



Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.  
ISSN 2707-2207 / ISSN 2707-2215 (en línea), julio-agosto 2026,  
Volumen 10, Número 4.

[https://doi.org/10.37811/cl\\_rcm.v10i4](https://doi.org/10.37811/cl_rcm.v10i4)

## **VALIDACIÓN O AUTOVALIDACIÓN: EL MODELO CAVAS-INSCRIM EN LA PRUEBA PERICIAL DE DELITOS SEXUALES**

**VALIDATION OR SELF-VALIDATION: THE CAVAS-INSCRIM  
MODEL IN FORENSIC EVIDENCE OF SEXUAL CRIMES**

**Marcelo Meriño Aravena**  
Investigador independiente

## Validación o autovalidación: el modelo CAVAS-INSCRIM en la prueba pericial de delitos sexuales

Marcelo Meriño Aravena<sup>1</sup>

[marcelo.merino@merinoyabogadosasociados.cl](mailto:marcelo.merino@merinoyabogadosasociados.cl)

<https://orcid.org/0000-0002-9186-1054>

Investigador independiente

Temuco, Chile

### RESUMEN

El modelo CAVAS-INSCRIM, del Instituto de Criminología de la Policía de Investigaciones de Chile, se emplea en la evaluación pericial de presuntas víctimas de delitos sexuales. Este artículo sostiene que bajo ese nombre conviven tres objetos distintos —una metodología integral, un sistema de análisis de credibilidad y un protocolo de entrevista— y que la evidencia disponible valida únicamente componentes procedimentales del tercero, sin autorizar las inferencias sustantivas del modelo. Mediante un análisis crítico documental de fuentes institucionales y de literatura forense y de la memoria revisada por pares, se muestra que la validación es de circuito cerrado: generada, aplicada, evaluada y publicada por la propia institución, sin replicación independiente ni tasa de error conocida. Se argumenta que su aceptación forense descansa en dos sesgos —autovalidación institucional y deferencia judicial hacia el origen policial— que sustituyen el control de fiabilidad por un argumento de autoridad. Se muestra, además, que la Ley N.º 21.057 desplazó la evaluación de credibilidad hacia el tribunal, dejando indeterminado el objeto de la pericia. El uso probatorio del modelo plantea problemas de admisibilidad y de garantías del debido proceso.

**Palabras clave:** CAVAS-INSCRIM; evaluación de credibilidad; prueba pericial psicológica; validez y confiabilidad; admisibilidad.

---

<sup>1</sup> Autor principal

Correspondencia: [marcelo.merino@merinoyabogadosasociados.cl](mailto:marcelo.merino@merinoyabogadosasociados.cl)

## Validation or self-validation: the CAVAS-INSCRIM model in forensic evidence of sexual crimes

### ABSTRACT

The CAVAS-INSCRIM model, of the Institute of Criminology of the Chilean Investigative Police, is used in the forensic assessment of alleged victims of sexual offences. This article argues that the label conceals three distinct objects —an integral methodology, a credibility-analysis system, and an interview protocol— and that available evidence validates only procedural components of the third, without warranting the model's substantive inferences. Through a critical documentary analysis of institutional sources and peer-reviewed forensic and memory literature, it shows that validation is closed-loop: generated, applied, assessed, and published by the institution itself, without independent replication or a known error rate. Its forensic acceptance rests on two biases —institutional self-validation and judicial deference to its police origin— that replace reliability scrutiny with an argument from authority. It further shows that Law No. 21,057 shifted credibility assessment to the court, leaving the object of the expert examination undefined. The model's evidentiary use raises problems of admissibility and due-process guarantees.

**Keywords:** CAVAS-INSCRIM; credibility assessment; forensic psychological evidence; validity and reliability; admissibility.

*Artículo recibido 20 mayo 2026  
Aceptado para publicación: 20 junio 2026*



## 1. INTRODUCCIÓN

En Chile, buena parte de la prueba pericial psicológica sobre credibilidad del testimonio y daño en causas por delitos sexuales proviene del Centro de Asistencia a Víctimas de Atentados Sexuales (CAVAS), unidad del Instituto de Criminología (InsCrim) de la Policía de Investigaciones (PDI), creada en 1987. El CAVAS opera con una metodología propia, denominada CAVAS-INSCRIM, que sus documentos describen como un sistema integral que abarca el análisis criminológico del caso, la entrevista pericial, el análisis del testimonio mediante criterios de credibilidad, la evaluación del funcionamiento psicológico, la valoración del daño y una integración transversal de las fuentes (Blanco, 2016). Sus conclusiones se incorporan a juicios orales y contribuyen a decisiones condenatorias y absolutorias.

