salud individual como la cohesión social, y que requiere respuestas coordinadas desde la justicia, la salud y la protección de la infancia (Girón, 2018). A diferencia de otros delitos, el testimonio del menor, que ha sido víctima del ASI, es la prueba principal a la hora de verificar los hechos judicialmente, incluso muchas veces es la única prueba que está disponible, Alonso García (2022).

El testimonio del menor adquiere un papel central en el proceso penal, ya que en la mayoría de los casos los abusos se cometen sin que existan testigos y en un entorno privado e íntimo. Además, el testimonio infantil no puede ser evaluado con los mismos parámetros que el de un adulto, ya que los menores presentan limitaciones cognitivas y lingüísticas propias de la infancia que condicionan la forma en que los menores relatan los hechos (López, 2002).

También se suma el tiempo prolongado entre el hecho y la denuncia formal, debido muchas veces, a la dificultad de transmitir la experiencia traumática por miedo, vergüenza o amenazas, lo que lleva a dificultar la investigación de los hechos debido a la poca probabilidad de obtener pruebas físicas o biológicas que ayuden a corroborar el hecho que se ha producido (Ministerio Público de Chile, 2014; UNODC, 2010).

Aunque a lo largo de los años se han producido avances significativos en la protección del menor a nivel legislativo y social, siguen existiendo numerosos casos que no se llegan a denunciar o no son llevados a juicio debido a la falta de pruebas, miedo al entorno o a las dificultades para verificar la credibilidad del testimonio del menor, Vara (2021).

A pesar de los avances en cuestiones metodológicas, la credibilidad del testimonio es una de las áreas más complejas dentro del ámbito clínico-forense, sobre todo debido a una gran falta de procedimientos adecuados y a la escasa credibilidad que se le suele otorgar al testimonio infantil, aumentando, por tanto, la vulnerabilidad de los menores en los procesos judiciales Vara (2021).

En la práctica, no existe un consenso concreto en cuanto a la metodología ni instrumento estandarizado a utilizar en la valoración de la credibilidad del testimonio infantil. Esto solamente conlleva a que los informes periciales puedan ser interpretados de manera subjetiva en función de la experiencia y el criterio de cada evaluador. Esta ausencia de estandarización es la que compromete la objetividad de las conclusiones, e incrementa la inseguridad jurídica al influir directamente en la decisión judicial (Posada & Cañón, 2022).

En este sentido, Vara (2021) señala que “diversos factores intervienen en la valoración de la credibilidad, destacando desde las limitaciones cognitivas y emocionales propias de la infancia hasta la forma en que se formulan las preguntas, la repetición excesiva de entrevistas y la falta de protocolos adaptados a esta población”.

Estas circunstancias también pueden afectar, por parte de los jueces, a la percepción de su veracidad, siempre teniendo en cuenta el testimonio del menor, puesto que no siempre cuentan con los

conocimientos técnicos que son necesarios para interpretar de manera correcta y adecuada los informes periciales, lo cual incrementa el riesgo de cometer ciertos errores judiciales, como pueden ser, sobre todo, condenar inocentes o absolver culpables (Ramírez Rojas, 2018).

“Todas estas particularidades incrementan la complejidad de la valoración psicológica infantil y refuerzan la necesidad de protocolos y peritos especializados, capaces de adaptar los métodos de entrevista y análisis a las capacidades evolutivas del menor” expresa López, (2002).

El uso incorrecto de dichos procedimientos, así como la falta de formación del evaluador o la inadecuación de los instrumentos y protocolos, puede no solo comprometer la calidad del testimonio, sino también afectar negativamente su estado emocional (Vara, 2021).

Con el fin de garantizar la protección de la declaración del menor y, al mismo tiempo, la validez procesal de su testimonio ha sido reconocida la llamada prueba preconstituida, de tal manera que se toma como una medida procesal fundamental en los casos de ASI.

Su importancia, sobre todo, radica en que garantiza que el testimonio del menor pueda ser recogido de una forma segura tan solo con una declaración, por lo que se evitan confusas repeticiones que puedan afectar a la memoria o generar victimización secundaria, de manera que se preserva y se mantiene tanto la fiabilidad del relato como el bienestar de los menores, en todo momento (Sotoca et al., 2012).

En este marco, la evaluación psicológica de la credibilidad se define y se entiende como una parte sumamente valiosa y esencial del procedimiento, ya que cumple con una doble función: uno, probatoria, ya que valora la veracidad del relato en ausencia de pruebas físicas; y dos, protectora, al prevenir entrevistas inadecuadas o reiteradas.

Tanto la prueba preconstituida como el análisis de credibilidad se enfocan en un mismo objetivo que es ni más ni menos que asegurar la calidad del procedimiento y, por otro lado, la protección de la infancia mediante el uso de unos protocolos claros y profesionales, de tipo especializado (González, 2016). En este sentido los peritajes psicológicos cumplen un papel de gran

valor, sobre todo cuando no existen pruebas físicas ni testigos directos. Su importancia se fundamenta en utilizar métodos rigurosos, actualizados y multidimensionales que incluyan entrevistas especializadas, observación conductual y pruebas validadas, evitando depender de una sola técnica.

Entre los protocolos de entrevista estructurada destaca el modelo NICHD, considerado un referente internacional por su eficacia en la obtención del relato del menor mediante preguntas abiertas y neutrales. Su correcta aplicación, a cargo de profesionales entrenados, mejora la calidad y validez del testimonio al reducir la sugestión y reforzar su fiabilidad (Gutiérrez Morante, 2021).

Técnicas como las proyectivas gráficas, pese a haber sido cuestionadas por su subjetividad y falta de estandarización, pueden cumplir un papel complementario, ofreciendo una alternativa menos invasiva de expresión y reduciendo el riesgo de revictimización (Pinto, 2020).

De igual manera, metodologías como el *Statement Validity Assessment* (SVA) y el *Criteria-Based Content Analysis* (CBCA) aportan criterios objetivos para diferenciar entre testimonios infantiles verdaderos y falsos; sin embargo, hay que tener en cuenta que no ofrecen certezas absolutas, y que su validez depende tanto de la calidad de la entrevista como de la pericia del evaluador en cada caso.

Todas estas herramientas deben aplicarse con suma cautela, siempre dentro de un modelo integral de evaluación y en todo momento con una constante actualización científica de los conocimientos, ya que solo así se logrará reforzar la credibilidad del testimonio infantil, en el ámbito judicial (Alonso García, 2022; Vara, 2021).

En definitiva, el ASI es un problema de salud pública y social que, debido a la ausencia de pruebas físicas, convierte al testimonio del menor en la base central de la investigación. La fiabilidad de ese testimonio del menor depende totalmente de protocolos estandarizados y de la capacitación de los profesionales, siendo los dos elementos clave para garantizar evaluaciones objetivas de los casos y la protección del interés superior del niño.

1.1 Justificación

A raíz de esta problemática cada vez más visible, la evaluación del testimonio en casos de ASI se ha convertido en uno de los campos de mayor interés en el ámbito profesional y científico, especialmente en los entornos clínico y forense. En consecuencia, se están realizando estudios e investigaciones para poder evaluar el testimonio de las víctimas de forma que se apliquen los instrumentos adecuados para poder obtener mayor certeza y fiabilidad.

En este contexto, la siguiente investigación se justifica en la necesidad de determinar, a través de una revisión sistemática, la fiabilidad de los instrumentos utilizados en la evaluación del testimonio de menores víctimas de ASI, así como identificar los factores externos que influyen tanto en el relato del menor como en la percepción de credibilidad del testimonio.

1.2 Objetivo de la investigación

Esta revisión parte de una pregunta de investigación centrada en la credibilidad del testimonio del menor y la fiabilidad de las herramientas empleadas para su evaluación, con el propósito de identificar qué instrumentos resultan más adecuados según su etapa evolutiva.

