

# La calidad de la formulación de riesgo de violencia y un cuasi-experimento de entrenamiento en salud mental forense

## *The Quality of Violence Risk Formulation and a Quasi-Forensic Mental Health Training Experiment*

Débora Moyano<sup>1</sup>, Alejandra Estrada<sup>2</sup>, Arturo Klenner<sup>3</sup>, Adolfo Méndez<sup>4</sup>, Marlene Vidal Quinde<sup>5</sup>, Jorge O. Folino<sup>6</sup>

### Resumen

**Objetivos:** El estudio se propuso medir el impacto de un entrenamiento sobre formulación pericial del caso forense en salud mental y extender evidencias relacionadas con la confiabilidad de la TEC-F. **Método:** Nueve psiquiatras y seis psicólogos de diversos países latinoamericanos participaron de un cuasi-experimento pedagógico. Se midió independientemente y a ciegas la calidad de las formulaciones con la TEC-F pre y post intervención y también se manipuló aleatoriamente la asignación de dos viñetas estándares. Se calcularon diferencias de medias e indicadores de confiabilidad del instrumento. **Resultados:** Los valores de coeficientes de correlación intraclass fueron 0,92; 0,94; 0,83; 0,93 y 0,95 y los valores del coeficiente Alfa de Cronbach fueron 0,83; 0,94; 0,63; 0,77 y 0,93, para las dimensiones transparencia, especificidad, comunicación, fundamentación y para el total TEC-F respectivamente. Los resultados del test-retest a 19 días fueron excelentes. La media de calidad total TEC-F pre-curso fue 31,4 y la media post curso, 38,4 ( $p = 0,003$  y  $p = 0,001$  para prueba grupal y para prueba apareada respectivamente). **Conclusiones:** La intervención pedagógica produjo una significativa mejoría en la calidad de las formulaciones periciales. El estudio adicionó evidencias sosteniendo la confiabilidad de la TEC-F.

**Palabras clave:** Formulación de caso forense - Evaluación de riesgo de violencia - Informe pericial - Práctica forense - Confiabilidad - Salud mental.

### Abstract

**Objectives:** The study was aimed at measuring the impact of training on forensic case formulation in mental health and to provide more evidence on the reliability of the TEC-F. **Method:** Nine psychiatrists and six psychologists from various Latin American countries participated in a quasi-pedagogical experiment. The quality of formulations was independently and blindly measured pre- and post-intervention with the TEC-F and the assignment of two standard vignettes was also randomly manipulated. Quality mean differences and instrument reliability indicators were calculated. **Results:** The values of intraclass correlation coefficients were 0.92; 0.94; 0.83; 0.93 and 0.95 and the values of Cronbach's alpha coefficient were 0.83; 0.94; 0.63; 0.77 and 0.93, for the dimensions transparency, specificity, communication, reasoning and for the total TEC-F respectively. The results of the 19-day test-retest were excellent. The mean

RECIBIDO 30/10/2022 - ACEPTADO 18/11/2023

<sup>1</sup>Psiquiatra. Mag. (C) Salud Mental Forense, Universidad Nacional de La Plata. Perito Oficial, Poder Judicial de Córdoba, Argentina. ORCID 0000-0002-6762-1579

<sup>2</sup>Magister Psicología Jurídica y Forense, Universidad de la Frontera. Defensoría Penal Pública, Chile. ORCID 0000-0001-9027-4213

<sup>3</sup>Psicólogo. Abogado. Mag. (C) Salud Mental Forense, Universidad Nacional de La Plata. Poder Judicial de Chile, Chile. ORCID 0000-0002-0103-9869

<sup>4</sup>Psicólogo. Mag. (C) Salud Mental Forense, Universidad Nacional de La Plata. Gendarmería de Chile. ORCID 0000-0002-9925-2579

<sup>5</sup>Magister en Salud Mental Forense, Universidad Nacional de La Plata. Perito acreditada por el Consejo de la Judicatura, Ecuador. ORCID 0000-0003-0189-2461

<sup>6</sup>Doctor en Medicina. Universidad Nacional de La Plata, Argentina. Director de la Maestría en Salud Mental Forense. ORCID 0000-0001-5834-2212

### Autor correspondiente:

Jorge O. Folino

folino@med.unlp.edu.ar



TEC-F total quality pre-course was 31.4 and the mean post-course, 38.4 ( $p = 0.003$  and  $p = 0.001$  for group test and paired test respectively). Conclusions: The pedagogical intervention produced a significant improvement in the quality of the expert's formulations. The study added evidence supporting the TEC-F reliability.

**Keywords:** Forensic case formulation - Violence risk assessment - Expert report - Forensic practice - Reliability - Mental health.

## Introducción

La formulación del caso en salud mental forense es una explícita parte de la tarea pericial tendiente a exponer una fundada comprensión de las condiciones de la persona evaluada y su impacto con determinado estado o funcionalidad que le interesa al sistema judicial. La formulación eficiente contribuye al ejercicio de la justicia, pero cuando es deficiente y produce asesoramiento erróneo a los operadores judiciales puede impactar en eventos indeseables, como son la reincidencia delictiva, las injurias significativas y los traumas a terceros, el gran costo personal por privación de libertad o por tratamientos de largo plazo injustificados (Sturmey & McMurrin, 2011). La prueba pericial de baja confiabilidad, sin el suficiente sustento científico y los sesgos cognitivos de los expertos aumentan el riesgo de condenas erróneas (Duce, 2018).

Tanta trascendencia motivó la búsqueda de evidencias empíricas relacionada con la medición de la calidad de la formulación forense (Folino et al., 2022; McMurrin & Bruford, 2016), sin embargo aún permanecen poco explorados los modos de capacitación específica y sus impactos en las habilidades profesionales (Brown et al., 2018; Minoudis et al., 2013).

La adquisición de habilidades profesionales relacionadas con formulación del caso en salud mental, el ejercicio de prácticas basadas en evidencias y la eficiencia en el servicio pericial conforman una concatenada expresión del profesionalismo. Las acciones sistemáticas que tiendan a perfeccionar cualquiera de esos eslabones tienen su justificación en la búsqueda del mejor y más seguro servicio. En este artículo comunicamos los hallazgos de una intervención pedagógica destinada a la capacitación avanzada en formulación del caso forense, específicamente, en la formulación de riesgo de violencia, y de la confiabilidad de la TEC-F (Folino, 2022) para medir su calidad. A continuación, describimos el contexto legal, la práctica pericial, los antecedentes pedagógicos y la estrategia de medición de la calidad de la formulación.

## El contexto legal y la formulación pericial

Desde tiempos inmemoriales personas o "partes" en disputa han recurrido a un tercero, objetivo e imparcial, conocedor de la ley y el derecho: el juez; para resolver sus diferencias a través de reglas preestablecidas: el proceso. En muchos casos, por la naturaleza del conflicto, el objeto en que recae, la solicitud formulada -entre otros variados motivos-, ese conocimiento jurídico y legal especializado es insuficiente, por lo que se recurre a la opinión de otro tercero, experto en una ciencia, arte u oficio, en que el juez es lego: el perito.

El rol de los jurados, también como "terceros", consistió en aportar la experiencia del "sentido común" a la decisión jurisdiccional, fue cediendo terreno a los peritos, en la medida en que el conocimiento científico y tecnológico se fue consolidando y especializando. La regla general es que, tanto el jurado -en su conjunto, sin perjuicio del conocimiento experto de alguno de sus miembros- como el juez, son legos en todas las materias distintas de su quehacer profesional (Kiely, 2001; Mnookin, 2008; Shelton, 2011). La excepción a tal regla no es nueva. Hand documenta la existencia, en Gran Bretaña, de jurados expertos ya en el siglo XIV (Hand, 1900).

Desde siempre se ha reconocido, pues, una tensión en el proceso: por un lado, la inevitabilidad de que el lego juzgue la credibilidad de la opinión del experto y resuelva el conflicto en virtud de ella; y por otro, la necesidad de evitar que la opinión del experto, introducida para eliminar la o las incógnitas fácticas del caso, introduzca "el oscurantismo de la supuesta ciencia de un perito" (Nieva-Fenoll, 2018, p. 13). En otras palabras, la participación de un perito en un proceso judicial no es, por sí misma, garantía ni fuente de certeza de que se incorpore al juicio información con el mejor y más actualizado conocimiento de una ciencia, arte u oficio. Los motivos de la incertidumbre son múltiples y van desde la desconfianza y parcialidad que puede producir el pago de los servicios periciales, hasta cuestiones complejas de carácter epistemológico (Mnookin, 2008; Nieva, 2010; Niño & Marrero, 2015).