El problema que aborda este artículo no es la existencia del modelo ni la legitimidad de peritar en este ámbito, sino la brecha entre el uso probatorio del CAVAS-INSCRIM y la demostración pública de su validez conforme a los estándares que su función forense exige. Esa brecha tiene un rasgo peculiar que condiciona incluso su estudio: la institución no publica de forma completa y verificable el instrumento que declara aplicar. Los manuales, las pautas de codificación y los estudios de validación no son accesibles por fuentes públicas, lo que impide la evaluación independiente, la replicación de los procedimientos y el contraperitaje. Como se argumentará, esa opacidad no es un detalle administrativo: es parte constitutiva del problema científico y jurídico.

La tesis de este trabajo es doble. En su formulación moderada, sostiene que la evidencia disponible valida únicamente componentes procedimentales del protocolo de entrevista —el tipo y la distribución de las intervenciones del entrevistador— y no autoriza las inferencias del modelo sobre credibilidad, ocurrencia del hecho o causalidad del daño. En su formulación crítica, sostiene que la expresión “validación del CAVAS-INSCRIM” encubre objetos metodológicos distintos y un circuito cerrado de legitimación, y que su aceptación forense se apoya en dos sesgos —la autovalidación institucional y la deferencia judicial hacia el origen policial del modelo— que reemplazan el escrutinio de fiabilidad por un argumento de autoridad.

El criterio que ordena todo el análisis es el principio de replicabilidad independiente: una metodología forense es científicamente válida en la medida en que está especificada con transparencia suficiente para

que un investigador ajeno a la institución que la creó pueda aplicarla, reproducir sus resultados y estimar su tasa de error. Este criterio, corolario del núcleo epistemológico de los estándares de admisibilidad (Daubert v. Merrell Dow, 1993), permite evaluar el modelo sin necesidad de disponer de datos primarios y sin incurrir en el error simétrico de exigirle una perfección de la que carecen también sus referentes internacionales.

El artículo procede por análisis crítico documental. No constituye una revisión sistemática ni pretende agotar la literatura, sino reconstruir el modelo desde sus propias fuentes, contrastarlo con estándares y evidencia indexada, e identificar los sesgos que operan en su recepción. Como traza empírica complementaria, se presentó ante la PDI una solicitud de acceso a la información pública (Ley N.º 20.285) requiriendo los manuales, los estudios de validación, los estudios de confiabilidad interevaluador y de tasa de error del método (folio AD010T0033870, ingresada el 6 de agosto de 2026; plazo de respuesta al 3 de septiembre de 2026). Su resultado —entrega, entrega parcial o denegación— constituirá, en sí mismo, un dato sobre el grado de publicidad del modelo. A la fecha de cierre de este manuscrito, dicha solicitud se encuentra pendiente de respuesta; sus resultados serán objeto de un artículo posterior, en el que se contrastarán las tesis aquí sostenidas con la documentación que eventualmente se obtenga y se revisará, confirmando o corrigiendo, lo aquí planteado.

## **2. El objeto: tres cosas bajo un mismo nombre**

Una fuente recurrente de confusión consiste en tratar al CAVAS-INSCRIM como una entidad única. Conviene, primero, delimitarlo del área reparatoria del CAVAS —orientada a la intervención psicosocial—, pues la crítica recae sobre el área pericial. Pero incluso dentro de lo pericial, la denominación agrupa al menos tres objetos metodológicos con lógicas y con estados de validación distintos, y la apariencia de solidez del conjunto depende de no distinguirlos.

El primer objeto (A) es la metodología integral: el sistema clínico-pericial que va del análisis criminológico y la formulación de hipótesis a la entrevista, el análisis del testimonio, la evaluación del funcionamiento psicológico, la valoración del daño y la integración transversal de la información (Blanco, 2016). El segundo objeto (B) es el sistema de análisis de credibilidad: los diecisiete criterios discursivos de forma y de contenido con los que se infiere la credibilidad del relato, tributarios del Criteria-Based Content Analysis (CBCA) y de la hipótesis de Undeutsch (Navarro, 2006). El tercer

objeto (C) es el protocolo de entrevista pericial: la guía semiestructurada que ordena las fases, los tipos de pregunta y las reglas de la entrevista (Vega et al., 2022).

Esta distinción es la contribución analítica central del artículo, porque la evidencia empírica invocada para “validar” el modelo se refiere, casi exclusivamente, al objeto C —y aun allí, a un aspecto acotado—, mientras las inferencias con peso probatorio provienen de los objetos B y A. Sostener que el modelo está validado porque se validó un componente de su entrevista equivale a una falacia de transferencia de validez, que se examinará en el apartado 4.