Habiendo descrito el problema anteriormente, la cuestión que se nos plantea es la siguiente: ¿Cuál es la fiabilidad y validez de los instrumentos habitualmente utilizados en la evaluación del testimonio de menores víctimas de abuso sexual, teniendo en cuenta su aplicación en el ámbito clínico y forense?

| Elementos | Definición |
|-------------------------|---|
| P (Población) | Menores víctimas de ASI |
| I (Intervención) | Análisis del uso de instrumentos y protocolos de evaluación de credibilidad más utilizados en ASI. |
| C (Comparación) | - Aplicación aislada de un único instrumento. - Métodos no estandarizados o con baja validación. |
| O (Resultados) | Nivel de fiabilidad y validez, limitaciones de estandarización e influencia de factores externos en el relato del menor y en la percepción de credibilidad. |

Para identificar y desarrollar los objetivos específicos se plantean las siguientes preguntas:

1. ¿Existe evidencia acerca de la validez y fiabilidad de los instrumentos más utilizados en la evaluación del testimonio en menores víctimas de abuso sexual?
2. ¿Se aplican correctamente estos protocolos e instrumentos en el ámbito clínico y forense?
3. ¿Qué factores afectan negativamente a la evaluación del testimonio del menor en casos de abuso sexual?

Objetivo general

Analizar la fiabilidad de los instrumentos empleados habitualmente en la evaluación del testimonio de menores víctimas de abuso sexual.

Objetivos específicos

- Investigar la validez y fiabilidad de los instrumentos más utilizados en la evaluación del testimonio en menores víctimas de ASI.
- Analizar la aplicación de los protocolos e instrumentos empleados en la evaluación del testimonio del menor víctima de ASI.
- Examinar los factores que afectan negativamente en la evaluación del testimonio del menor en casos de abuso sexual, dentro del ámbito clínico y forense.

2- Método

Esta revisión sistemática se ha llevado a cabo a través del protocolo PRISMA 2020 (Page et al., 2021) y las recomendaciones del Manual Cochrane (Higgins & Green, 2011), con el objetivo de que el proceso fuera riguroso, transparente y replicable, proporcionando una explicación más detallada sobre los aspectos clave de la metodología y la revisión sistemática.

Para comenzar, se formularon diferentes preguntas relacionadas con el tema de investigación con el fin de definir de forma concreta el objeto de estudio.

Tras marcar los objetivos para la revisión sistemática, se definieron diferentes criterios de inclusión y exclusión, con el fin de reducir la búsqueda y obtener resultados que estuvieran correctamente centrados en la credibilidad del testimonio en ASI.

Una vez elegidos los criterios (de inclusión y exclusión) y para que la búsqueda fuera aún más específica se utilizaron palabras clave en español e inglés (Abuso sexual infantil (ASI), fiabilidad de las pruebas, testimonio, entrevista forense, niños) combinadas con operadores booleanos como AND y OR, con el fin de obtener artículos más incluyentes para la investigación.

Cuando se llevaron a cabo las diferentes metodológicas utilizadas para la selección y análisis de los estudios incluidos, se realizó una revisión inicial de la literatura encontrada a través de la lectura de títulos y resúmenes de los artículos, seleccionando aquellos que tuvieran mayor relevancia con el objeto de estudio y que cumplieran con los criterios de inclusión establecidos y descartando los restantes.

Una vez seleccionados, se realiza un análisis más detallado de cada artículo, destacando la población de estudio, la metodología utilizada y los resultados del estudio con el fin de extraer de forma detallada los datos más relevantes de cada uno de ellos.

Para finalizar este proceso, se examinaron los resultados obtenidos a través de un análisis de tipo descriptivo y se establecieron las conclusiones principales obtenidas a través de los diferentes estudios.

2.1. Criterios de inclusión y exclusión

Para la siguiente revisión sistemática se tienen en cuenta a la hora de seleccionar los artículos los siguientes criterios de inclusión y exclusión con el fin de acotar la búsqueda y seleccionar artículos concretos relacionados con el objeto de estudio, en este caso la validez de la credibilidad del testimonio en niños víctimas de ASI.

Tabla 1:

Criterios de selección

| Categoría | Criterios de Inclusión | Criterios de Exclusión |
|-----------------------------|--|---|
| Idioma | Estudios publicados a texto completo en cualquier idioma. | |
| Fecha de publicación | Publicaciones con una antigüedad máxima de 10 años (a partir de 2015). | Publicaciones anteriores a 2015. |
| Tipo de publicación | Estudios empíricos y artículos de revistas científicas. | Estudios cualitativos, artículos teóricos, revisiones narrativas y revisiones sistemáticas. |
| Población | Muestras de menores de entre 3 y 18 años víctimas o presuntas víctimas de abuso sexual. | Estudios que no incluyan población de menores víctimas de abuso sexual. |
| Contenido/objetivo | Investigaciones cuantitativas que evalúen la fiabilidad y/o validez de instrumentos utilizados en la recogida del testimonio infantil en contextos jurídicos o clínico-forenses. | Investigaciones centradas en instrumentos no aplicables al ámbito jurídico o forense, o que no estén orientados a la evaluación del testimonio. |

2.2 Búsqueda de artículos

La búsqueda bibliográfica de los artículos se ha llevado a cabo en las bases de datos PudMed, Dialnet y Psycinfo durante los meses de mayo a agosto.

Una vez elegidas las bases de datos, se han realizado estrategias de búsqueda a través de ecuaciones de palabras combinadas que han sido adaptadas a las características específicas de cada base de datos. Para ello, se han seleccionado términos controlados (palabras estandarizadas)

combinados con términos libres, para ampliar el alcance de la revisión e incluir aquellos artículos relevantes que no se encuentren clasificados bajo estos términos.

La estrategia de búsqueda se ha llevado a cabo en todos los idiomas, y se han utilizado operadores booleanos como AND y OR para combinar los conceptos clave de forma precisa en las ecuaciones planteadas. Las palabras clave utilizadas en español fueron: abuso sexual infantil, testimonio, fiabilidad de las pruebas, entrevistas y niños. Las palabras clave correspondientes en inglés fueron: child sexual abuse, testimony, reliability of the instruments, interviews y childrens.

Para organizar y clasificar los artículos finales de la revisión, se ha utilizado el software Zotero, que permite identificar y eliminar artículos duplicados.

Tabla 2:

Palabras clave

| Base de datos | Cadena de búsqueda |
|-----------------|---|
| PudMed | Child sexual abuse AND (Testimony OR Reliability of the instruments OR Childrens OR Interviews). |
| Dialnet | Abuso sexual infantil AND (testimonio OR/ AND entrevista) Abuso sexual infantil AND fiabilidad de las pruebas AND testimonio |
| Psycinfo | Abuso sexual infantil AND Fiabilidad de las pruebas AND/OR testimonio Abuso sexual infantil AND entrevista AND testimonio |

2.3 Proceso de selección de estudios

Para la elección de los estudios, se han seguido las indicaciones del protocolo PRISMA a través de varias fases.

Primero, se llevó a cabo una búsqueda exhaustiva con las palabras clave seleccionadas, obteniendo un total de n= 10.321 (Pubmed: 8.116, Psycinfo: 1.405, Dialnet: 800).

Luego, se aplicaron los filtros y se revisaron los títulos y resúmenes de los artículos, con el objetivo de descartar los estudios que no cumplen los criterios de inclusión obteniendo un total de n= 1092. (Pubmed: 790, Psycinfo: 262, Dialnet: 40).

Tras la primera criba, los artículos seleccionados fueron leídos y examinados de forma concreta, y finalmente tras un análisis se incluyeron 13 n= 13 (Pubmed: 6, Psycinfo: 4, Dialnet: 3).