La ley, y con ella la jurisprudencia y la doctrina, incluso con independencia del sistema judicial en que nos encontremos en occidente, han procurado establecer múltiples exigencias, tanto para el juez como para el perito, que han decantado en el establecimiento de “estándares de prueba”. Lo anterior no es sinónimo de afirmar que la existencia de tales estándares y su cumplimiento sea la clave para la obtención de información atingente y de calidad, para formar convicción en el juez; ni mucho menos, que la ponderación de la prueba no pueda exigir más que los estándares, que suelen operar como requisitos o niveles mínimos, por debajo de los cuales, no es posible arribar a un juicio razonable. De hecho, por mucho que se aspira a contar con “estándares de prueba” y a cumplirlos, no están exentos de controversia en cuanto son un resabio de la prueba legal o tasada, que originalmente procuraba controlar las facultades jurisdiccionales, y que, en el sistema de jurados, intenta guiar a sus miembros. La anhelada certeza que descansa en los “estándares de prueba” no es, ni tan estándar, ni tan certeza (Nieva, 2020).

En este contexto es insoslayable -aunque inagotable en este artículo- el denominado “Caso Daubert” y las sentencias que han dado lugar, en Estados Unidos, a las “Federal Rules of Evidence”, particularmente a su artículo 702 (Kiely, 2001). Sin negar el aporte y los efectos de este caso -por controvertidos que han resultado-, particularmente en el debate de admisibilidad de la prueba (Dixon & Gill, 2002; Gatowski et al., 2001; Groscup et al., 2002; Jurs & Devito, 2013), se puede afirmar que la cuestión está muy lejos de estar zanjada (Bernstein & Lasker, 2015), mucho menos en cuanto a la ponderación de la prueba pericial, en sí y en relación con los otros medios de convicción jurisdiccional.

Eso no implica claudicar en la búsqueda de criterios que permitan estandarizar el rendimiento y la funcionalidad de la prueba pericial. Es así como en el contexto de la investigación empírica sobre formulación del caso forense, específicamente en el área de la salud mental, se han generado recientemente valiosos aportes (Bucci et al., 2016; Hart et al., 2011; Hopton et al., 2018; McMurrin & Bruford, 2016; Ryan, 2020; Wheable & Davies, 2020), aunque aún insuficientes para dar por cumplida la tarea de establecimiento de criterios de calidad del desempeño pericial, ampliamente consensuados y científicamente validados, que ofrezcan garantías a los intervinientes de cualquier proceso judicial, de que la opinión experta efectivamente lo es.

En el ámbito teórico, la cuestión se define conceptualmente en los textos procesales mediante las exigencias dirigidas a los peritos, de imparcialidad y objetividad, o conceptos afines. Determinar si estos criterios se cumplen y cómo, queda a criterio del juzgador del caso particular, con todas las dificultades teóricas y prácticas que ello conlleva y que se han esbozado previamente. En el ámbito práctico y casuístico, el dilema a resolver y el desafío a vencer son mucho mayores: no existe una regla de oro ni un baremo omnicompreensivo para ponderar la calidad y fiabilidad de las pericias; la ley no es ni el único criterio y muchas veces tampoco el más eficaz. La valoración del caso forense demanda múltiples tareas y las herramientas prácticas y epistemológicas son menos que escasas. De ahí que en el contexto forense cada aporte es bienvenido y celebrado.

Dicho lo anterior y conociendo, o a lo menos sospechando el mapa conceptual de la prueba pericial, la TEC-F resulta una herramienta novel y novedosa, que pretende facilitar el tránsito por la accidentada geografía real que corresponde al aludido mapa. De hecho, aspira a orientar al lector inexperto y enriquecer en su conocimiento al lector experimentado; espera, del mismo modo, guiar apropiadamente a quien pretenda examinar un producto pericial en términos de la calidad de su formulación.

En el contexto descrito, la Guía para la Valoración y Formulación del Caso en Salud mental forense, no establece estándares, aunque aporta valiosos elementos de juicio para ponderar si la formulación del caso forense o prueba pericial en salud mental forense satisface las dimensiones de transparencia; especificidad; comunicación y fundamentación (Folino, 2022). Estas dimensiones son plenamente coherentes con cualquier proceso judicial, en cuanto facilitan que la actuación del perito sea trazable, es decir transparente para todos los intervinientes; su intervención profesional responda a la pregunta psico-jurídica o psico-legal; tanto su informe escrito como su comparecencia ante estrado comunique con eficacia un contenido experto en un contexto lego; y, en todo este quehacer, fundamente las decisiones y actuaciones técnicas, los hallazgos, conclusiones y sugerencias, a fin de cumplir con los criterios previos y enmarcarse en la razonabilidad del procedimiento técnico y científico (Folino et al., 2022).

## La formulación pericial sobre riesgo de violencia

El desarrollo de las estrategias para evaluar el riesgo de violencia y la disponibilidad de evidencias cientí-

ficas han mostrado un incremento notable a partir de fines del previo milenio en variadas regiones de mundo, incluyendo América del Sur (Ochoa Balarezo et al., 2016). Desde la teoría se reformuló la conceptualización del tema y se estimuló la exploración y manejo de los factores de riesgo dinámicos y protectores (Douglas & Kropp, 2002; Greer et al., 2020; Monahan & Steadman, 1994). Desde los emprendimientos empíricos se iniciaron líneas de investigación en Argentina (Folino et al., 2004), en Brasil (de Borba Telles et al., 2009; de Borba Telles et al., 2012), en Colombia (Escobar-Córdoba et al., 2015; Escobar-Córdoba, 2006; Escobar-Córdoba et al., 2015), en Chile (León-Mayer et al., 2014) y en Ecuador (Ochoa-Rosales, 2007). La creciente literatura de investigación suministró bases sólidas para el desarrollo de procedimientos de evaluación de riesgo basados en evidencia, de manera tal que los evaluadores disponen actualmente de sistemáticas para coleccionar, balancear y combinar la información acorde a la mejor práctica profesional (Hart & Logan, 2011).

En lo específico de la formulación pericial del riesgo de violencia diversos hitos han marcado una historia de perfeccionamiento. Hubo tempranos aportes relacionados con la relación entre factores individuales y ecológicos (Mulvey & Lidz, 1984; Silver et al., 1999) y con la necesidad de contemplar factores históricos, clínicos y medioambientales (Webster et al., 1995; Webster et al., 1997) y de integrar evidencias normotéticas y particularidades idiosincráticas a las condiciones medioambientales (Folino, 2004b, 2004a). Finalmente, las últimas versiones de guías de evaluación basadas en el juicio profesional estructurado ofrecieron modos explícitos para la formulación del caso sobre riesgo de violencia (Douglas et al., 2013; Kropp & Hart, 2015, 2021). La formulación del caso está actualmente reconocida como una práctica crítica para las tareas de evaluación, asesoramiento y prevención del riesgo de violencia (Ogloff & Davis, 2020).

## Entrenamiento profesional sobre formulación pericial

El entrenamiento profesional relacionado con las evaluaciones periciales en salud mental ha sido tradicionalmente preocupación y foco de propuestas para las asociaciones científicas internacionales (Wettstein, 2005) e, incluso, para las instituciones académicas latinoamericanas (Folino & Raverta, 2006). Aunque en las dos últimas décadas hubo aumento de las oportunidades de entrenamiento, hay hallazgos latinoamericanos que indican la necesidad de mejorar la calidad del proceso de evaluación y de elaboración de los informes periciales (Beleño, 2016).

El interés en acondicionar habilidades para la valoración racional de los peritajes por parte de los operadores de justicia también ha crecido últimamente, pero aún mantiene un predominante estatus de aspiración. Gascón Abellán (2016) sostiene que la valoración que puede llevar a cabo el juez sobre una prueba científica dependerá si éste es capaz de entender cómo se han alcanzado los datos, es decir qué técnica se utilizó, cual es la validez científica de ésta, los análisis efectuados y cómo ha sido la cadena de custodia y, en consecuencia, qué fiabilidad posee la prueba presentada. También alerta que los vacíos cognoscitivos pueden llevar a error de parte del juez si las pruebas no se presentan en los términos apropiados y con las características anteriormente indicadas. Si bien el autor destaca que los jueces carecen de la formación necesaria para entender la base cognoscitiva de las pruebas que deben valorar, no ofrece evidencias relacionadas con la extensión de esa formación.

En el ámbito correccional, los entrenamientos focalizaron en la formulación del caso bajo el supuesto que la capacitación mejoraría la comprensión de los trastornos de personalidad y la gestión de las personas liberadas bajo condiciones (Minoudis et al., 2013; Radcliffe et al., 2018). Los entrenamientos a oficiales de gestión de casos de liberados duraron entre ocho horas y seis días; algunos tuvieron supervisión subsiguiente y utilizaron formulaciones de casos reales, mientras otros utilizaron viñetas ficticias (Wheable & Davies, 2020). Los hallazgos de esos estudios sobre entrenamiento fueron mixtos y para Wheable & Davies (2020) sería necesario suministrar una base firme de conocimientos a los oficiales a cargo de la gestión de personas liberadas bajo condiciones.