### **3. Marco de estándares: validez, admisibilidad y replicabilidad**

La evaluación de un instrumento pericial exige distinguir tipos de validez y no dar por probado el conjunto cuando solo se ha aportado uno de ellos. La validez de contenido pregunta si las categorías cubren lo relevante según expertos; la de constructo, si esas categorías representan constructos psicológicos definidos; la de criterio, si los resultados se relacionan con un criterio externo confiable; la discriminante, si el instrumento distingue una experiencia real de una sugerida, de un error sincero o de una falsedad; la ecológica, si funciona en casos reales; y la incremental, si aporta algo por sobre una buena entrevista y el examen del expediente. Un método de consenso entre expertos puede aportar validez de contenido, pero no sustituye a las restantes.

En el plano jurídico, el estándar Daubert (Daubert v. Merrell Dow, 1993) exige que la técnica sea falsable, que tenga una tasa de error conocida, que haya sido sometida a revisión por pares y publicación, y que goce de aceptación en la comunidad científica. Su núcleo epistemológico es la replicabilidad intersubjetiva: un método científico debe poder ser reproducido y potencialmente refutado por terceros.

En el proceso penal chileno, la admisibilidad se articula por la vía del control de fiabilidad y de la valoración conforme a la sana crítica; la doctrina nacional ha subrayado que la confiabilidad es el requisito que más dudas despierta en el peritaje psicológico (Horvitz & López, 2004) y que, sin un informe con contenido suficiente —metodología, validez, márgenes de error, fuentes y respaldo bibliográfico—, no es posible controlar racionalmente la prueba; una lectura minimalista del artículo 315 del Código Procesal Penal “favorecería la admisión de informes de mala calidad” (Duce, 2025).

A ese marco se añade un dato institucional documentado en literatura indexada: los laboratorios periciales chilenos, incluida la PDI, carecen de la acreditación forense internacional pertinente (ISO/IEC



17025:2017) y dependen orgánicamente de las agencias de persecución penal, lo que constituye una fuente estructural de sesgo; además, los peritos no son inmunes a los sesgos cognitivos, contra la extendida “falacia de la inmunidad del experto”, y la libre valoración judicial puede degradarse en una “deferencia epistémica” hacia el experto (Duce, 2026, siguiendo a Taruffo, 2008). De aquí se sigue el criterio rector: un método que solo “funciona” en manos de sus creadores, y cuyo protocolo no es público, confunde consenso interno con validez y vulnera la igualdad de armas.

#### **4. La evidencia empírica del modelo: análisis crítico**

##### **4.1. El estudio fundacional de credibilidad (Navarro, 2006)**

El único estudio que aborda la validez del análisis de credibilidad del CAVAS-INSCRIM es la tesis de magíster de Navarro (2006). Su diseño ilustra, a la vez, un esfuerzo pionero y las limitaciones que impiden tratarlo como una validación diagnóstica. Sobre una muestra de treinta y dos relatos de niños, niñas y adolescentes —veintidós de casos considerados comprobados por prueba externa y diez de casos no confirmados—, tres juezas expertas codificaron cada relato con el CBCA y con los criterios CAVAS-INSCRIM. El estudio informó una validez concurrente alta entre ambos sistemas ( $R_{sp} = 0,87$ ), concordancia interjueces elevada y diferencias significativas entre los dos grupos.

Tres problemas limitan el alcance de estas conclusiones. El primero es la circularidad institucional: la autora es co-creadora del método, las tres juezas pertenecían al mismo Equipo Pericial del InsCrim y habían sido entrenadas en la técnica, y la muestra provino de las propias pericias del centro. La propia autora advierte que la alta concordancia podría explicarse por la formación común de las juezas y no por el instrumento, y sugiere replicar el estudio con codificadores sin entrenamiento (Navarro, 2006) —replicación que, hasta donde alcanza la evidencia pública, no se ha realizado—. El segundo problema es el criterio de verdad: el grupo “no confirmado” no equivale a relatos falsos; la propia tesis reconoce que “ni los resultados negativos... ni la ausencia de medios de prueba... serían suficientes para establecer la no ocurrencia” del abuso. Sin un grupo de declaraciones falsas conocidas no pueden estimarse sensibilidad, especificidad ni valores predictivos, que son precisamente los parámetros que un uso probatorio requeriría. El tercero es lógico: la validez del CAVAS-INSCRIM se infiere de su correlación con el CBCA y de allí se extrapola la validez que el CBCA habría mostrado en el extranjero; pero la correlación entre dos listados aplicados por las mismas juezas a los mismos relatos no demuestra

validez de criterio, y la escala de puntajes utilizada figura, en la propia tesis, con “fuente desconocida” y sin evidencia empírica de validez.