Finalmente, los artículos seleccionados en la revisión han sido registrados y representados en un diagrama de flujo PRISMA, reflejando detalladamente las fases de exclusión y los motivos correspondientes basados en los criterios de inclusión y exclusión, así como los artículos finalmente seleccionados.

2.4 Evaluación del riesgo de sesgo

Para determinar la calidad y validez de los estudios seleccionados, se realizó la evaluación de sesgos utilizando la herramienta de evaluación “Joanna Briggs Institute” (JBI) ya que proporciona un conjunto de 15 listas de verificación según el tipo de artículo, que nos permite evaluar la confiabilidad, relevancia y resultados de los artículos seleccionados para la revisión.

Concretamente, en esta revisión se han utilizado 3 de las 15 listas de verificación:

- Lista de verificación JBI para la evaluación crítica de estudios transversales analíticos, empleada en los estudios observacionales transversales.

- La lista de verificación crítica de JBI para estudios de precisión de pruebas diagnósticas, que se emplea en los proyectos de creación y validación de instrumentos (como FIT, CSAVC o CBCA).

- JBI Critical Appraisal Checklist for Case Series, empleada en la investigación que examinaba casos judiciales en Colombia.

El empleo de estas listas permitió que la evaluación se adaptara a cada estudio, garantizando un análisis detallado y conforme al diseño metodológico de los artículos que se incluían.

2.5 Procedimiento y análisis de datos

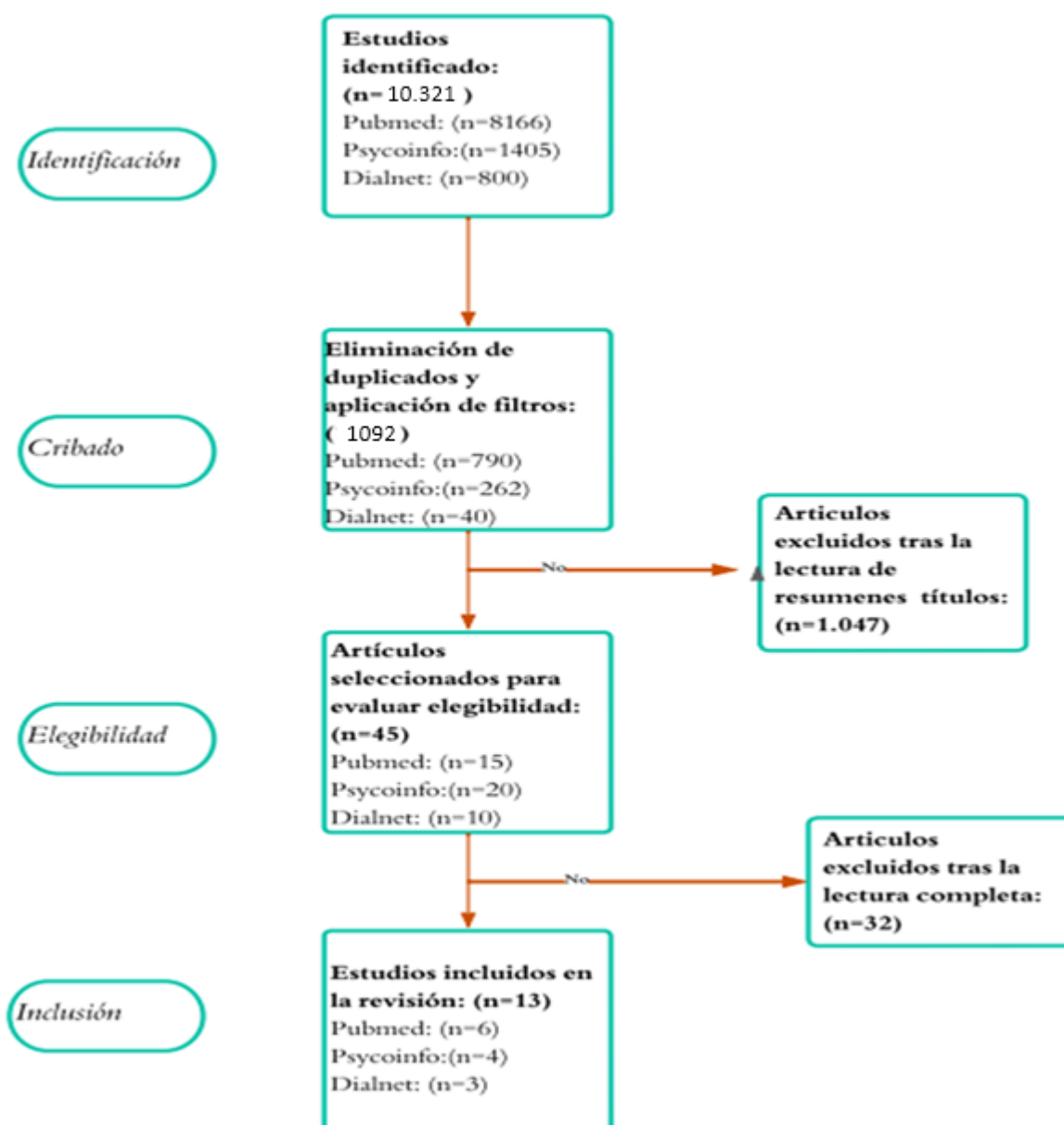
Una vez seleccionados los estudios de la revisión, se ha realizado una extracción de los datos más relevantes de cada uno de los estudios, mediante una tabla *ad ho*. En ella se han recogido aspectos como los autores, año de publicación, país, tipo de instrumento evaluado, dimensiones analizadas y los resultados principales.

Después, se ha realizado un análisis comparativo de los estudios incluidos, con el fin de determinar similitudes y diferencias en los protocolos e instrumentos utilizados, así como en su aplicación, identificando factores que influyen en la veracidad del testimonio. El objetivo era identificar patrones, fortalezas y limitaciones metodológicas en la evaluación de la credibilidad del testimonio infantil en casos de abuso sexual.

3- Resultados

3.1 Selección de estudios

En el siguiente diagrama de flujo se refleja las bases de datos utilizadas para realizar la búsqueda. Además, permite conocer el proceso de selección realizado para determinar los artículos primarios de la revisión sistemática. Este diagrama refleja el número de artículos encontrados en cada base de datos, el cribado utilizado para la selección de artículos y la exclusión e inclusión de los artículos finales.



3.2 Características de los estudios

Tras la búsqueda y análisis de los resultados encontrados en las bases de datos se seleccionan un total de 13 estudios primarios, que se incluyen en la revisión sistemática con el fin de estudiar la fiabilidad y validez de las herramientas, pruebas y protocolos usados para evaluar la credibilidad del testimonio de niños y niñas víctimas de ASI.

En los estudios, la población objetivo, está formado por menores de entre 3 y 18 años víctimas de ASI que han participado en procedimientos judiciales reales o simulados. Además, los estudios e investigaciones seleccionadas se han realizado en diferentes países como Israel, Portugal, Francia, Suecia, Países Bajos o Estados Unidos, lo que ayuda a proporcionar una visión más internacional acerca de las prácticas forenses utilizadas en la credibilidad del testimonio del menor.

De forma genérica, los trabajos que han sido investigados y seleccionados analizan la aplicación de protocolos estructurados como el NICHD, el uso de herramientas psicométricas, el análisis basado en criterios como el CBCA o cómo pueden influir las variables procesales (tipo de pregunta, consistencia del relato, rol del menor) en la percepción de credibilidad (Niveau et al., 2015; Ernberg et al., 2018; Voogt et al., 2017).