En la limitada producción de estudios educativos sobre evaluación de riesgo de violencia con medidas de su eficacia se destaca el realizado por McNiel et al. (2008), pero a diferencia del presente estudio que presenta una medida sofisticada de la calidad de la formulación, en aquel los autores utilizaron un sencillo índice de apreciación global de la calidad con rango 1 (pobre) a 7 (excelente). Consideramos que la actual disponibilidad de una medida como la TEC-F (Folino et al., 2022; Folino, 2022) facilitará la apreciación más minuciosa y sistemática del impacto de las intervenciones pedagógicas en el aprendizaje de formulación pericial.

## La valoración de la calidad de la formulación con la TEC-F

En los últimos años ha habido un creciente interés en el campo forense por la formulación de casos pues es considerada una competencia central en la práctica

profesional (Hart & Logan, 2011; Wheable & Davies, 2020). La capacitación sobre la formulación es un camino para el avance en la investigación y en los servicios forenses, aunque aún, lamentablemente, no haya tenido la necesaria propagación. Actualmente se dispone de la Guía para la Valoración y Formulación del Caso en Salud Mental Forense – TEC-F (Folino, 2022), que orienta sistematizadamente a quien quiera examinar y valorar la calidad de la formulación del caso en un informe pericial relacionado con temas de salud mental y del comportamiento, sin perjuicio de que podría servir de marco de referencia más allá de esta materia. La guía ofrece también una avenida plena de sugerencias para que sea recorrida por quien construye su informe pericial o quiere revisarlo en búsqueda de debilidades susceptibles de fortalecimiento.

La TEC-F se utiliza como instrumento para valorar la calidad de la formulación del caso forense atendiendo a las dimensiones de transparencia, especificidad, comunicación y fundamentación. Si bien exige respetar sus definiciones operacionales expresadas en diversos criterios, permite el ejercicio del juicio profesional al estimular el descubrimiento de indicadores de calidad, además de orientar la verificación de otros indicadores sugeridos. Mayores detalles se describen en la sección de método, pero en este apartado destacamos la rápida asimilación que tuvo el instrumento en la comunidad forense. Desde la publicación de las primeras evidencias empíricas (Folino et al., 2022) se produjo rápida demanda de capacitación (Asociación Nacional de Consejeros/as Técnicos/as del Poder Judicial -ANCOT-, 2022; Poder Judicial República de Chile, 2022) y el desarrollo de nuevas investigaciones además de la que estamos informando en el presente artículo, precisamente con Consejeros Técnicos de los Tribunales de Familia chilenos; y con Jueces y funcionarios judiciales de Chile, Argentina y Ecuador.

## El presente estudio

La información pertinente a la interfaz del derecho y la salud mental que hemos comentado arriba resalta la relevancia que tienen los programas de entrenamiento específicos sobre formulación y sobre riesgo de violencia y, especialmente, aquellos que permiten apreciar el impacto formativo. Asimismo, esa información pone de manifiesto la utilidad que tiene un instrumento como la TEC-F que guía el proceso de formulación y la medición de su calidad. En este estudio nos propusimos medir el impacto de un entrenamiento sobre formulación pericial del caso forense en general, y sobre riesgo de violencia en particular, y

extender las evidencias relacionadas con la confiabilidad de la TEC-F.

El intento tiene pocos antecedentes en América Latina y está llamado a construir conocimiento en áreas especialmente sensibles de las prácticas periciales y, consiguientemente, al ejercicio de la justicia. Tiene, además, un particular desafío que es la involucración de participantes de la psicología y de la psiquiatría provenientes de diversas tradiciones y de diversos países latinoamericanos. Ese aspecto es relevante pues hay antecedentes que alertan sobre influencias que pueden depender de algunas características de los evaluadores (de Vogel & de Ruitter, 2004), del contexto judicial (Miller et al., 2012), de las tradiciones profesionales (Folino et al., 2022) y de la experiencia y tipos de factores de riesgos evaluados (Penney et al., 2014). El presente estudio, al incorporar comparaciones entre profesionales según el procedimiento que se describe en la sección siguiente, aporta información pertinente a actuales prácticas interdisciplinarias que en diversos países están establecidas por ley o resoluciones de carácter reglamentario (i.e. Congreso de la Nación Argentina, 2010; Servicio Médico Legal - Ministerio de Justicia - República de Chile, 2021).

## Material y método

### Diseño y participantes

La investigación tuvo el diseño de cuasi-experimento pues se realizaron medidas pre y post intervención por investigadores independientes y a ciegas y se manipuló la asignación de determinadas condiciones aleatoriamente. Los participantes fueron quince profesionales de la psiquiatría (nueve) y de la psicología (seis) con domicilios laborales en diversos países latinoamericanos (Argentina, Bolivia, Chile, Colombia, Ecuador, Perú y Uruguay) y con experiencia en trabajo pericial forense.

Los evaluadores independientes fueron una profesional de la psiquiatría y otra de la psicología con experiencia en actividad pericial y estudios de postgrado relacionados con salud mental forense.

El grupo formó parte de la actividad académica de postgrado que fue dirigida por el último autor y diseñada para el entrenamiento avanzado en formulación del caso forense desde la perspectiva de la salud mental.

### Viñetas

Se utilizaron dos viñetas ficticias como bases para las formulaciones periciales. Las viñetas fueron originalmente utilizadas en previa investigación en el Reino Unido, y construidas con información del tipo de la que es encontrada en los informes psicológicos y

psiquiátricos del Offender Assessment System (Minoudis et al., 2013). Ambas fueron suministradas por Philip Minoudis al último autor del presente artículo juntamente con la autorización para su uso en la investigación. Una de las viñetas representaba el caso de alguien con severos rasgos de personalidad antisocial y la otra, el caso de alguien con rasgos de trastorno borderline de la personalidad.

### Instrumento de medición

Para valorar la calidad de las formulaciones se utilizó la TEC-F que orienta el procedimiento contemplando las dimensiones denominadas *Transparencia*, *Especificidad*, *Comunicación* y *Fundamentación* (Folino, 2022). La TEC-F cuenta con extensa descripción de las dimensiones y con detallada definición de quince criterios diseñados para ser contrastados con la formulación del caso y, subsiguientemente, valorar el grado de calidad con el que están satisfechos. Asimismo, expone diversos indicadores que ejemplifican detalles técnicos en los que se puede apreciar el grado de calidad. La guía suministra un minucioso manual explicativo que intenta superar los problemas emergentes de escuetas definiciones operativas de previos instrumentos (McMurrin & Bruford, 2016) y que debe ser estudiado para su correcto uso. Las descripciones abreviadas de cada dimensión son las siguientes:

- Transparencia*: la formulación requiere ser pasible de examen; debe presentar información suficiente de tal manera que pueda ser puesta a prueba por las partes;
- Especificidad*: la formulación está destinada a satisfacer una necesidad judicial determinada por quien requiere el servicio. Responde a puntos de pericia concretos;
- Comunicación*: la formulación debe transmitir un mensaje eficaz, que llegue con claridad y significación al receptor extra disciplinar y
- Fundamentación*: la formulación requiere estar sostenida en un razonamiento con bases fácticas y argumentos técnicos, propios del conocimiento científico aplicado. La lista de los criterios consta en la *Tabla 1*.

Los criterios se miden con una escala de tipo Likert con 4 categorías ordenadas. Las instrucciones indican que el evaluador seleccione en qué grado la formulación del caso revisada satisface cada criterio. Las opciones de elección son las siguientes: 0) No lo satisface en absoluto; 1) Lo satisface poco; 2) Lo satisface bastante; 3) Lo satisface completamente. Los indicadores, que no se exponen en el anexo, orientan de manera concreta a la verificación de los constructos y a su valoración. Esos indicadores no son exhaustivos y dejan oportunidad para que el evaluador aplique su juicio profesional si verificare algún otro.