#### **4.2. La descripción institucional y la defensa epistemológica (Huellas, 2016)**

El volumen institucional Huellas (PDI, 2016) reúne la descripción oficial del modelo y su defensa. Blanco (2016) reconoce que los criterios “surgen de un proceso inductivo... basado en la constatación de regularidades observadas en los relatos” a lo largo de más de una década de casuística, y que su valoración se realiza “desde un enfoque comprensivo interpretativo... sin realizarse, bajo ninguna circunstancia, un análisis en términos cuantitativos”, sin estándar uniforme ni punto de corte; una nota del propio texto admite que el protocolo de entrevista estaba entonces “en proceso de validación”. Romo (2016), por su parte, no aporta evidencia de validación: argumenta que a la pericia no cabe exigirle validez y confiabilidad cuantitativas, por tratarse de un estudio de caso único, cualitativo y hermenéutico, y propone sustituir esos criterios por los de Guba y Lincoln, con la triangulación como garantía —una “póliza de seguros”, en sus términos—. El mismo trabajo reconoce que las pericias de credibilidad han sido progresivamente desestimadas en juicios orales.

El movimiento merece ser tomado en serio y, a la vez, mostrado en su insuficiencia. El enfoque comprensivo es legítimo para formular hipótesis clínicas, pero no exime de demostrar poder discriminativo a un dispositivo que se ofrece al tribunal precisamente para distinguir relatos creíbles de no creíbles. La triangulación controla la consistencia interna del análisis, no su validez frente a un criterio de verdad independiente. Reencuadrar la ausencia de validación como una virtud del paradigma cualitativo no responde a la pregunta forense: cuál es la probabilidad de que la conclusión pericial sea errónea.

#### **4.3. El protocolo de entrevista y su “validación” (Vega et al., 2022; Silva et al., 2021)**

El desarrollo más reciente corresponde al objeto C. Vega et al. (2022) documentan la construcción del Protocolo de Entrevista Pericial Psicológica CAVAS-INSCRIM mediante un “método de consensos”: grupos nominales integrados por peritos de la propia institución que, a partir de la literatura y de su experiencia, acordaron los lineamientos de la entrevista. Este procedimiento puede aportar estandarización y validez de contenido, pero el consenso experto no es validación empírica: no demuestra que las conclusiones derivadas del protocolo discriminen hipótesis verdaderas de falsas.



La validación propiamente tal se atribuye a un estudio pre/post (Silva et al., 2021) publicado en una revista policial interna. Según la descripción disponible, el estudio comparó cincuenta y cuatro entrevistas —veintisiete antes y veintisiete después de la capacitación en el protocolo— y halló una reducción cercana al 23% de las intervenciones del entrevistador y un aumento de las preguntas apropiadas, sin incremento significativo de las categorías de información investigativamente relevante ni de la información sobre daño. Con la cautela de que estas cifras deben verificarse en la fuente, su alcance es claro: se trata de evidencia preliminar sobre la conducta verbal del entrevistador —adherencia procedimental y calidad formal de las preguntas—. No demuestra que los relatos obtenidos sean verdaderos ni más exactos, que disminuyan las falsas alegaciones, que el análisis posterior de credibilidad sea válido, ni que la evaluación de daño establezca causalidad.

#### **4.4. La falacia de transferencia de validez**

De lo anterior se sigue el argumento central. La “validación de 2021” pertenece al objeto C y se refiere a un aspecto acotado de la entrevista; sin embargo, tiende a presentarse como aval del modelo en su conjunto. Atribuir a los objetos B y A —el análisis de credibilidad y la metodología integral, que son los que sostienen las inferencias sobre ocurrencia del hecho y daño— la validez preliminar demostrada para un componente procedimental de C constituye una falacia de transferencia de validez. La pregunta pertinente es si las publicaciones posteriores respondieron a las objeciones sobre la validez de los criterios de credibilidad, y la respuesta es negativa: validar la entrevista no responde a la crítica sobre los criterios, que sigue abierta desde Macurán (2014), para quien la asumida equivalencia entre CAVAS-INSCRIM y CBCA es un “acto de fe”, dado que los criterios de ambos sistemas son sustancialmente distintos, el modelo carece de un sistema explícito de codificación y solo cubre parte de los criterios del CBCA.

#### **5. Los límites de la evaluación de credibilidad a la luz de la ciencia de la memoria**

El respaldo externo más robusto de la crítica proviene de la literatura indexada sobre memoria y testimonio, y es especialmente contundente porque parte de él lo aportan los propios difusores del análisis de validez de las declaraciones en lengua española. Manzanero y González (2015) sostienen que “el mero análisis de la presencia de los denominados criterios de credibilidad no es suficiente para discriminar las declaraciones reales de las que no lo son”, y que estos procedimientos “no son útiles para

valorar la veracidad de un relato concreto, sino sólo para... discriminar grupos” de relatos; añaden que no satisfacen dos de los criterios Daubert —la tasa de error es superior a la tolerable y no gozan de aceptación amplia— y que, en ocasiones, se han hallado más rasgos de realidad en relatos falsos que en reales. Estas conclusiones son concordantes con la revisión de Vrij (2005) sobre el CBCA.