Tabla 3:

Artículos

| Autores y año | País | Tipo de investigación | Objetivo | Muestra | Resultados y Conclusiones |
|----------------------|-------------|------------------------------|---|---|--|
| Niveau et al., 2015 | Suiza | Observacional cuantitativa | Evaluar la fiabilidad entre interevaluadores cuando se utilizaba el método CBCA en la evaluación del testimonio del menor | 95 declaraciones de niños de 3 a 17 años (edad media: 12,7 años). La muestra incluyó 17 niños varones y 78 niñas. | El estudio determina que la fiabilidad interevaluadores es moderada: algunos criterios muestran alta consistencia, pero otra baja concordancia, lo que refleja interpretaciones subjetivas que pueden afectar la evaluación. En consecuencia, el CBCA no resulta totalmente fiable, ya que la validez de la declaración depende en parte del juicio del evaluador. |

| | | | | | |
|--------------------------|----------|---|--|--|---|
| Ruiz-Tejero et al., 2016 | España | Estudio analítico con un diseño ex post facto | Analizar la estructura factorial y consistencia interna de un protocolo clínico-pericial (CAS). | N= 50 peritajes psicológicos y entrevistas forenses de casos abiertos Edad= 2 y 16 años. Edad media: 6,8 años; 86% mujeres. | El estudio refleja que el protocolo de evaluación (CAS) tiene dos factores generales: variables criminológicas y socio-familiares, y otro con el abuso y el relato del menor. Tras el análisis se encontró una consistencia interna moderada para los dos factores, concluyendo que el protocolo presenta una alta fiabilidad en la detección de falsas alegaciones, pero se necesitan investigaciones con muestras más amplias para mejorar su validez discriminante. |
| Welle et al., 2018 | Suecia | Estudio de campo | Evalúa la fiabilidad y valides del CBCA | Muestra inicial: 225 testimonios. Finalmente 60 declaraciones de niños de 3,5 a 17,5 años. 76% eran mujeres (n = 46). Rango de edad promedio: 13 años. | La investigación refleja que la fiabilidad interevaluadores del CBCA es elevada, sobre todo en las puntuaciones totales, aunque algunos criterios, como "Detalles inusuales" y "Detalles característicos del abuso", mostraron baja fiabilidad. Las puntuaciones totales fueron más altas en las alegaciones confirmadas que en las no confirmadas, y la edad de los niños influyó en los puntajes del CBCA, sobre todo en los casos no confirmados. En conclusión, el CBCA es fiable, pero se necesitan más estudios que proporcionen mayor consistencia y realizar modificaciones en algunos criterios. |
| Peixoto et al. (2017) | Portugal | Estudio de campo | Analizar las entrevistas judiciales a menores víctimas de abuso sexual o físico con el fin de medir con precisión la calidad y cantidad de los detalles proporcionados por los niños en sus testimonios. | 137 entrevistas, de las cuales 60 fueron analizadas completamente. Edad: 3–17 años. Edad media: 11 años, 5 meses Género: 85 mujeres, 48 hombres | Tras el análisis de las entrevistas judiciales, se determina un uso excesivo de preguntas cerradas y sugestivas, escasa utilización de abiertas. Las preguntas con opción fueron las más utilizadas, seguidas de preguntas muy directivas que reflejaban en las menores respuestas responsivas. Lo que se concluye limita la oportunidad de los niños para proporcionar detalles completos proponiendo el protocolo NICH como alternativa para un relato más completo y detallado |

| | | | | | |
|-----------------------------|----------|--|--|--|--|
| Rodríguez Cely et al., 2018 | Colombia | Análisis de contenido basado en el uso de tablas de revisión de categorías | Evaluar la calidad de los informes de evaluación psicológica forense en casos de ASI. | 53 expedientes de menores víctimas de ASI. | Los resultados muestran uso limitado de protocolos en Colombia, deficiencias en evidencia demostrativa y necesidad de fortalecer prácticas forenses. |
| Ernberg et al., 2018 | Suecia | Análisis de contenido cualitativo y cuantitativo | Analizar como los tribunales evalúan la credibilidad del testimonio del menor | 70 sentencias judiciales de casos reales. Edad: 3 a 7 años. 88% niñas, 12% niños | Las conclusiones y resultados del estudio reflejan que la credibilidad del menor se ve influenciada no solo por las herramientas utilizadas, también por variables subjetivas y psicosociales, lo que refleja una carencia de criterios basados en evidencia científica. |
| Karzt et al., 2019 | Israel | Observacional, transversal | Estudiar las estrategias de confrontación utilizadas en las entrevistas forenses basadas en el protocolo NICH a menores víctimas de ASI | Muestra: 224 transcripciones de las entrevistas Edad: 4 a 14 años | Los resultados del estudio muestran que el 61,2 % de las entrevistas utilizaban estrategias de confrontación, lo que pueden generar en el menor miedo, confusión e incluso retracciones, llegando a comprometer el relato del menor. Además, señala que hay variables como la edad, e tipo de abuso y la relación con el agresor que también incrementan las dificultades del relato del menor |
| Karni-Visel et al., 2019 | Israel | Cuasi-experimental comparativo | Analiza y compara dos protocolos de entrevista (Standard NICHD Protocol – SP vs Revised Protocol con el fin de evaluar el efecto del tipo de protocolo de entrevista en la expresión emocional y la cantidad de información proporcionada por niños víctimas de abuso. | Muestra: 278 entrevistas Edad: 4 a 14 años | Los resultados del estudio reflejan que cuando se aplica el factor emocional en las entrevistas, mejora su calidad permitiendo que el menor se exprese de forma más libre y detallada. Se concluye que el uso adecuado del protocolo NICH mejora la evaluación de la credibilidad del testimonio. |
| Denne et al., 2021 | EEUU | Estudio experimental | Analizar si la presencia de un testimonio experto ayuda a la credibilidad del menor frente a una retracción. | Muestra: 260 adultos Edad: Mayores de 18 años. | Los resultados reflejaron que los testigos expertos, sustentados en pruebas sólidas o en protocolos detallados, influían en la veracidad del testimonio. La forma del testimonio y el sexo de los jurados |

| | | | | | |
|----------------------------|-------------|-----------------------------------|---|--|---|
| | | | | | incidieron en la seguridad al dar el veredicto, la asignación de la responsabilidad y la ponderación de las pruebas. |
| Voogt et al., 2021 | Australia | Instrumental psicométrico | Diseñar y validar (CSAVCS), con el fin de medir de forma concreta la percepción de la credibilidad de los niños víctimas de abuso sexual a través de. | Muestra: 1Fase: 42 adultos (6 hombres y 36 mujeres) 2Fase: 210 adultos (61 hombre y 141 mujeres) | Los resultados obtenidos tras el análisis indican que el CSAVCS presenta una estructura sólida, reflejando buena fiabilidad y validez factorial. Se concluye por tanto que puede ser una herramienta útil para medir siendo adecuado para medir la percepción de la credibilidad de los niños víctimas de abuso sexual |
| Farrugia et al. (2018) | Reino Unido | Instrumental | Desarrollar y proponer una herramienta (FIT) para registrar y evaluar entrevistas forenses de manera integra | No se reporta una muestra cerrada de entrevistas; se menciona pilotaje en organizaciones | Los resultados concluyen que, aunque FIT se encuentra en una fase inicial y requiere pruebas empíricas y validación, se espera que el FIT ayude a realizar evaluaciones efectivas de entrevistas forense |
| Cabeza-García et al., 2022 | España | Cuasi-experimental | Investigar acerca de los factores psicosociales asociados a la credibilidad del testimonio en ASI | Muestra: 99 casos Edad: 4 a 17 años 85.9% eran niñas. | Los testimonios creíbles se caracterizaron tardar más en contar los hechos, se daban de manera más natural, mantenían una línea coherente y añadían datos secundarios., mientras que los no creíbles mostraron inconsistencias, influencia de terceros y más retracciones. Los factores psicosociales permiten diferenciar la credibilidad y complementar la evaluación pericial. |
| Szojka et al., 2025 | Australia | Diseño cuantitativo experimental. | Diseñar de un modelo automatizado para la clasificación y codificación de preguntas en entrevistas para la evaluación del testimonio | Muestra: entrevistas forenses (n = 1.435) y testimonios en juicio (n = 416) Edad: 3 y 17 años | Los resultados reflejan que el modelo presenta una alta fiabilidad 98% de acuerdo; $\kappa = .97$), evidenciando que la codificación automatizada del tipo de preguntas se puede considerar una herramienta prometedora debido a su alta precisión y eficiencia frente a la codificación manual |

3.3 Calidad metodológica

Después de aplicar las herramientas JBI a los estudios que forman parte de esta revisión, se puede determinar que la mayoría de ellos tiene un bajo riesgo de sesgo y cumple adecuadamente con los criterios principales de calidad metodológica.