**Tabla 1.** Lista de criterios de la TEC-F

Dimensión	Criterio
<b>Transparencia</b>	A. Expone las fuentes y la información utilizadas posibilitando su verificación
	B. Transmite los suficientes detalles que hacen posible el examen del procedimiento
	C. Explica los cuidados éticos de la práctica
<b>Especificidad</b>	A. Está focalizada en el requerimiento
	B. Ofrece una respuesta útil para la toma de decisión judicial
	C. Entrega conclusión que tiene la potencialidad de ser contrastable en el futuro
<b>Comunicación</b>	A. Entrega una narración ordenada y comprensible
	B. Está libre de detalles innecesarios
	C. Está expresada con un lenguaje natural e incluye explicación de los términos técnicos que pudieran ser ineludibles
	D. Adiciona estrategia para facilitar la interpretación
<b>Fundamentación</b>	A. Analiza las piezas de información pertinentes y necesarias
	B. Expone los argumentos con coherencia y sin contradicciones
	C. Respeta el principio de temporalidad, promoviendo la comprensión según la secuencia temporal de pasado, presente y futuro
	D. Aplica teorías o modelos con sostén empírico
	E. Expresa conclusiones de manera congruente con la argumentación y con el grado de certeza reconocido

Dado que la formulación pericial es una práctica profesional sumamente compleja que, dependiendo de múltiples circunstancias concretas, se presenta con amplias variaciones en su configuración y que, por otra parte, previos estudios empíricos no pudieron superar las dificultades de una ilustración exhaustiva (Wheable & Davies, 2020), la TEC-F no está diseñada con la pretensión de alcanzar propiedades métricas rigurosas que posibiliten conformar normas o baremos. Sin embargo, diversos aspectos de su confiabilidad han sido revisados empíricamente en difícil prueba con profesionales y formulaciones provenientes de diferentes disciplinas y diversas tradiciones latinoamericanas. Los hallazgos principales de esa investigación fueron valores de congruencia interna con un rango 0,34-0,81 para las dimensiones y con valor de 0,85 para el total de la TEC-F y, por otra parte, valores de coeficiente de correlación intraclase de acuerdo absoluto entre evaluadores con similar tradición de 0,72; 0,52; 0,76 y 0,77 para las respectivas dimensiones, y de 0,84 para el total (Folino et al., 2022).

### Procedimiento

Las viñetas fueron distribuidas al grupo de participantes de manera aleatoria en el momento previo a la intervención pedagógica, y distribuidas de manera inversa posteriormente a la intervención. La consigna suministrada fue realizar una formulación de riesgo de recidiva violenta, considerando que los sujetos descriptos en la viñeta estaban cumpliendo condena privados de libertad y en tiempo de acceder a alguna forma de liberación anticipada.

La intervención didáctica consistió en cinco encuentros de entrenamiento sincrónico a través de plataforma de videoconferencia y en actividad práctica domiciliaria. El peso horario total fue de 12 h. Los contenidos didácticos abarcaron temas de formulación del caso forense en general y de riesgo de violencia, en particular. La intervención se realizó desde junio a julio de 2022.

Las formulaciones periciales pre y post intervención didácticas basadas en las viñetas fueron anonimadas de manera que no quedaran indicadores de sus autores y que pudieran ser valoradas a ciegas por dos investigadores independientes. Posteriormente a las valoraciones independientes, los investigadores hicieron valoraciones de consenso y, sobre una muestra aleatoria de 10 casos, hicieron nuevas valoraciones independientes de para estimar la estabilidad test-retest. Las puntuaciones individuales fueron cargadas en formulario Google con medidas de seguridad y luego integradas en base de dato para su análisis.

### Plan de análisis

La calidad de las formulaciones consensuadas fue descripta mediante frecuencias y estadísticos de tendencia central y dispersión. Asimismo, se calcularon el coeficiente Alfa de Cronbach y las correlaciones ítem e ítem-total.

La confiabilidad pertinente al acuerdo entre evaluadores fue puesta a prueba con el cálculo de los coeficientes de correlación intraclase con modelo de efectos aleatorios de dos factores para los tipos A,1 (acuerdo absoluto de medidas únicas) y C,1 (consistencia de medidas únicas). La confiabilidad test-retest se estimó calculando correlaciones y coeficientes de correlación intraclase.

Las medias pre y post intervención fueron comparadas con la prueba de *t* computada grupalmente y con la prueba de Wilcoxon de manera apareada, incluyendo la estratificación por profesión. Se consideraron estadísticamente significativos los resultados con  $p < 0,05$  y altamente significativos los resultados con  $p < 0,001$ .

### Resultados

Inicialmente se exponen los resultados pertinentes a la descripción de calidad de las formulaciones periciales y a la confiabilidad de la TEC-F. Posteriormente se exponen los resultados comparativos de la calidad pre y post intervención pedagógica.

#### Descriptivos y confiabilidad de la TEC-F

Los estadísticos descriptivos fueron calculados con las valoraciones consensuadas (ver Tabla 2). De los hallazgos se destacan los desvíos estándar amplios, indicadores de variada calidad que resulta importante a los efectos de estimar la confiabilidad inter evaluadores. Los coeficientes de correlación intraclase se calcularon en base al modelo de efectos aleatorios de dos factores; se calculó el tipo A1 que es apropiado para diseños que utilizan una muestra de evaluadores seleccionada de una población de potenciales evaluadores y en los que cada evaluador puntúa los mismos casos (Fleiss, 1986). Con la excepción del valor obtenido para el criterio F\_C, todos los valores de los coeficientes correspondientes a los criterios individuales se ubicaron entre las categorías de acuerdos bueno ( $0,60 < CCI < 0,75$ ) y excelente ( $\geq 0,75$ ). El  $CCI_{C,1}$ , que no considera las diferencias constantes como discordancia, tuvo valores prácticamente equivalentes. Ambos tipos de coeficientes para los totales de las dimensiones y para el total de la TEC-F se ubicaron en la categoría de excelente.

Las correlaciones ítem/total indican buena consistencia interna y que cada uno contribuye significativa-

**Tabla 2.** Estadísticos descriptivos de valoraciones consensuadas, correlación ítem-total y confiabilidad entre evaluadores para criterios y totales TEC-F

Criterio/escala	Frecuencia puntajes TEC-F				Descriptivos			Dos evaluadores	
	0	1	2	3	Media	Desvío estándar	r ítem/total	CCI <sub>A,1</sub>	CCI <sub>C,1</sub>
T_A	0	3	6	21	2,60	0,675	0,756**	0,937	0,935
T_B	1	3	5	21	2,53	0,819	0,889**	0,927	0,929
T_C	2	7	10	11	2,00	0,947	0,718**	0,724	0,733
E_A	1	4	5	20	2,47	0,860	0,872**	0,850	0,848
E_B	1	5	4	20	2,43	0,898	0,871**	0,881	0,878
E_C	2	4	7	17	2,30	0,952	0,824**	0,865	0,861
C_A	0	0	6	24	2,80	0,407	0,779**	0,659	0,665
C_B	0	0	5	25	2,83	0,379	0,500**	0,813	0,818
C_C	0	2	1	27	2,83	0,531	0,486**	0,627	0,634
C_D	4	7	12	7	1,73	0,980	0,648**	0,696	0,688
F_A	0	3	3	24	2,70	0,651	0,786**	0,891	0,887
F_B	0	2	4	24	2,73	0,583	0,667**	0,889	0,887
F_C	0	1	2	27	2,87	0,434	0,458*	0,261	0,273
F_D	5	4	5	16	2,07	1,172	0,698**	0,888	0,887
F_E	2	3	8	17	2,33	0,922	0,837**	0,884	0,882
T_total	-	-	-	-	7,13	2,129	-	0,924	0,930
E_total	-	-	-	-	7,20	2,565	-	0,944	0,943
C_total	-	-	-	-	10,20	1,710	-	0,827	0,823
F_total	-	-	-	-	10,37	2,125	-	0,926	0,924
TEC-F_total	-	-	-	-	34,90	7,554	-	0,948	0,947

Nota: N = 30. r: correlación de Pearson. \*: significativa en el nivel 0,01 bilateral. \*\*: significativa en el nivel 0,001 bilateral. CCI: Coeficiente de correlación intraclase modelo de efectos aleatorios de dos factores. A, 1: CCI tipo acuerdo absoluto de medidas únicas. C, 1: CCI tipo consistencia de medidas únicas. Las siglas de cada criterio se conforman con la primera letra de cada dimensión (T de Transparencia; E de Especificidad; C de Comunicación y F de Fundamentación- y con letra de secuencia alfabética).

mente con la puntuación total; todos los valores fueron positivos, altos y estadísticamente significativos (Tabla 2). La menor correlación encontrada fue la del criterio F\_C con el total y su valor fue 0,45 ( $p < 0,05$ ). La media de las correlaciones inter ítem fue 0,49, resultado que está bien por encima del valor de 0,20 que es el umbral recomendado para considerar que una escala es homogénea (Green, Lissitz, & Mulaik, 1977). Los valores del coeficiente Alfa de Cronbach obtenidos fueron 0,83; 0,94; 0,63; 0,77 y 0,93 para las dimensiones Transparencia, Especificidad, Comunicación y Fundamentación y para la TEC-F total respectivamente.