El problema se agrava porque los criterios de contenido presuponen una mentira intencional y consciente, de modo que “no son aplicables a la mayoría de los relatos falsos producto de errores... ya que los propios sujetos creen en su realidad” (Manzanero, 2004). En un estudio experimental, ese autor mostró que las escasas diferencias entre relatos reales y sugeridos se difuminan con el paso del tiempo —justo el escenario forense típico de las declaraciones tardías— y que los errores de clasificación del origen alcanzan entre un 36% y un 48%. La razón última es cognitiva: la maleabilidad de la memoria y el efecto de la información posterior al suceso hacen que memorias reales y sugeridas puedan ser fenomenológicamente indistinguibles (Loftus & Palmer, 1974; Loftus, 1993, 2005), y el juicio sobre el origen de un recuerdo es intrínsecamente falible (Johnson, Hashtroudi & Lindsay, 1993). Incluso los recuerdos de hechos traumáticos, largamente supuestos inmunes, incorporan información falsa sugerida (Manzanero, 2010), y la sugestionabilidad del testigo infantil está sólidamente documentada: la entrevista y la información post-suceso pueden generar relatos falsos que el propio niño cree verdaderos (Ceci & Bruck, 1993).

Dos consecuencias interesan al examen del CAVAS-INSCRIM. La primera es que ningún conjunto de rasgos internos del relato puede garantizar, en el caso individual, el origen real o sugerido de una memoria; los criterios de credibilidad correlacionan, en el mejor de los casos, con la exactitud a nivel de grupo, no con la verdad de una declaración concreta. La segunda es que la evaluación debe basarse en el análisis científico de las capacidades del declarante y no en juicios intuitivos: en poblaciones vulnerables, un déficit cognitivo puede confundirse con falta de credibilidad si no se evalúan antes las capacidades, por lo que debe primar el conocimiento científico sobre el estereotipo (Contreras, Silva & Manzanero, 2015). Es preciso, con todo, evitar el error simétrico: el punto no es que el CAVAS-INSCRIM sea inválido frente a un CBCA infalible —el CBCA tampoco prueba la verdad—, sino que el modelo chileno descansa en una base empírica mucho más restringida y, además, gris y endógena, mientras el CBCA acumula investigación indexada con sus límites reconocidos.



## 6. Dos sesgos que sostienen la aceptación del modelo

Si la evidencia de validez es fragmentaria y no replicada, cabe preguntarse por qué el modelo goza de aceptación forense. La respuesta remite a dos sesgos que se refuerzan mutuamente. El primero es la autovalidación institucional. El conocimiento que respalda el CAVAS-INSCRIM circula por un circuito cerrado: los criterios se generan inductivamente desde la casuística del propio centro (Blanco, 2016); el único estudio de credibilidad lo conduce una co-creadora del método con juezas del mismo equipo (Navarro, 2006); la defensa epistemológica se publica en el cuaderno de la propia PDI (Romo, 2016); y la validación del protocolo aparece en una revista policial interna (Silva et al., 2021). Ninguna réplica independiente, fuera de la institución, ha reproducido estos resultados. La autovalidación no es un defecto accesorio: es el mecanismo que hace pasar por validez lo que puede ser coincidencia entre personas formadas en la misma tradición.

El segundo sesgo es la deferencia judicial hacia el origen policial del modelo. La hipótesis —que los tribunales tienden a acreditar las conclusiones del CAVAS en parte por provenir de la PDI— es sostenible y contrastable. Cifuentes (2018) documenta que algunos magistrados manifiestan desconfianza precisamente del “alero institucional” del modelo, lo que confirma que la autoridad institucional opera como factor de valoración y que su influencia es objeto de disputa; la misma investigación registra la percepción de que los informes tienden a concluir de manera uniforme que el relato es creíble. Teóricamente, se trata de un sesgo de autoridad o heurístico de la fuente: la credibilidad de la conclusión se infiere del prestigio del emisor y no de la fiabilidad demostrada del método, fenómeno que la literatura describe como deferencia epistémica (Taruffo, 2008, en Duce, 2026) y como falacia de la inmunidad del experto (Duce, 2026). Este punto debe formularse como hipótesis a sustanciar con un corpus jurisprudencial —no como hecho probado—, distinguiendo dos efectos opuestos y coexistentes: jueces que sobrevaloran la pericia por autoridad y jueces que, por lo mismo, sospechan del sesgo institucional.