En cuanto a los estudios observacionales transversales, todos (Peixoto et al., 2017; Karni-Visel et al., 2019; Katz et al., 2019; Niveau et al., 2015; Ernberg et al., 2018; Denne et al., 2021; Cabezas-García et al., 2022; Ruiz-Tejedor et al., 2016) incluyeron una definición precisa de los criterios de inclusión, las muestras y el contexto, y utilizaron medidas confiables y válidas.

No obstante, en múltiples casos se detectaron restricciones en los criterios 5 y 6 de la lista de verificación (las cuales son la identificación y el control de factores de confusión). En muchas situaciones se hallaron restricciones en los criterios 5 y 6 de la lista de verificación (identificación y control de factores confusos). Esto sitúa su riesgo de sesgo en un nivel medio-bajo, sin embargo, no compromete la validez general de los resultados.

Por otra parte, en los estudios de validación psicométrica y de nuevas herramientas (Farrugia et al., 2018; Szojka et al., 2025; Voogt et al., 2017; Welle et al., 2016), se determinó que la calidad metodológica en relación con el diseño y el análisis estadístico era alta. Szojka et al. (2025) y Farrugia et al. (2018), relejaron un riesgo bajo de sesgo. El estudio de Voogt et al. (2017) mostró ciertas limitaciones en la aleatoriedad; en cambio, el estudio de Welle et al. (2016), a pesar de ser metodológicamente sólido, empleó un diseño caso-control, lo cual representa un peligro para la validez externa y aumenta el riesgo de sesgo a un nivel moderado.

En relación con los estudios de series de casos, se incorporó el trabajo de Rodríguez Cely et al. (2018) con cautela. A pesar de que satisface la identificación válida de los casos y utiliza métodos correctos para medir la condición (procedimientos judiciales documentados), enfrenta restricciones en términos de la claridad de los criterios de inclusión, la completud de la muestra y la presentación de datos clínicos y demográficos. Estos elementos disminuyen la transparencia del estudio y aumentan el

riesgo de sesgo a un nivel medio, pese a que los resultados añaden valor contextual al examen de protocolos en el sistema judicial de Colombia.

También se descartaron dos trabajos de la evaluación con JBI porque no eran investigaciones empíricas primarias. Ambos eran de naturaleza teórica y crítica, sin ofrecer información original proveniente de entrevistas o muestras, así que no era apropiado utilizar las listas de verificación creadas específicamente para investigaciones empíricas.

En conclusión, la mayor parte de las investigaciones que componen esta revisión tienen un riesgo bajo o moderado de sesgo. Sus limitaciones más destacadas son el manejo deficiente de variables confusas en los estudios transversales y las limitaciones en el diseño experimental de ciertos trabajos de validación. Estos puntos deben tenerse en cuenta al analizar los resultados generales de la revisión, pues podrían afectar parcialmente la coherencia de la evidencia sintetizada.

Tabla 4

JBI – Checklist Analithical Cross Sectional Studies

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|------------------------------|---|---|---|---|---|---|---|---|
| Peixoto et al. (2017) | | | | | | | | |
| Karni-Visel et al. (2019) | | | | | | | | |
| Katz et al. (2019) | | | | | | | | |
| Niveau et al. (2015) | | | | | | | | |
| Ernberg et al. (2018) | | | | | | | | |
| Denne et al. (2021) | | | | | | | | |
| Cabezas-García et al. (2022) | | | | | | | | |
| Ruiz-Tejedor et al. (2016) | | | | | | | | |

Si Limitado No No aplicable

Tabla 5

JBI – “JBI Critical Appraisal Checklist for Diagnostic Test Accuracy Studies”

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-----------------------|---|---|---|---|---|---|---|---|---|----|
| Welle et al. (2016) | | | | | | | | | | |
| Farrugia et al., 2018 | | | | | | | | | | |
| Voogt et al. (2021) | | | | | | | | | | |
| Szójka et al. (2025) | | | | | | | | | | |

Si Limitado No No aplicable

Tabla 6
JB1 (Case Series)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|------------------------|---|---|---|---|---|---|---|---|---|----|
| Rodriguez et al., 2018 | | | | | | | | | | |

Si
 Limitado
 No
 No aplicable

3.4 Análisis de los resultados

Los estudios que se han revisado nos muestran diferentes factores, herramientas y procedimientos involucrados en la evaluación de la credibilidad del testimonio en caso de ASI, con el fin de poder determinar aspectos importantes que puedan influir en la valoración del testimonio. A continuación, se reflejan los resultados obtenidos tras el análisis descriptivo realizado.

Uso de protocolos, herramientas forenses y calidad de las entrevistas

Diversos estudios recientes han determinado que la entrevista forense presenta deficiencias en su aplicación, siendo considerada la herramienta más empleada para la evaluación de la credibilidad del testimonio del menor.

Autores como Peixoto et al. (2017) observaron en su estudio, un patrón similar al analizar las entrevistas que se habían realizado, identificando mayoritariamente el uso de preguntas cerradas y sugestivas, con pocas preguntas abiertas, lo que limita la cantidad y calidad de la información obtenida. Estos autores proponen la implementación del protocolo NICHD para favorecer el recuerdo libre y proteger mejor los derechos del menor.

En la misma línea, el estudio de Karni- Visel et al. (2019), quiso determinar cómo la expresión de las emociones en las entrevistas forenses que seguían el protocolo NICHD, influía en la calidad del testimonio de menores víctimas de ASI. Los autores estudiaron esto, realizando una comparación de las entrevistas en las que los entrevistadores invitaban de forma directa a los menores a expresar cómo se sentían durante los hechos y aquellas que no lo hacían.

Los resultados de tales entrevistas reflejan que añadir preguntas emocionales a las entrevistas forenses que utilizan el protocolo NICHD, ayudó a los menores a relatar un testimonio más espontáneo, detallado y rico en contenido. Por todo ello, los autores concluyeron que utilizar la expresión

emocional de forma controlada en las preguntas, puede aumentar la calidad y cantidad del testimonio. Además, se mostró que incitar a los menores a que revelen como se sintieron durante el abuso sufrido, puede incrementar la percepción de credibilidad del testimonio en los procedimientos judiciales.

En el mismo año, Katz et al. (2019) realizaron un estudio muy complejo y profundo con el fin de ver las estrategias de confrontación usadas en entrevistas forenses reales donde se aplicaba el protocolo NICHD a niños y niñas presuntamente víctimas de abuso físico o sexual y su efecto en la fiabilidad del testimonio infantil. En los resultados de este estudio, se pudo observar cómo, en el 61,2 % de las entrevistas se emplearon estrategias de contrainterrogatorio, por lo que se determina que estas prácticas pueden generar miedo, confusión e incluso retracciones, llegando a comprometer directamente la fiabilidad del testimonio.

