



Universidad de Valladolid

Facultad de Derecho

Grado en Derecho

¿La ciencia alumbra nuestro camino hacia el futuro? Sobre las implicaciones de la Neurociencia en el Derecho Penal

Presentado por:

Miguel Cuesta Álvarez

Tutelado por:

José Mateos Bustamante

Valladolid, 30 de junio de 2023

Resumen - *Este Trabajo pretende analizar los vínculos existentes entre la Neurociencia y el Derecho Penal desde una perspectiva teórico-práctica, abordando cómo puede afectar la primera al concepto de delito e indagando en los diferentes modos en que puede contribuir la investigación científica a nuestro campo, en particular en el ámbito de impartición de justicia y en la rehabilitación de los delincuentes. Por último, profundizaremos en la posible colisión de estas y otras tecnologías con los Derechos Humanos, analizando aspectos controvertidos como la posibilidad de atribuir un estatus propio a la Inteligencia Artificial. Los resultados de nuestro análisis muestran que estas innovaciones están cambiando nuestra sociedad de un modo nunca antes visto; por ello, es menester advertir sobre los desafíos que traen consigo, razón por la cual está aumentando el interés de los gobiernos en legislar sobre estas materias.*

Palabras clave: neurociencia, derecho penal, teoría del delito, determinismo, Inteligencia Artificial, justicia predictiva, política criminal, neuroderechos

Abstract – This Project aims to analyse the existing bonds between Neuroscience and Criminal Law in a theoretical-practical fashion. This issue will be tackled both by emphasizing how the former might thoroughly alter the concept of crime and exploring the different ways scientific research can be of assistance in our field of knowledge, particularly in terms of dispensing justice and rehabilitating criminals. Lastly, we will provide insight into the potential collision of these and other technologies and Human Rights, addressing several controversial aspects such as the plausibility of Artificial Intelligence being given their own status. Our findings suggest that these innovations are reshaping our society in an unprecedented way; consequently, concern should be raised about the challenges they pose, hence the increasing will of governments to promote legislation regarding this matter.

Keywords: neuroscience, criminal law, theory of crime, determinism, Artificial Intelligence, predictive justice, criminal policy, neurorights

ÍNDICE

1. Introducción	4
2. Neurociencia y Teoría del Delito	6
2.1 La chispa que desató el incendio: el (re)surgimiento del determinismo	6
2.2 Los fundamentos del castigo, en llamas: presupuestos del delito amenazados	8
2.3 Y tras la tormenta, llegó la calma: ¿llegó a prender la mecha en realidad?	13
3. Neurociencia, Proceso Penal y Justicia Predictiva	15
3.1 Nuevos medios de prueba para nuevos tiempos	15
3.1.1 <i>Su acogida en el ámbito forense</i>	17
3.1.1.1 <u>El origen: la neurociencia en la jurisprudencia italiana</u>	17
3.1.1.2 <u>La prueba neurocientífica en España: un caso paradigmático</u>	19
3.2 De la cualidad de imprescindibilidad del ser humano en la impartición de justicia y de cómo un robot podría mejorar al juez: la justicia predictiva.....	21
3.2.1 <i>No es oro todo lo que reluce: críticas a la justicia predictiva</i>	22
3.2.2 <i>Después de todo, ¿ha perdido la partida la justicia predictiva?</i>	25
4. Neurociencia y Política Criminal	27
4.1 Un inesperado y poderoso aliado en la prevención del delito.....	27
4.2 Distintos caminos, un mismo propósito: herramientas neurocientíficas en la rehabilitación del delincuente	29
4.3 La Inteligencia Artificial en acción: evaluaciones sistemáticas de riesgo	34
4.3.1 <i>Riesgo de reincidencia y proceso penal: COMPAS</i>	34
4.3.2 <i>¿Predecir la escena del crimen antes de que ocurra? Geolítica</i>	37
4.3.3 <i>El último eslabón de la cadena: de cómo auxiliar a las víctimas. VioGén</i>	40
4.3.4 <i>La evaluación del riesgo al otro lado de los barrotes: RisCanvi</i>	44
5. Desafíos ético-penales de la Neurociencia	49
5.1 ¿La IA, responsable de hechos punibles? Estado actual y obstáculos para su regulación	49
5.2 ¿Clara como el cristal? Responsabilidad del programador y falta de transparencia	52
5.2.1 <i>¿Los sistemas de evaluación de riesgo, a prueba? El caso State vs. Loomis</i>	54
5.3 ¿Y nuestro papel en esta Revolución? Relación de la neurociencia con los Derechos Humanos	58
5.3.1 <i>Nuevos avances, nuevos derechos: la perspectiva de Ienca y Andorno (2022)</i>	61
5.3.2 <i>Un rayo de esperanza: mirada al derecho comparado</i>	62
6. ¿Qué podemos aprender de todo esto? Conclusiones	67
7. Bibliografía	70