### Test-retest

Se seleccionaron al azar 10 casos que fueron reevaluados por los dos investigadores independientes luego de un período de 19 días en promedio. Los valores de las correlaciones en las valoraciones de los dos evaluadores conjuntamente fueron 1; 0,98; 0,93; 0,95 y

0,99, para las dimensiones T, E, C y F y para el total TEC-F respectivamente (todas  $p < 0,01$  bilateral). Similares excelentes indicadores se obtuvieron en las correlaciones de las valoraciones individuales de cada evaluador; baste con mencionar que la correlación test-retest para el total de la TEC-F fue 0,98 para cada evaluador. Asimismo, se calculó el CCI<sub>A,1</sub> con modelo aleatorio de dos factores obteniéndose como resultado 0,99 y 0,81 para las valoraciones del total TEC-F del evaluador A y del evaluador B, respectivamente. Con las valoraciones conjuntas, los resultados obtenidos para test-retest del total de la TEC-F fue 0,98 tanto para el CCI<sub>A,1</sub> como para el CCI<sub>C,1</sub>.

### Valoración de las formulaciones con la TEC-F y grupos profesionales

Contemplando la hipótesis de que las profesiones de psiquiatría y de psicología constituyeran factores con influencias diferenciales en las valoraciones de calidad o en la capitalización de la intervención, com-

paramos las medias de calidad estratificando por profesión, para todas las formulaciones, para las previas a la intervención y para las posteriores. Asimismo, comparamos las medias pre y post intervención para cada uno de los grupos con y sin apareamiento.

Al comparar todas las formulaciones de cada grupo profesional, encontramos que ambas medias fueron prácticamente idénticas y sólo tuvieron una diferencia no significativa de apenas 3 centésimos ( $p = 0,98$ ). Tampoco hubo diferencias significativas entre psicólogos y psiquiatras en las medias pre intervención ni en las medias post intervención, consideradas grupalmente.

En cambio, al comparar las medias pre y post intervención separadamente para psicólogos y psiquiatras, emergieron algunas pequeñas diferencias en el patrón de capitalización de la experiencia. Grupalmente, entre los psicólogos, la mejoría más significativa se produjo en la dimensión de comunicación y entre los psiquiatras, en las dimensiones de comunicación y fundamentación y en el total TEC-F. Los resultados de la comparación más pertinente, que es la apareada, puso de manifiesto que mientras entre los psicólogos las mejorías más significativas se produjeron en las dimensiones de transparencia y comunicación ( $p = 0,03$  para ambas), entre los psiquiatras se produjo en comunicación ( $p = 0,03$ ).

### Valoración y cantidad de palabras

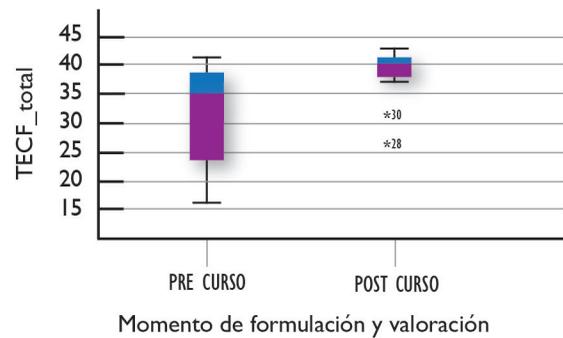
Partiendo de la hipótesis que la extensión de los informes se relacionara con la calidad de alguna de las dimensiones o de la formulación en general, exploramos la cantidad de palabras y las respectivas correlaciones. La media de palabras de los 30 informes fue 1918 (DE 1211,2; mínimo 818; máximo 5847). Las correlaciones entre la cantidad de palabras y la valoración, si bien fueron positivas no fueron ni fuertes ni estadísticamente significativas. Los valores fueron 0,34; 0,32; 0,27; 0,29 y 0,35 para las dimensiones T, E, C, F y para el total TEC-F respectivamente, todos con una  $p$  bilateral mayor que 0,05.

También exploramos las diferencias entre medias de cantidad de palabras para los grupos de psicólogos y de psiquiatras, para la viñeta 1 y la viñeta 2 y para los informes pre intervención y post intervención. Ninguna de las diferencias fue estadísticamente significativa.

### Resultados pre y post intervención de entrenamiento

Los resultados de la intervención pedagógica indicaron que los participantes alcanzaron una relevante mejora en la calidad de la formulación, tanto a nivel grupal como a nivel individual. La media de los valores totales TEC-F pre-curso fue de 31,4 y la media post curso fue

**Gráfico 1.** Distribución de valores TEC-F totales pre y post intervención pedagógica



de 38,4; la diferencia entre las medias fue significativa tanto en la prueba de grupos independientes ( $p = 0,003$ ) como en la prueba apareada ( $p = 0,001$ ). En el *Gráfico 1* se ilustra la amplia diferencia en las distribuciones pre y post intervención que se pone de manifiesto no sólo en el aumento del valor medio sino, también, en la disminución de la dispersión.

### Discusión

La meta de este estudio fue medir el impacto de un entrenamiento sobre formulación pericial del caso forense desde la perspectiva de la salud mental y extender las evidencias relacionadas con la confiabilidad de la TEC-F. El estudio se desarrolló con un diseño de cuasi-experimento, con mediciones de la calidad pericial previas y posteriores a la intervención pedagógica y con cuidados para evitar sesgos anonimizando las formulaciones, asignando viñetas equivalentes y alternadas a los participantes y produciendo mediciones independientes y consensuadas. Los resultados pusieron de manifiesto una significativa mejora en la calidad de las formulaciones periciales posteriores a la intervención y permitieron sostener diversos indicadores de la confiabilidad de la TEC-F. A continuación, profundizamos en la discusión iniciando por la confiabilidad del instrumento.

Los instrumentos que intentan medir fenómenos complejos o que intentan guiar el juicio profesional estructurado requieren de acumulación de evidencias para que puedan ser considerados confiables. Mientras que previo estudio había ofrecido evidencias germinales en cuanto a la confiabilidad de la TEC-F (Folino et al., 2022), en el presente estudio encontramos que a nivel de los criterios, los indicadores de acuerdo inter evaluadores fueron, en general, entre buenos y excelentes. La excepción fue el hallazgo de un bajo acuerdo para el criterio C de la dimensión Comunicación. Es posible que ese criterio presente alguna

vulnerabilidad especial a la influencia de diferentes profesiones y tradiciones, pues este estudio, al igual que en parte de las valoraciones realizadas en el de Folino et al. (2022), los evaluadores independientes no sólo eran de profesiones diferentes sino, también, de países diferentes. Si bien mayores consideraciones se podrán hacer a partir de nuevos estudios con muestras que permitan análisis factorial, mientras tanto, cabe resaltar que ese bajo acuerdo a nivel de uno de los criterios se mejora notablemente cuando se considera a la dimensión comunicación en su totalidad. Más destacable aún es que el acuerdo se ubicó en la categoría de excelente para los totales de todas las dimensiones y de la TEC-F. En la investigación también se obtuvieron positivos resultados relacionados con la consistencia interna y con la homogeneidad.

El presente estudio ofrece resultados indicadores de la estabilidad de la TEC-F. El instrumento hasta ahora no había sido sometido a la prueba de test-retest y en esta oportunidad se hizo midiendo la correlación individual y conjunta de las valoraciones y calculando dos tipos de coeficientes de correlación intraclase entre el test y el retest. Todos los hallazgos indicaron una excelente confiabilidad. Por otra parte, cabe destacar que el plazo intermedio entre la primera y la segunda valoración fue de casi 3 semanas, a diferencia de previo estudio con otro instrumento de medición de calidad que utilizó una semana (Minoudis et al., 2013).

En síntesis, esta investigación puso a prueba de manera sumamente exigente a la TEC-F porque tanto los evaluadores como los participantes tuvieron muy diferentes antecedentes y porque se utilizó un período prolongado entre test y retest. Por lo tanto, las evidencias que adicionan al sostén de la confiabilidad del instrumento pueden ser consideradas robustas.

La exploración de la relación entre cantidad de palabras y la valoración de calidad correspondientes a cada dimensión y al total de la TEC-F produjo resultados consistentes con previas investigaciones. Todo ello a pesar de que la muestra de formulaciones tuvo amplio desvío estándar y amplio rango tanto en la cantidad de palabras como en la valoración de la calidad, lo que constituía una oportunidad sensible para captar la eventual relación, si hubiera existido. Al igual que informaron Folino et al. (2022), con el mismo instrumento, en el presente estudio no se obtuvieron correlaciones significativas con la cantidad de palabras. El resultado también fue homólogo al que obtuvieran Hopton et al. (2018) con otro instrumento. Por lo tanto, los presentes hallazgos ratifican que la formulación pe-

ricial no es pasible de prescripciones estrictas en cuanto a su extensión y que, de alguna manera, refleja una resolución individualizada del caso en términos de la tensión entre complejidad y simplicidad (Eells, 2010). Estas consideraciones pueden ser relevantes para la práctica profesional pues orientan a buscar el apropiado balance entre ambas cualidades y a evitar, al mismo tiempo, omitir aspectos importantes del caso e interferir en su función comunicacional (Folino, 2022).