Por tanto, el estudio demuestra que la fiabilidad del testimonio infantil puede verse comprometida cuando se emplean estrategias de contrainterrogatorio, y que variables como la edad, el tipo de abuso o la cercanía con el agresor pueden influir en la capacidad del menor para mantener su relato. Concluyen que el protocolo NICHD, ayuda a mejorar la estructura, la calidad y la obtención de información relevante pero que su efectividad puede verse afectada si se introducen técnicas de confrontación directa, ya que interfieren en el relato libre y fiable que proporciona dicho protocolo

Por otro lado, Rodríguez et., al (2018), concluyeron que, aunque la entrevista era la herramienta más utilizada, presentaban problemas éticos y metodológicos como: ausencia de consentimiento informado, hipótesis mal formuladas, escasa replicabilidad y falta de integración de múltiples fuentes. Además, destacan que en la aplicación de protocolos como el SATAC, donde las entrevistas realizadas fueron básicas sin análisis sistemático del discurso, y técnicas como SVA y CBCA se aplicaron de forma poco rigurosa, afectando su validez como prueba.

Farrugia et al., (2018) desarrollaron y evaluaron en su estudio el Forensic Interview Trace (FIT), con el fin de que analizara la calidad de entrevistas forenses con menores víctimas de abuso y proporcionar un método sistemático y fiable para valorar las entrevistas. Una vez diseñada la

herramienta FIT se aplicó a las entrevistas realizadas en el contexto judicial, ayudando a identificar de forma precisa fortalezas y debilidades de las entrevistas. Los resultados mostraron que muchas entrevistas no llevaban a cabo las buenas prácticas recomendadas internacionalmente, resaltando el uso insuficiente de preguntas abiertas o la tendencia a preguntas sugestivas. Los autores tras los resultados concluyen, que la herramienta FIT es práctica y fiable para evaluar la calidad de las entrevistas forenses, ya que detecta malas praxis, ofreciendo datos objetivos, claros y específicos.

En la misma línea Szojka et al. (2025) desarrollaron en Australia un sistema automatizado para codificar el tipo de preguntas que se hacía, demostrando que los resultados obtenidos podían compararse con la codificación humana, lo que permite el control de calidad de las entrevistas. Además, estos autores detectaron una alta frecuencia en el uso de preguntas cerradas o sugestivas, lo que puede generar distorsiones en el recuerdo y afectar negativamente la credibilidad.

Evaluación de la fiabilidad y validez de los métodos y herramientas para evaluar la credibilidad del testimonio del menor

En cuanto a la evaluación de la fiabilidad de las herramientas específicas, Ruiz-Tejedor et al., (2016) en su estudio evaluaron la fiabilidad interna y la estructura factorial de la Credibility Assessment Scale (CAS), a través del análisis de 50 entrevistas a menores víctimas de abuso sexual. Los resultados del análisis factorial reflejaron que CAS presenta una estructura sólida, con cuatro dimensiones principales (coherencia interna, cantidad/calidad de detalles, comportamiento y emociones durante la entrevista y ajuste del relato a pruebas externas) y una alta consistencia interna (α de Cronbach elevado). Además, se observó un alto porcentaje de acuerdo entre los interevaluadores, lo que respalda la fiabilidad de la prueba.

Por todo ello, los autores concluyen que la CAS es una herramienta útil y prometedora en la evaluación de la credibilidad del testimonio infantil, pero recomiendan ampliar la investigación con más muestras que ayuden a reforzar su validez externa.

El estudio realizado en Suiza por Niveau et al., 2015, pretende determinar la validez y fiabilidad de la herramienta CBCA, analizando la fiabilidad de los interevaluadores cuando aplicaban la prueba CBCA, donde se demostró que dicha fiabilidad era moderada, ya que hubo puntuaciones altas de concordancia en algunos criterios, pero había puntuaciones bajas de concordancia en otros muchos, lo que indica que hay criterios en la prueba que pueden ser interpretados de forma subjetiva, pudiendo afectar en la evaluación del testimonio. Además, los resultados indicaron que la formación y experiencia de los evaluadores influye en gran medida en la fiabilidad de las evaluaciones. Todo esto subraya la necesidad de realizar más investigaciones para optimizar la herramienta y garantizar su validez en diferentes contextos forenses.

Welle et al., (2016) tienen como objetivo evaluar la validez y fiabilidad del *Criteria-Based Content Analysis* (CBCA) en un contexto forense real, determinando qué criterios predicen mejor la veracidad del testimonio y analizando el efecto de la edad. Los resultados del estudio mostraron que CBCA presenta una fiabilidad interevaluador moderada ($ICC = 0,76$) y que las declaraciones confirmadas obtuvieron puntuaciones más altas que las no confirmadas.

Además, la herramienta reflejó una alta capacidad en la identificación de los testimonios verdaderos (hasta un 90% en casos confirmados), pero se detectó una baja precisión para detectar testimonios falsos (40–50% en casos no confirmados). Por tanto, los autores concluyen, que la fiabilidad de la prueba es moderada, pero recomiendan usarla como herramienta complementaria ya que presenta limitaciones para diferenciar de forma fiable entre relatos verídicos y falsos y sugieren criterios que ayuden a medir y detectar posibles testimonios falsos, además de recomendar que los casos sean analizados por dos evaluadores independientes

Por otro lado, Voogt et al. (2021) desarrollaron y validaron la Child Sexual Assault Victim Credibility Scale (CSAVC) con el fin de medir la credibilidad del testimonio desde cuatro dimensiones clave: veracidad percibida, coherencia emocional, consistencia del relato y credibilidad general. Para medir la validez de la prueba, los autores realizaron un análisis estadístico de fiabilidad interna,

demostrando una estructura sólida y afirmando una alta consistencia entre los ítems. Aunque no se probó con entrevistas reales, los autores resaltan su fiabilidad como instrumento complementario, animando a realizar futuras investigaciones para determinar su uso en contextos reales.

Factores que influyen en la percepción de la credibilidad

En relación con los factores que influyen en la percepción de la credibilidad, Ernberg et al. (2018), realizaron un estudio en Suecia, para determinar si la valoración de los jueces se veía influenciada por sesgos y criterios no uniformes.

Tras los resultados obtenidos a través de un análisis de contenido, los autores afirman, que además de los protocolos y herramientas utilizadas en la evaluación de la credibilidad del testimonio, existen factores como la edad del menor, la coherencia del relato y la existencia de pruebas de apoyo, que influyen considerablemente en la credibilidad percibida por el evaluador. Además, concluyen que no existen criterios uniformes entre los jueces y que las valoraciones judiciales analizadas, reflejan, la carencia de criterios basados en evidencia científica, recalcando que no solo se da por herramientas objetivas, sino también por variables subjetivas y psicosociales.

Por otro lado, Denne et al. (2021) pretenden analizar y determinar si la presencia de un testimonio experto puede ayudar a la credibilidad del menor frente a una retracción. Los resultados obtenidos tras el análisis mostraron que la retracción del testimonio reduce notablemente la credibilidad percibida por los jueces, aun explicando su causa por un testimonio experto.

Además, los autores observaron que un lenguaje más emocional o la descripción de un abuso más grave puede aumentar la credibilidad percibida y que los testimonios de niños y niñas más pequeños eran más creíbles que los relatos de los adolescentes. Por todo ello los autores concluyen que factores como la edad o la retracción del testimonio influye negativamente en la percepción del evaluador sobre la veracidad del testimonio.

En España, Cabezas-García et al. (2022) realizaron un estudio para identificar factores psicosociales que influyen en la credibilidad percibida en el testimonio de ASI. Para llevar a cabo el

estudio, se codificaron variables como edad, género, tipo de abuso, relación con el agresor, existencia de pruebas médicas o psicológicas, denuncias previas y consistencia del relato. El análisis multivariante de los resultados obtenidos mostró que la edad del menor, la coherencia del testimonio y la presencia de pruebas de respaldo tiene una alta relación con la credibilidad del testimonio del menor.

Los autores concluyen que, además de protocolos y herramientas estandarizadas, las evaluaciones judiciales están condicionadas por factores subjetivos y psicosociales, lo que evidencia la necesidad de instrumentos objetivos y específicos para garantizar decisiones justas y coherentes.