Ambos sesgos se anudan. La autovalidación priva a los tribunales de la información necesaria para controlar la fiabilidad —no hay manual público, tasa de error ni estudio de confiabilidad externo—, y en ese vacío la deferencia a la autoridad institucional ocupa el lugar del escrutinio. Adicionalmente, la propia arquitectura del modelo —que integra antecedentes de la carpeta, hipótesis, entrevista y daño sin



procedimientos de ceguera ni búsqueda activa de evidencia desconfirmatoria— favorece un sesgo de confirmación: el relato se juzga compatible por contener ciertos criterios, los síntomas se leen como compatibles con el relato y el contexto como compatible con la victimización, y la coincidencia se presenta como convergencia, cuando los tres elementos pueden provenir de la misma fuente, la misma entrevista y la misma hipótesis (Kassin, Dror & Kukucka, 2013). Sin independencia entre las fuentes, no hay corroboración.

## **7. Implicancias de admisibilidad, garantías y daño**

Las consecuencias son jurídicas y éticas. En el plano de la admisibilidad, un informe que no explicita su metodología, su validez y su margen de error, y que se apoya en un protocolo no público, no es susceptible de control racional de fiabilidad (Duce, 2025) y compromete la igualdad de armas: la defensa no puede replicar ni contraperitar aquello que no conoce. La opacidad del modelo, lejos de ser neutra, obstruye el derecho a la contradicción de la prueba. En el plano del daño, el modelo incorpora la valoración del daño psicológico como fase propia, pero los síntomas asociados a la agresión sexual son inespecíficos; de allí dos máximas que la práctica pericial no debería omitir: la presencia de daño no prueba el hecho y su ausencia no lo descarta, y la compatibilidad clínica no equivale a atribución causal individual. El mismo cuidado corresponde a los denominados elementos de relevancia forense, que orientan líneas investigativas pero no deben presentarse —ni ser recibidos judicialmente— como corroboración del hecho.

Conviene subrayar que esta crítica protege, no perjudica, a las víctimas. Una prueba pericial científicamente frágil es impugnabile, y su caída puede arrastrar consigo la persecución de agresiones efectivamente ocurridas; la revictimización que produce un peritaje mal fundado, al ser desestimado en juicio, es tan real como la que se pretende evitar. El examen recae sobre los estándares del instrumento, no sobre las intenciones de los profesionales que lo aplican.

## **8. Un objeto en crisis: la evaluación de credibilidad tras la Ley N.º 21.057**

Hay una paradoja que atraviesa todo lo anterior y que conviene explicitar. El CAVAS-INSCRIM nació para evaluar la credibilidad del testimonio: su estudio fundacional se titula, precisamente, evaluación de la credibilidad discursiva (Navarro, 2006), y sus diecisiete criterios son criterios de credibilidad tributarios del CBCA. Sin embargo, la entrada en vigor de la Ley N.º 21.057, sobre entrevistas

investigativas videograbadas para niños, niñas y adolescentes víctimas de delitos sexuales y otros delitos graves, ha desplazado esa función. La entrevista investigativa videograbada tiene una finalidad distinta de la pericia psicológica y no constituye, por sí misma, una evaluación de credibilidad; más aún, la discusión legislativa corrigió la pretensión inicial del proyecto de que el peritaje de credibilidad del relato se hiciera sobre la base del registro de esa entrevista, reafirmando que la valoración de la credibilidad es una función privativa del tribunal y no del perito (Fuentes et al., 2025; concordante con Duce, 2010).

De ahí se sigue una consecuencia que la práctica no ha resuelto: si la pericia ya no debe pronunciarse sobre la credibilidad —que era su objeto original y aquello para lo cual el modelo fue construido—, ¿qué es, entonces, lo que la pericia debe establecer? El objeto legítimo del peritaje ha quedado indeterminado. Los objetos residuales que podrían justificarlo —la evaluación de las capacidades para testificar o la valoración del daño psicológico— son tareas distintas, con instrumentos y estándares propios, para las cuales los criterios de credibilidad del CAVAS-INSCRIM no son aptos; y volver a solicitar un relato sustantivo después de la entrevista videograbada reintroduce justamente los riesgos que la ley quiso evitar: la duplicación de entrevistas y la contaminación post-acontecimiento del testimonio. Se configura así una metodología en busca de un objeto: un dispositivo diseñado para inferir credibilidad, aplicado en un marco normativo que ha retirado la credibilidad del ámbito pericial, sin que se haya redefinido con claridad qué debe hacer en su lugar.