La discusión del aporte que este estudio hace a la educación en temas de salud mental forense y, específicamente, a la evaluación de riesgo de violencia, se enmarca en un campo de la pedagogía especializada plétórico de necesidades. Tempranamente se dudó que los programas de entrenamiento dispusieran del tiempo suficiente para cubrir las dimensiones implícitas en la elaboración y comunicación de un informe pericial (Griffith & Baranoski, 2007). Para la misma época también se llamaba la atención sobre las carencias que tenía la formación de los residentes de psiquiatría en evaluación de riesgo de violencia (Oldham, 2007) y sobre la disparidad de los servicios educativos disponibles en varios contextos internacionales (Folino & Pezzotti, 2008). Si bien en América Latina se contaba con algunos foros formativos específicos (Folino & Raverta, 2006), la difusión e implementación de estrategias de evaluación de riesgo de violencia basadas en evidencia seguía siendo dispar, lenta y dificultosa varios años después (Folino, Villaba, & Singh, 2016; Rettenberger et al., 2017; Singh et al., 2014; Singh, Condemarin, & Folino, 2013). Que el presente estudio haya vehiculizado el entrenamiento profesional de una manera altamente estructurada, involucrando a los participantes en un diseño cuasi-experimental y que haya podido producir y medir un gran impacto formativo es una muy buena noticia para la interfaz de la salud mental y el derecho. El impacto formativo quedó puesto de manifiesto en la significativa diferencia de la calidad registrada antes y después de la intervención pedagógica. Los hallazgos cuantitativos no sólo reflejaron la mejoría de calidad indicada por las medias, sino, también reflejaron una notable disminución en la dispersión de las valoraciones de la calidad. En la sección de resultados, ambas mejoras pueden ser fácilmente visualizadas en los gráficos de caja focalizando en las significativas diferencias de las medias, de las cajas y de los bigotes (Gráfico 1). Nuestros hallazgos suman evidencias a previos estudios que sostuvieron la utilidad de los entrenamientos específicos para mejorar la calidad de las formulaciones (Eells et al., 2015; Flinn et al., 2015) y, considerando el diseño

utilizado, estimula la generación de planificaciones que combinen entrenamiento e investigación.

Por otra parte, la comparación de valoraciones y de mejorías de las formulaciones según grupos profesionales suministró evidencias de muy parecidas performances. Los profesionales de ambas profesiones lograron diferencias positivas entre el antes y el después de la intervención, no obstante, se detectó que las diferencias más significativas entre los psicólogos se produjeron en las dimensiones de transparencia y de comunicación, y entre los psiquiatras, en la dimensión de comunicación. Estos resultados minuciosos, que por el momento resulta razonable considerarlos preliminares, orientan una heurística interesante: la acumulación de evidencias de este tipo podría configurar necesidades disciplinares diferenciales y, consiguientemente, la posibilidad de planificaciones curriculares más ajustadas.

El procedimiento de cuantificación de la calidad no impidió que los investigadores encargados de la valoración pudieran detectar algunas particularidades cualitativamente. Como apreciación cualitativa positiva se destacó que los profesionales aplicaron los aprendizajes del curso aún manteniendo características reconocibles de un estilo propio. La apreciación resulta interesante pues capta un efecto personalizado que tuvo la intervención didáctica y la adaptabilidad del funcionamiento de la TEC-F como medida de la calidad. La interpretación que hacemos del hallazgo es que, tanto en la planificación de la intervención como en el instrumento de medición, subyacen constructos esenciales, adaptables a eventuales peculiaridades de las evaluaciones de casos y a estilos individuales de comunicación de la formulación. Como apreciación negativa, se destacó la debilidad de la comunicación de cuidados éticos de la práctica y la escasa utilización de exposiciones que atendieran a lo estratégico y didáctico; ambos hallazgos cualitativos resultan completamente consistentes con los resúmenes cuantitativos de los correspondientes criterios. El impacto formativo que deriva de la utilización de la TEC-F responde a la necesidad de generar métodos que mejoren la confiabilidad y eviten sesgos profesionales (Guarnera et al., 2017).

El estudio tuvo la debilidad de no haber utilizado una medida de la satisfacción de los participantes para con el curso. Si bien la apreciación cualitativa de lo expresado al cierre del mismo dio cuenta de generalizada satisfacción, su apropiada cuantificación hubiera permitido explorar su relación con el rendimiento individual, entre otras variables.

Entre las fortalezas contamos los cuidados metodológicos, tales como las valoraciones independientes a ciegas y el uso cruzado de viñetas estándar.

Finalmente, más allá de los aportes específicos que el estudio pudo ofrecer y del intento de responder a la expectativa de promover mejores servicios periciales (Folino et al., 2022), queremos expresar nuestra aspiración a que el presente estudio también ilustre el beneficio que conlleva la integración de formación e investigación (Bloom, 2007; Wiese et al., 2010).

**Conflicto de intereses:** Jorge O. Folino cobra derechos de autor y Arturo Klenner es uno de los autores contribuyentes en la edición de la TEC-F Guía para la valoración y formulación del caso en salud mental forense: Transparencia; Especificidad; Comunicación y Fundamentación.

**Agradecimiento:** Los autores agradecen a los siguientes profesionales por su responsabilidad al ejecutar las formulaciones durante el período de estudio: Águila Zúñiga, Alejandro R.; Ambroggio, Daniel R.; Barrachina, Alicia; Benítez, Julián M.; da Silva, Sandra; Daneck Muñoz, María C.; Follado, Ana C.; Guillén, Ximena; Hinojosa, González Valeska A.; Huber Pérez, Anne M.; Julca Belloda, Rossana M.; Mendoza Arango, Catalina; Porteiro, María F.; Vega, Raúl; Vila, Atenas. Asimismo, a Philip Minoudis por compartir material de su investigación.

## Referencias bibliográficas

- Asociación Nacional de Consejeros/as Técnicos/as del Poder Judicial -AN-COT-. (2022). Consejeras y Consejeros Técnicos se Capacitan en Uso de Guía de Salud Mental Forense. Retrieved October 7, 2022, <https://ne-np.facebook.com/296526407189552/posts/consejeras-y-consejeros-técnicos-se-capacitan-en-uso-de-guia-de-salud-mental-for/2271446893030817/>
- Beleño, M. (2016). *Análisis de la calidad de las pericias psicológicas realizadas a niños, niñas y adolescentes en un Centro Zonal ICBF-Regional Cesar*. Tesis de Maestría. Facultad de Psicología. Universidad Santo Tomás.
- Bernstein, D., & Lasker, E. (2015). Defending Daubert: It's time to amend federal Rule of Evidence 702. *William and Mary Law Review*, 57(1), 1–48. [https://www.law.gmu.edu/assets/files/publications/working\\_papers/LS1529.pdf](https://www.law.gmu.edu/assets/files/publications/working_papers/LS1529.pdf)
- Bloom, J. D. (2007). Commentary: authorship and training in forensic psychiatry. *J Am Acad Psychiatry Law*, 35(1), 32–33.
- Brown, S., Beeley, C., Patel, G., & Völlm, B. (2018). Training probation officers in case formulation for personality disordered offenders. *Criminal Behaviour and Mental Health : CBMH*, 28(1), 50–60. <https://doi.org/10.1002/cbm.2006>
- Bucci, S., French, L., & Berry, K. (2016). Measures Assessing the Quality of Case Conceptualization: A Systematic Review. *Journal of Clinical Psychology*, 72(6), 517–533. <https://doi.org/10.1002/jclp.22280>
- Congreso de la Nación Argentina. Ley N° 26.657. Ley Nacional de Salud Mental. (2010). Argentina: Boletín Oficial de la Nación Argentina, 2010/12/03.
- de Borba Telles, L. E., Day, V. P., Folino, J. O., & Taborda, J. G. V. (2009). Reliability of the Brazilian version of HCR-20 Assessing Risk for Violence. *Revista Brasileira de Psiquiatria*, 31, 253–256.
- de Borba Telles, L. E., Folino, J. O., & Taborda, J. G. V. (2012). Accuracy of the Historical, Clinical and Risk Management Scale (HCR-20) in predicting violence and other offences in forensic psychiatric patients in Brazil. *International Journal of Law and Psychiatry*, 35, 427–431.