4- Discusión

El ASI es un gran problema global que impacta de manera muy negativa no solo en la vida de los menores sino también en la sociedad, la justicia y, por supuesto, la salud pública. Dado que la mayoría de los casos de abuso sexual hacia menores ocurre en contextos privados, el testimonio del menor suele ser la única prueba que suele estar disponible para analizar. Sin embargo, este relato puede llegar a estar condicionado por factores como: la vergüenza que siente el menor, la culpa o la demora en la revelación de abuso. Todo ello entorpece y complica su valoración judicial.

De hecho, todos estos elementos junto con la necesidad de protocolos estandarizados y profesionales especializados evidencian lo siguiente: que la evaluación del testimonio infantil sigue siendo uno de los mayores retos en el ámbito jurídico-forense.

Los estudios que se han revisado muestran que, para evaluar la credibilidad del testimonio en los menores, la entrevista forense sigue siendo, en todo momento, la herramienta más empleada por los expertos. No obstante, se trata de una herramienta que presenta deficiencias metodológicas, las cuales pueden comprometer la fiabilidad del relato, lo cual genera una cierta inseguridad en la valoración pericial y judicial.

Investigaciones como las de Peixoto et al. (2017) o Farrugia et al. (2018) reflejan que la mayoría de las preguntas empleadas son de dos tipos: cerradas y sugestivas. Esto supone una restricción en cuanto a la posibilidad de obtener un relato que sea detallado y mucho más espontáneo.

Esta importante limitación impide que se pueda obtener una calidad de la información de gran fiabilidad y también aumenta el riesgo de distorsiones, debido a la posible contaminación de la memoria del menor, pudiendo inducir una serie de falsos recuerdos. En este sentido, la implementación del protocolo NICHD es una propuesta mucho más sólida, ya que utiliza preguntas abiertas que favorecen la narración libre, mejorando así la validez del testimonio.

Karni et al. (2019) destacan que incluir la expresión emocional en las entrevistas forenses puede llegar a enriquecer el testimonio del menor, por lo que se promueve una mayor conexión emocional

con el relato, lo que aumenta su percepción de credibilidad. Esta idea refuerza la importancia de integrar toda clase de cuestiones de tipo emocional, no solo con la intención de mejorar la calidad del testimonio, sino que también con el fin de facilitar que, en todo momento, los menores se sienten cómodos y con menos presión para contar mejor sus relatos, lo cual es vital para conseguir fortalecer la fiabilidad del testimonio en el ámbito judicial. En contraposición a lo anteriormente mencionado, Katz et al. (2019) advierten que el uso de estrategias de confrontación directa en las entrevistas estructuradas bajo el protocolo NICHD puede ser contraproducente, llegando a generar confusión, miedo e incluso retracciones.

Aunque la inclusión de elementos emocionales en la entrevista puede llegar a ser beneficiosa, las estrategias empleadas deben usarse con suma cautela y mucha precaución. Los resultados de ambos estudios destacan que la formación del entrevistador es un elemento totalmente crucial para conseguir gestionar de una manera adecuada las emociones del menor y también para evitar una serie de técnicas que no sean las más idóneas y distorsionen el relato, garantizando la fiabilidad del testimonio. En este sentido, el valor probatorio de la entrevista depende de dos factores, siendo uno de ellos el instrumento que se ha empleado para realizarla, mientras que el otro es la propia competencia técnica, la formación y la sensibilidad del entrevistador. También se ha de decir que, cuando hablamos de los instrumentos para analizar la credibilidad, la evidencia es bastante variada.

Por un lado, la CAS ha demostrado tener una base realmente sólida en cuanto a psicométrica y una alta consistencia interna, esto se refleja en un gran nivel de acuerdo entre todos los evaluadores que la han empleado, aumentando así su fiabilidad (Ruiz-Tejedor et al., 2016) lo que la ha convertido en una herramienta prometedora para el campo del testimonio infantil.

Otro de los instrumentos más utilizados es el CBCA, que presenta una fiabilidad entre inter-evaluadores de tipo moderada, demostrando una alta variabilidad en la interpretación de criterios y disminuyendo su fiabilidad (Niveau et al., 2015; Welle et al., 2016).

Además, como muestran Welle et al. (2016), aunque discrimina con una precisión de hasta un 90% relatos verdaderos, muestra baja capacidad para identificar testimonios falsos. Esto determina que hay que entenderlo como una herramienta de tipo complementario, no principal ni de base, sino adicional en todo momento.

En este contexto, el modelo alternativo como el FIT (Farrugia et al., 2018) ofrece un avance significativo en la evaluación de la credibilidad del testimonio infantil. FIT es una herramienta para auditar entrevistas forenses de manera objetiva, permitiendo detectar malas prácticas y verificar si las entrevistas se llevan a cabo siguiendo los estándares internacionales de buenas prácticas. Ambos modelos, han incorporado mayor control en la estandarización, marcando un avance en la evaluación forense, asegurando que las entrevistas se realicen de manera adecuada y respetuosa.

De igual manera, se han creado sistemas automatizados, como el desarrollado por Szojka et al. (2025), que aportan innovación en la evaluación del testimonio infantil, permitiendo la codificación sistemática de las preguntas realizadas durante las entrevistas forenses. Estos sistemas ayudan a identificar si las preguntas son abiertas, cerradas o sugestivas, detectando posibles errores metodológicos y facilitando un control de calidad más objetivo y consistente. Todo esto ayuda a reducir la subjetividad del evaluador y aumenta la fiabilidad de la información obtenida.

Por otra parte, escalas como la CSAVC (Voogt et al., 2021) incorporan nuevas dimensiones de análisis, como la coherencia emocional, la consistencia percibida y la veracidad subjetiva, enriqueciendo la evaluación del testimonio y ofreciendo criterios más amplios para determinar su credibilidad. La fusión de estos sistemas tecnológicos con herramientas psicométricas permite un enfoque multidimensional, que combina precisión y rigor científico con la valoración clínica y ética del profesional, mejorando así la calidad y fiabilidad de la evaluación en contextos judiciales.

Otro aspecto que se debe tener en cuenta y que es sumamente frecuente en la mayoría de los estudios es la falta de estandarización y de rigor metodológico en muchas entrevistas. Rodríguez et al. (2018) tras su estudio señalaron problemas como la ausencia de consentimiento informado, hipótesis

mal formuladas y escasa replicabilidad, lo que lleva a cuestionarnos que a veces evaluamos más la calidad de la práctica profesional que la credibilidad real del menor.

Por su parte, las investigaciones de Ernberg et al. (2018), Denne et al. (2021) y Cabezas-García et al. (2022) reflejan que factores externos como son: edad, coherencia del relato, existencia de pruebas de apoyo y uso de un lenguaje más emocional. Todo ello influye de manera significativa en la percepción de credibilidad, más allá de los protocolos o técnicas utilizadas.

Estos hallazgos indican claramente que, aunque se implementen protocolos estrictos como NICHD, SVA o CBCA, la evaluación del testimonio no se puede basar exclusivamente en la herramienta utilizada, sino también en cómo las variables externas influyen en la percepción del evaluador. Esto resalta la importancia de considerar de manera integral los instrumentos, supervisando y asegurando que los profesionales estén bien preparados para interpretar estas variables y minimizar los sesgos en la valoración de la credibilidad.

5- Conclusiones

La credibilidad del testimonio de un menor no se determina solo por criterios técnicos, sino que también está influenciada por factores sociales y psicosociales que, inevitablemente, afectan la evaluación del relato del menor.

En términos prácticos, se concluye que la fiabilidad del testimonio infantil solo se puede asegurar a través de un enfoque multidimensional combinando en la evaluación: modelos como Barnahus, que nos proporcionan entrevistas bien estructuradas (como las del protocolo NICHHD), medidas de protección procesal como la prueba preconstituida y herramientas de análisis como CAS, CBCA o FIT, siempre bajo la supervisión de profesionales capacitados y en constante actualización.