Esta indeterminación agrava el problema de admisibilidad en lugar de atenuarlo. No es posible controlar la fiabilidad de un instrumento cuyo propósito legítimo no está definido: la pregunta previa a la validez —validez ¿para qué?— queda sin respuesta. La crisis del objeto debe delimitarse en su alcance: el desplazamiento de la credibilidad es más nítido para niños, niñas y adolescentes bajo el régimen de la entrevista videograbada que para las víctimas adultas; pero el problema conceptual de fondo —que la credibilidad es competencia del tribunal— persiste con independencia de la edad de la víctima. Resulta llamativo, por lo demás, que en la literatura reciente favorable al sistema se exija y documente profusamente la base empírica del protocolo NICHD, mientras el CAVAS-INSCRIM se menciona nominalmente y su única referencia disponible remite a un documento de construcción por consenso (Vega et al., 2022): una asimetría de estándar probatorio que el propio campo aplica sin advertirlo.

## 9. Conclusiones y agenda de validación

El CAVAS-INSCRIM evolucionó desde una metodología clínico-pericial de análisis integrado hacia un protocolo de entrevista más estructurado. Sin embargo, la evidencia disponible valida únicamente componentes procedimentales de la entrevista y no permite sostener una validación integral de sus inferencias sobre credibilidad, ocurrencia del hecho o causalidad del daño. La expresión “validación del CAVAS-INSCRIM” encubre objetos metodológicos distintos, y la validación preliminar de un protocolo de entrevista no puede transferirse al análisis de credibilidad, a la integración transversal de antecedentes ni a las conclusiones sobre el daño. La aceptación forense del modelo descansa, en medida no menor, en una autovalidación de circuito cerrado y en la deferencia a su origen institucional, que sustituyen el control de fiabilidad por un argumento de autoridad. A ello se suma una paradoja de fondo: la reforma de la entrevista investigativa videograbada retiró la credibilidad del ámbito pericial y la restituyó al tribunal, de modo que el modelo —concebido para inferir credibilidad— ha quedado sin un objeto claramente definido, lo que agrava, y no atenúa, la dificultad de controlar su fiabilidad.

La afirmación defendible no es que el modelo sea inválido —eso exigiría datos primarios—, sino que no ha demostrado públicamente su validez conforme a los estándares que su uso probatorio exige, y que su diseño de validación no satisface el requisito de replicabilidad independiente. Revertir esta situación no requiere abandonar la pericia psicológica del testimonio, sino someterla al escrutinio ordinario de la ciencia. Ello supone, al menos: publicar íntegramente el protocolo y las pautas de codificación, de modo que un investigador externo pueda aplicarlos; realizar estudios de confiabilidad interevaluador con codificadores ajenos a la institución; construir muestras con un criterio de verdad independiente que permitan estimar sensibilidad, especificidad y tasa de error; someter los estudios a revisión por pares externa en revistas indexadas; y avanzar hacia la acreditación forense y hacia diseños que controlen el sesgo de confirmación. Mientras esa agenda no se cumpla, la prudencia aconseja que los tribunales traten las conclusiones del modelo como hipótesis clínicas orientadoras y no como prueba de la ocurrencia del hecho, y que su valoración no derive del prestigio de la institución que las emite, sino de la fiabilidad demostrada del método.

La solicitud de acceso a la información pública presentada a la Policía de Investigaciones (folio AD010T0033870) pondrá a prueba, en los hechos, el grado de publicidad del modelo. Su respuesta —o



su denegación— informará no solo sobre la disponibilidad de los documentos, sino sobre la disposición institucional a permitir la evaluación independiente que la validez científica, en último término, requiere. Al tiempo de esta publicación, la diligencia continúa en curso —con plazo legal de respuesta al 3 de septiembre de 2026—; sus resultados se comunicarán en un trabajo ulterior, que permitirá confirmar, matizar o corregir las conclusiones del presente artículo a la luz de la documentación que la institución entregue o de las razones que invoque para reservarla.

## REFERENCIAS BIBLIOGRÁFICAS

Blanco González, A. (2016). Descripción de la metodología de evaluación pericial CAVAS-INSCRIM.

En Policía de Investigaciones de Chile, *Huellas: cuadernos de criminodinámicas y fenómenos emergentes* (Vol. 2, pp. 14–32). PDI.

Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis.

*Psychological Bulletin*, 113(3), 403–439. <https://doi.org/10.1037/0033-2909.113.3.403>

Cifuentes Güenes, M. (2018). *Percepción de los jueces del Tribunal Oral en lo Penal respecto del rol del perito forense y su concordancia en los factores que definen sentencias condenatorias y absolutorias en casos de delitos sexuales* [Tesis de magíster, Universidad de Valparaíso].

Contreras, M. J., Silva, E. A., & Manzanero, A. L. (2015). Evaluación de capacidades para testificar en víctimas con discapacidad intelectual. *Anuario de Psicología Jurídica*, 25, 87–96. <https://doi.org/10.1016/j.apj.2015.02.006>

Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993).