- de Vogel, V., & de Ruiter, C. (2004). Differences between clinicians and researchers in assessing risk of violence in forensic psychiatric patients. *Journal of Forensic Psychiatry and Psychology*, 15(1), 145–164. <https://doi.org/10.1080/14788940410001655916>
- Dixon, L., & Gill, B. (2002). Changes in the standards for admitting expert evidence in federal civil cases since the Daubert decision. *Psychology, Public Policy, and Law*, 8(3), 251–308. <https://doi.org/10.1037/1076-8971.8.3.251>
- Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). HCR-20 v3 Assessing Risk for Violence. Mental Health, Law, and Policy Institute, Simon Fraser University.
- Douglas, K. S., & Kropp, P. R. (2002). A Prevention-Based Paradigm for Violence Risk Assessment - Clinical and Research Applications. *Criminal Justice and Behavior*, 29(5), 617–658.
- Duce J., M. (2018). Una aproximación empírica al uso y prácticas de la prueba pericial en el proceso penal chileno a la luz de su impacto en los errores del sistema. *Política Criminal*, 13(25), 42–103. [http://www.politicacriminal.cl/Vol\\_13/n\\_25/Vol13N25A2.pdf](http://www.politicacriminal.cl/Vol_13/n_25/Vol13N25A2.pdf)
- Eells, T. D., Lombart, K. G., Kendjelic, E. M., Turner, L. C., & Lucas, C. P. (2005). The quality of psychotherapy case formulations: A comparison of expert, experienced, and novice cognitive-behavioral and psychodynamic therapists. *Journal of Consulting and Clinical Psychology*, 73(4), 579–589. <https://doi.org/10.1037/0022-006X.73.4.579>
- Escobar-Córdoba, F. (2006). *Riesgo para cometer homicidio en jóvenes bogotanos: Estudio multimétodo [Homicide risk in youth from Bogota: Multi-method study]*. Doctoral dissertation. Director Prof. J. O. Folino. Departamento de Postgrado Facultad de Ciencias Médicas; Universidad Nacional de La Plata.
- Escobar-Córdoba, F., Acero-González, A. R., & Folino, J. O. (2015). Homicidas juveniles en Bogotá, estudio de grupos focales. *Revista de La Facultad de Medicina*, 63(3), 389–398.
- Escobar-Córdoba, F., Acero-González, Á. R., & Folino, J. O. (2015). Young murderers in Bogotá, a focus groups study. *Revista Facultad de Medicina*, 63(3). <https://doi.org/10.15446/revfacmed.v63n3.50632>
- Fleiss, J. L. (1986). *The design and analysis of clinical experiments*. New York: Wiley.
- Flinn, L., Braham, L., & Das Nair, R. (2015). How reliable are case formulations? A systematic literature review. *British Journal of Clinical Psychology*, 54(3), 266–290. <https://doi.org/10.1111/bjc.12073>
- Folino, J. O. (2004a). La emergencia y el riesgo para terceros. En M. Suárez Richards (Ed.), *Psiquiatría en Medicina General* (pp. 367–378). Editorial Pólemos.
- Folino, J. O. (2004b). *Nuevos paradigmas en la evaluación de peligrosidad [New paradigms in dangerousness assessment]*. Doctrina Judicial - Editorial La Ley, XX, 871–876.
- Folino, J. O. (2022). *TEC-F Guía para la valoración y formulación del caso en salud mental forense: Transparencia; Especificidad; Comunicación y Fundamentación (eBook)*. La Plata: Amazon. Retrieved from [https://www.amazon.com/dp/B0B4ZDXY6W/ref=sr\\_1\\_2?crid=18X8CJS8T7P-FI&keywords=jorge+folino&qid=1656070232&s=books&prefix=jorge+folino%2Cstripbooks%2C84&sr=1-2&asin=B0B4ZDXY6W&revision-Id=9562beb2&format=1&depth=1](https://www.amazon.com/dp/B0B4ZDXY6W/ref=sr_1_2?crid=18X8CJS8T7P-FI&keywords=jorge+folino&qid=1656070232&s=books&prefix=jorge+folino%2Cstripbooks%2C84&sr=1-2&asin=B0B4ZDXY6W&revision-Id=9562beb2&format=1&depth=1)
- Folino, J. O., Castillo, J. L., De Borja Telles, L. E., Jiménez Jiménez, W. A., Klenner, A., León-Mayer, E. M., ... Velázquez Zilberberg, J. A. (2022). La calidad de la formulación del caso forense en salud mental. *Vertex Revista Argentina de Psiquiatría*, 33(156), 25–34. <https://doi.org/10.53680/vertex.v33i156.176>
- Folino, J. O., Marengo, C. M., Marchiano, S. E., & Ascazibar, M. (2004). The Risk Assessment Program and the Court of Penal Execution in the Province of Buenos Aires, Argentina. *International Journal of Offender Therapy and Comparative Criminology*, 48(1). <https://doi.org/10.1177/0306624X03257526>
- Folino, J. O., & Pezzotti, L. C. (2008). Education in forensic psychiatry. *Current Opinion in Psychiatry*, 21(5). <https://doi.org/10.1097/YCO.0b013e328305e46d>
- Folino, J. O., & Raverta, M. J. (2006). Forensic Psychiatry in Argentina: A Training Forum For Latin America. *American Academy of Psychiatry and the Law Newsletter*, 31, 28–29.
- Folino, J. O., Villaba, D. K., & Singh, J. P. (2016). Violence Risk Assessment in Latin America. In J. P. Singh, S. Bjorkly, & S. Fazel (Eds.), *International Perspectives on Violence Risk Assessment* (pp. 181–191). New York: Oxford University Press.
- Gascón Abellán, M. (2016). Conocimientos expertos y deferencia del juez (Apunte para la superación de un problema). *Doxa. Cuadernos de Filosofía Del Derecho*, 39, 347–365. <https://doi.org/10.14198/doxa2016.39.18>
- Gatowski, S., Dobbin, S., Richardson, J., Ginsburg, G., Merlino, M., & Dahir, V. (2001). Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-Daubert World. *Law and Human Behavior*, 25(5), 433–458. Retrieved from <https://link.springer.com/article/10.1023/A:1012899030937>
- Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of unidimensionality. *Educational and Psychological Measurement*, (37), 827–838.
- Greer, B., Taylor, R. W., Cella, M., Stott, R., & Wykes, T. (2020). The contribution of dynamic risk factors in predicting aggression: A systematic review including inpatient forensic and non-forensic mental health services. *Aggression and Violent Behavior*, 53(April), 101433. <https://doi.org/10.1016/j.avb.2020.101433>
- Griffith, E. E. H., & Baranoski, M. V. (2007). Commentary: The Place of Performative Writing in Forensic Psychiatry. *Bulletin of the American Academy of Psychiatry and the Law*, 35, 27–31.
- Groscup, J., Studebaker, C., Penrod, S., Huss, M., & O'Neil, K. (2002). The effects of Daubert on the admissibility of expert testimony in state and federal criminal cases. *Psychology, Public Policy, and Law*, 8(4), 339–372. <https://doi.org/10.1037/1076-8971.8.4.339>
- Guarnera, L. A., Murrie, D. C., & Boccaccini, M. T. (2017). Why do forensic experts disagree? Sources of unreliability and bias in forensic psychology evaluations. *Translational Issues in Psychological Science*, 3(2), 143–152. <https://doi.org/10.1037/tps0000114>
- Hand, L. (1900). *Historical and practical considerations regarding expert testimony* (Reprinted). Albany: Gale, Making of Modern Law.
- Hart, S. D., & Logan, C. (2011). Formulation of Violence Risk Using Evidence Based Assessment: The Structured Professional Judgment Approach. In P. Sturmey & M. McMurran (Eds.), *Forensic Case Formulation* (pp. 83–106). West Sussex, UK: John Wiley & Sons, Ltd.
- Hart, S., Sturmey, P., Logan, C., & McMurran, M. (2011). Forensic case formulation. *International Journal of Forensic Mental Health*, 10(2), 118–126. <https://doi.org/10.1080/14999013.2011.577137>
- Hopton, J., Cree, A., Thompson, S., Jones, R., & Jones, R. (2018). An Evaluation of the Quality of HCR-20 Risk Formulations: A Comparison between HCR-20 Version 2 and HCR-20 Version 3. *International Journal of Forensic Mental Health*, 17(2), 195–201. <https://doi.org/10.1080/14999013.2018.1460424>
- Jurs, A., & Devito, S. (2013). Et Tu, Plaintiffs? An Empirical Analysis of Daubert's Effect on Plaintiffs, and Why Gatekeeping Standards Matter (a Lot). *Arkansas Law Review*, 66, 975–1005. [https://heinonline.org/hol-cgi-bin/get\\_pdf.cgi?handle=hein.journals/arklr66&section=39](https://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/arklr66&section=39)
- Kiely, T. (2001). *Forensic evidence: science and the criminal law*. Florida: CRC Press.
- Kropp, P. R., & Hart, S. D. (2015). *The Spousal Assault Risk Assessment Guide version 3 (SARA-V3)*. ProActive Resolutions Inc.
- Kropp, P. R., & Hart, S. D. (2021). *SARA V3- Guía de Valoración del Riesgo de Violencia Contra la Pareja - Versión 3* (1ra ed.; K. Arbach-Lucioni, I. Loinaz, & A. Andrés Pueyo, Eds.). Editorial Brujas.
- León-Mayer, E., Cortés Olavarría, M. S., & Folino, J. O. (2014). Descripción multidimensional de población carcelaria chilena. *Psicoperspectivas*, 13, 68–81.