El verdadero desafío no radica en crear nuevos métodos, sino en garantizar que los que ya existen se apliquen con rigor, minimizando la variabilidad entre evaluadores y la inseguridad jurídica que aún persiste en este ámbito. A pesar de los avances, las lagunas son evidentes: falta de consenso entre los profesionales, la necesidad de replicar estudios con muestras más amplias y la ausencia de validación intercultural. Este vacío, en mi opinión, no es solo un asunto científico, sino también ético y jurídico, ya que mientras no haya un consenso, las decisiones judiciales seguirán dependiendo en gran medida de la subjetividad de peritos o jueces.

La prioridad para el futuro debería ser la validación internacional de protocolos, la mejora en la formación profesional y el desarrollo de herramientas complementarias —incluyendo las tecnológicas— que permitan objetivar el análisis sin perder de vista la protección del menor.

6- Bibliografía

- Alonso García, C. R. (2022). La prueba del abuso sexual infantil: Posibilidades y límites. *Anuario Jurídico y Económico Escurialense*, LV, 177–204.
- Cabezas-García, M., de la Peña Fernández, M. E., & Andreu Rodríguez, J. M. (2022). Psychosocial factors associated with the assessment of testimony credibility in child sexual abuse. *Anales de Psicología*, 38(2), 307–315
- Delgado, C., Almonte, C., & Araya, C. (2015). *Abuso sexual infantil: características y consecuencias en una muestra de víctimas en España*. *Estudios Jurídicos*, 15(2), 83–108.
- Denne, L. D., Bottoms, B. L., & Goodman, G. S. (2021). The effects of evidence-based expert testimony on perceptions of child sexual abuse involving recantation. *Law and Human Behavior*, 45(2), 95–110.
- Ernberg, E., Magnusson, M., Landström, S., & Akehurst, L. (2018). Court evaluations of young children's testimony in child sexual abuse cases. *Psychology, Crime & Law*, 24(5), 509–527.
- Farrugia, L. D., Oxburgh, G. E., & Gabbert, F. (2018). Effective evaluation of forensic interviews: The Forensic Interview Trace (FIT). *International Journal of Speech, Language and the Law*, 25(2), 245–270.
- García, A. Z. (2016). Evaluación psicosocial del abuso sexual infantil: factores de riesgo y protección, indicadores, técnicas, y procedimientos de evaluación. *Apuntes de psicología*, 201-209.
- García, C. R. A. (2022). La prueba del abuso sexual infantil. Posibilidades y límites. *Anuario jurídico y económico escurialense*, (55), 177-204.
- Girón, R. (2018). *El testimonio infantil en casos de abuso sexual: entre las garantías procesales y la protección del menor*. *Revista de la Universidad Femenina del Sagrado Corazón*, 12(1), 45–60.
- González, A. (2016). *Testimonio del menor víctima de abuso sexual: técnicas de credibilidad y prevención de la victimización secundaria*. [Trabajo de grado, Universidad Cooperativa de Colombia]. Repositorio Institucional UCC.

- Gutiérrez Morante, L. (2021). Evaluación del abuso sexual infantil en la actualidad: Revisión teórica y aspectos metodológicos. *Revista de Victimología*, (13), 45–66.
- Higgins, J. P. T., & Green, S. (Eds.). (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* [updated March 2011]. The Cochrane Collaboration.
- Karni-Visel, Y., Hershkowitz, I., Lamb, M. E., & Blasbalg, U. (2019). *Facilitating the expression of emotions by alleged victims of child abuse during investigative interviews using the Revised NICHD Protocol*. *Child Maltreatment*, 24(3), 310–318
- Katz, C., & Barnett, Z. (2018). “I know it because it happened to me!” Confrontations of children within forensic investigations. *Child Abuse & Neglect*, 79, 203–211.
- López, F. (2002). *La credibilidad del testimonio infantil en los delitos de abuso sexual*. *Revista de Derecho y Ciencias Penales*, 5(10), 203–220.
- Losada, A. V. (2012). Epidemiología del abuso sexual infantil. *Revista de Psicología GEPU*, 3(1), 201-229.
- Niveau, G., Lacasa, M.-J., Berclaz, M., & Germond, M. (2015). Inter-rater Reliability of Criteria-Based Content Analysis of Children’s Statements of Abuse. *Journal of Forensic Sciences*, 60(5), 1247–1252.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). *Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas*. *Revista Española de Cardiología*, 74(9), 790–799.
- Peixoto, C. E., Fernandes, R. V., Almeida, T. S., Silva, J. M., La Rooy, D., Ribeiro, C., Magalhães, T., & Lamb, M. E. (2017). Interviews of children in a Portuguese special judicial procedure. *Behavioral Sciences & the Law*, 35(3), 189–203
- Pineda, D., Muris, P., Martínez-Martínez, A., & Piqueras, J. A. (2023). Prevalence of child sexual abuse in Spain: A survey study. *European Journal of Psychology Applied to Legal Context*, 15(2), 83-88.

- Pinto Moura, A. G. L., & Pérez Zambón, S. (2020). La confiabilidad de las técnicas proyectivas gráficas en la pericia psicológica del abuso sexual. *Revista Subjetividad y Procesos Cognitivos*, 24(1), 42-63.
- Posada, N. E., & Cañón, C. I. (2022). *Criterios de valoración del psicólogo jurídico en el abordaje del testimonio infantil en casos de posible abuso sexual como parte de un concepto pericial*. *Revista Criminalidad*, 64(1), 197–213.
- Ramírez Rojas, I. E. (2018). *Valoración judicial del examen de credibilidad del testimonio realizado a menores entre 4 y 7 años en delitos de abuso sexual infantil* (Memoria para optar al Grado Académico de Licenciado en Ciencias Jurídicas y Sociales, Universidad de Chile). Facultad de Derecho.
- Rodríguez Cely, L. A., Ochoa-Carrascal, M. S., Perilla, P. A., & Amaya-Nassar, S. (2018). Estudio sobre la evidencia demostrativa y el uso de protocolos en casos de abuso sexual infantil en Colombia. *Criminalidad*, 60(1), 133–151.
- Ruiz-Tejedor, M. P., Andreu, J. M., & Peña, M. E. (2016). Análisis preliminar de la estructura y consistencia interna de un protocolo clínico-pericial para la identificación de falsas alegaciones de abuso sexual infantil. *International Journal of Psychology and Psychological Therapy*, 16(1), 73–82.
- Sotoca, F., Gómez, M., & González, A. (2012). *La prueba preconstituida en los delitos de abuso sexual infantil: Garantías procesales y protección de la víctima*. *Revista de Derecho Penal y Criminología*, (8), 93–118.
- Szojka, Z. A., Yashraj, S., & Lyon, T. D. (2025). Automated question type coding of forensic interviews and trial testimony in child sexual abuse cases. *Law and Human Behavior*, 49(2), 163–172.
- United Nations Office on Drugs and Crime. (2010). *Handbook for professionals and policymakers on justice in matters involving child victims and witnesses of crime*. United Nations.

- Vara, A., Manzanero, A. L., & Vallet, R. (2021). Víctimas de abuso sexual especialmente vulnerables: obtención del testimonio. *Derecho & Sociedad*, (57), 1-39.
- Voogt, A., Klettke, B., & Mohebbi, M. (2021). The development and validation of the Child Sexual Assault Victim Credibility Scale: An instrument to measure laypersons' perceptions of victim credibility. *Journal of Interpersonal Violence*, 36(1–2), NP850–NP870.
- Welle, I., Berclaz, M., Lacasa, M.-J., & 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*, 44, 124–128
- World Health Organization. (2002). *World report on violence and health*. World Health Organization.
- World Health Organization. (2017). *Responding to children and adolescents who have been sexually abused: WHO clinical guidelines*. World Health Organization.