Duce, M. (2025). Los informes periciales escritos en el proceso penal en Chile: análisis crítico de su regulación y práctica. *Quaestio Facti*, (10), 1–38. [https://doi.org/10.33115/udg\\_bib/qf.i10.23191](https://doi.org/10.33115/udg_bib/qf.i10.23191)

Duce, M. (2026). Sesgos cognitivos de los peritos en la justicia penal: una visión panorámica del problema e identificación de algunas herramientas para minimizarlo. *Política Criminal*, 21(41), 97–128.

Fuentes Rebolledo, A., Duque Duvauchelle, C., Rosati Jerez, N., & Sovino Meléndez, M. (2025). Transformaciones procesales y metodológicas en la investigación penal de delitos sexuales



- contra NNA: entrevista investigativa videograbada, valoración en juicio oral y distinción con las pericias psicológicas. *Revista del Instituto de Estudios Judiciales*, 734–772.
- Horvitz Lennon, M. I., & López Masle, J. (2004). *Derecho procesal penal chileno* (Tomo II). Santiago, Chile: Editorial Jurídica de Chile.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114(1), 3–28. <https://doi.org/10.1037/0033-2909.114.1.3>
- Kassin, S. M., Dror, I. E., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, 2(1), 42–52. <https://doi.org/10.1016/j.jarmac.2013.01.001>
- Köhnken, G., Manzanero, A. L., & Scott, M. T. (2015). Análisis de la validez de las declaraciones (SVA): mitos y limitaciones. *Anuario de Psicología Jurídica*, 25, 13–19. <https://doi.org/10.1016/j.apj.2015.01.004>
- Loftus, E. F. (1993). The reality of repressed memories. *American Psychologist*, 48(5), 518–537. <https://doi.org/10.1037/0003-066X.48.5.518>
- Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory*, 12(4), 361–366. <https://doi.org/10.1101/lm.94705>
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, 13(5), 585–589. [https://doi.org/10.1016/S0022-5371\(74\)80011-3](https://doi.org/10.1016/S0022-5371(74)80011-3)
- Macurán Nodarse, G. (2014). Una prueba de fe: la metodología CAVAS-INSCRIM para la evaluación de la credibilidad de los testimonios en casos de delitos sexuales. *Doctrina y Jurisprudencia Penal*, 4(15), 3. Thomson Reuters. Cita online CL/DOC/311/2014.
- Manzanero, A. L. (2004). ¿Son realmente diferentes los relatos sobre un hecho real y los sugeridos? *Anuario de Psicología Jurídica*, 14, 115–139.
- Manzanero, A. L. (2010). Recuerdo de hechos traumáticos: de la introspección al estudio objetivo. *Psicopatología Clínica, Legal y Forense*, 10, 149–164.



- Manzanero, A. L., & González, J. L. (2015). Modelo holístico de evaluación de la prueba testifical (HELPT). *Papeles del Psicólogo*, 36(2), 125–138.
- Navarro Medel, C. (2006). *Evaluación de la credibilidad discursiva de niños, niñas y adolescentes víctimas de agresiones sexuales* [Tesis de magíster, Universidad de Chile].
- Policía de Investigaciones de Chile. (2016). *Huellas: cuadernos de criminodinámicas y fenómenos emergentes. Comprensión de los delitos sexuales: una mirada desde el CAVAS pericial* (Vol. 2). PDI–Santillana del Pacífico. ISSN 0719-6881.
- Romo Fernández, V. (2016). Rigurosidad científica de las evaluaciones periciales psicológicas en presuntas víctimas de agresión sexual. En Policía de Investigaciones de Chile, *Huellas* (Vol. 2, pp. 33–51). PDI.
- Silva Contreras, S., Vega Lizana, A., Avilés Viveros, G., Camplá Bolívar, X., & Adrián Serrano, L. (2021). Validación del protocolo de entrevista pericial psicológica CAVAS-INSCRIM en las intervenciones con niños, niñas y adolescentes. *Formación y Desarrollo Policial*, 4, 53–79.
- Taruffo, M. (2008). *La prueba*. Madrid, España: Marcial Pons.
- Vega, A., Avilés, G., Adrián, L., & Camplá, X. (2022). Construcción del protocolo de entrevista pericial psicológica CAVAS-INSCRIM: método de consensos. *Revista de Investigación Forense*, (10), 5–11. Instituto Dr. Carlos Ybar, Servicio Médico Legal.
- Vrij, A. (2005). Criteria-Based Content Analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law*, 11(1), 3–41. <https://doi.org/10.1037/1076-8971.11.1.3>