- McMurrin, M., & Bruford, S. (2016). Case formulation quality checklist: a revision based upon clinicians' views. *Journal of Forensic Practice*, 18(1), 31–38. <https://doi.org/10.1108/JFP-05-2015-0027>
- McNeil, D. E., Chamberlain, J. R., Weaver, C. M., Hall, S. E., Fordwood, S. R., & Binder, R. L. (2008). Impact of clinical training on violence risk assessment. *Am J Psychiatry*, 165(2), 195–200. <https://doi.org/10.1176/appi.ajp.2007.06081396>
- Miller, C. S., Kimonis, E. R., Otto, R. K., Kline, S. M., & Wasserman, A. L. (2012). Reliability of risk assessment measures used in sexually violent predator proceedings. *Psychological Assessment*, 24(4), 944–953. <https://doi.org/10.1037/a0028411>
- Minoudis, P., Craissati, J., Shaw, J., McMurrin, M., Freestone, M., Chuan, S. J., & Leonard, A. (2013). An evaluation of case formulation training and consultation with probation officers. *Criminal Behaviour and Mental Health*, 23, 252–262. <https://doi.org/10.1002/cbm.1890>
- Mnookin, J. (2008). Expert Evidence, Partisanship, and Epistemic Competence. *Brook. L. Rev*, 73(3), 1009–1033. [https://heinonline.org/hol/cgi-bin/get\\_pdf.cgi?handle=hein.journals/brklr73&section=30](https://heinonline.org/hol/cgi-bin/get_pdf.cgi?handle=hein.journals/brklr73&section=30)
- Monahan, J., & Steadman, H. (1994). *Violence and Mental Disorder - Developments in Risk Assessment*. Chicago: The University of Chicago Press.
- Mulvey, E. P., & Lidz, C. W. (1984). Clinical considerations in the prediction of dangerousness in mental patients. *CLIN-PSYCHOL-REV*, 4, 379–401.
- Nieva-Fenoll, J. (2018). Repensando Daubert: La Paradoja de la Prueba Pericial. *Civil Procedure Review*, 9(1), 11–26. <https://civilprocedurereview.com/revista/article/view/156>
- Nieva, J. (2010). *La valoración de los deferentes medios de prueba*. In *La valoración de la prueba* (pp. 209–355). Madrid: Marcial Pons.
- Nieva, J. (2020). Carga de la prueba y estándares de prueba: dos reminiscencias del pasado. *InDret*, 3, 404–437. <https://doi.org/10.31009/indret.2020.i3.13>
- Niño, D., & Marrero, D. (2015). El testimonio de un experto. Una perspectiva argentina. In A. Páez (Ed.), *Hechos, evidencia y estándares de prueba: ensayos de epistemología jurídica* (pp. 185–203). Bogotá: Ediciones Uniandes.
- Ochoa-Rosales, M. (2007). *Factores de riesgo en la población carcelaria del Cantón Cuenca Ecuador [Risk factors of the prison population in Cuenca district, Ecuador]*. [Ms Cs dissertation] National University of La Plata, La Plata, Argentina.
- Ochoa Balarezo, J., Guillén, X., Ullauri, D., Narvaez, J., León-Mayer, E., & Folino, J. O. (2016). Relevancia de la evaluación del riesgo de violencia y de la psicopatía en el ámbito forense. *Juris*, 1(16), 113–127.
- Ogloff, J. R. P., & Davis, M. R. (2020). From Predicting Dangerousness to Assessing and Managing Risk for Violence: A Journey Across Four Generations. In J. S. Wormith, L. A. Craig, & T. E. Hogue (Eds.), *The Wiley Handbook of What Works in Violence Risk Management. Theory, Research and Practice*. West Sussex, UK: John Wiley & Sons, Ltd.
- Oldham, J. (2007). Violent behavior. *J Psychiatr Pract*, 13(3), 139.
- Penney, S. R., McMaster, R., & Wilkie, T. (2014). Multirater reliability of the historical, clinical, and risk management-20. *Assessment*, 21(1), 15–27.
- Poder Judicial República de Chile. (2022). Jueces y juezas de Familia se capacitan en uso de guía de salud mental forense. *Prensa y Comunicaciones*. Retrieved from <https://www.pjud.cl/prensa-y-comunicaciones/noticias-del-poder-judicial/79798>
- Radcliffe, K., McMullan, E., & Ramsden, J. (2018). Developing offender manager competencies in completing case formulation: An evaluation of a training and supervision model. *Probation Journal*, 65(1), 27–38. <https://doi.org/10.1177/0264550517744272>
- Rettenberger, M., Eher, R., Desmarais, S. L., Hurducas, C., Arbach-Lucioni, K., Condemarin, C., ... Singh, J. P. (2017). Risk Assessment in clinical practice - Results of the international risk survey (IRiS) from Germany. *Diagnostica*, 63(1). <https://doi.org/10.1026/0012-1924/a000168>
- Ryan, T. J. (2020). *An Evaluation of the Reliability and Quality of Expert and Novice Forensic Case Formulations*. PhD Thesis. Department of Psychology; Simon Fraser University.
- Servicio Médico Legal - Ministerio de Justicia - República de Chile. (2021). *Resolución Exenta N° 2.981. Modifica Normativa Técnica de salud mental aprobada por Resolución Exenta N° 2.349, en el sentido que indica*. Promulgada: 15 de diciembre de 2020. D. O.: 20 de enero de 2021. <https://www.bcn.cl/leychile/navegar?i=1154832&f=2021-01-20>
- Shelton, D. (2011). *Forensic Science in Court. Challenges in the twenty-first Century*. Maryland: Rowman & Littlefield Publishers, Inc.
- Silver, E., Mulvey, E. P., & Monahan, J. (1999). Assessing violence risk among discharged psychiatric patients: Toward an ecological approach. *Law and Human Behavior*, 23(2), 237–255. <https://doi.org/https://doi.org/10.1023/a:1022377003150>
- Singh, J. P., Condemarin, C., & Folino, J. O. (2013). El uso de instrumentos de evaluación de riesgo de violencia en Argentina y Chile [The use of violence risk assessment instruments in Argentina and Chile]. *Revista Criminalidad*, 55(3), 279–290.
- Singh, J. P., Desmarais, S. L., Hurducas, C., Arbach-Lucioni, K., Condemarin, C., Dean, K., ... Otto, R. K. (2014). International Perspectives on the Practical Application of Violence Risk Assessment: A Global Survey of 44 Countries. *International Journal of Forensic Mental Health*, 13(3). <https://doi.org/10.1080/14999013.2014.922141>
- Sturme, P., & McMurrin, M. (2011). *Forensic Case Formulation* (P. Sturme & M. MacMurrin, Eds.). West Sussex: John Wiley & Sons, Ltd.
- Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR - 20: Assessing Risk for Violence Version 2*. Burnaby: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Webster, C. D., Eaves, D., Douglas, K. S., & Wintrup, A. (1995). *The HCR-20 scheme: The assessment of dangerousness and risk*. Burnaby, British Columbia: Simon Fraser University and Forensic Psychiatric Services Commission.
- Wettstein, R. M. (2005). Quality and quality improvement in forensic mental health evaluations. *Journal of the American Academy of Psychiatry and the Law*, 33(2), 158–175.
- Wheable, V., & Davies, J. (2020). Examining the Evidence Base for Forensic Case Formulation: An Integrative Review of Recent Research. *International Journal of Forensic Mental Health*, 19(3), 315–328. <https://doi.org/10.1080/14999013.2019.1707331>
- Wiese, R., Pujol, S., Fernández-Busse, A., Kuczinski, M., Hernández, A., Ruiz, V., ... Folino, J. O. (2010). Innovación didáctica en psiquiatría y evaluación de su impacto. *Tercera Época - Revista Científica de La Facultad de Ciencias Médicas*, 2(2), 1–1.