1. INTRODUCCIÓN

La neurociencia es una de las disciplinas más pujantes y con un futuro más prometedor en la actualidad, atendiendo a las distintas aplicaciones que la sociedad puede dar a sus innovaciones: desde el ámbito estrictamente médico (con medios de diagnóstico más eficaces) al tecnológico, pasando por otros como el educativo, parece evidente que sus aportaciones pueden ser de extraordinaria utilidad para el progreso de la humanidad. No obstante, con cada nuevo descubrimiento surge un desafío; el de integrarlo en nuestras costumbres, con nuestra propia forma de entender el mundo. Esto ya ocurrió en épocas pretéritas; durante la Revolución Industrial, tuvo que adaptarse súbitamente a la llegada de cambios radicales en la producción de bienes o el transporte de los mismos, lo que acabó sentando las bases de disciplinas como el Derecho Laboral. La neurociencia en sus distintas expresiones ha llegado, y lo ha hecho para quedarse; consecuentemente, no podemos quedarnos de brazos cruzados, debemos anticiparnos a esos cambios en pos de un mayor beneficio común. Debemos afrontar este desafío sin paliativos con una amplitud de miras lo más extensa posible y con un enfoque multidisciplinar, tratando de buscar una vía de salida que satisfaga el interés común.

Es aquí donde entra en juego el Derecho, como ciencia social encargada de la inserción de estos hallazgos en un sistema dotando de validez a un conjunto de normas que permitan la inclusión de las herramientas neurocientíficas en nuestro día a día sin atentar contra el orden establecido. Especial relevancia cobran en este sentido ramas como el Derecho Penal, y también otras como el Derecho Procesal y los Derechos Humanos, tan en boga en la actualidad. ¿Cómo pueden cambiar las consideraciones sobre el fundamento del delito y la pena que desde los inicios de esta disciplina han permanecido indiscutidas a la luz de las diferentes aportaciones de la neurociencia? ¿Cómo pueden sernos de ayuda en diferentes ámbitos jurídicos, como el criminológico o el judicial? ¿Qué régimen debemos establecer para la convivencia con los sistemas de Inteligencia Artificial? Y, finalmente, ¿cómo afectará todo esto a la especie humana? ¿Cambiarán el orden del mundo estos descubrimientos...o seremos nosotros los que cambiemos con su llegada?

Todos estos interrogantes son solo algunas de las líneas de discusión que el mundo neurocientífico puede ofrecernos como juristas. Se pretende con esta contribución realizar un análisis de las múltiples aristas de la neurociencia, relacionando los distintos puntos de

conexión que esta disciplina tiene con el Derecho (y, en particular, sus implicaciones en el Derecho Penal) y las posibilidades y limitaciones de estas innovaciones, a fin de arrojar un pequeño destello de luz sobre este terreno, menos explorado de lo que creemos, y con un lado oscuro.

De esta forma, comenzaremos la exposición analizando cómo pueden influir las herramientas neurocientíficas en la Teoría del Delito, profundizando en la dogmática jurídico-penal sobre sus elementos, en pos de obtener una respuesta a la cuestión sobre si la neurociencia realmente puede quebrar los cimientos del fundamento de la pena o si, por el contrario, permanecen intactos los postulados clásicos acerca de la conducta estrellada de nuestra disciplina. En siguiente lugar, centraremos el foco de análisis en la faceta judicial de los instrumentos neurocientíficos y la Inteligencia Artificial, indagando en las distintas formas en que pueden ayudarnos en el proceso penal, tanto desde el punto de vista de las partes (con la posible admisión de informes basados en evidencia neurocientífica como prueba en el seno de un proceso) como del órgano decisor (con la denominada justicia predictiva, en torno a la posibilidad de sustituir al juez por un sistema de aprendizaje autónomo). En tercer lugar, dedicaremos otro espacio a estudiar la faceta asistencial de la neurociencia en la política criminal y las formas en que pueden auxiliarnos en la prevención del delito, con ejemplos reales de su utilización. Por último, nos detendremos a examinar los desafíos que la neurociencia ha traído consigo y las distintas valoraciones que se pueden efectuar al respecto desde la óptica de los Derechos Humanos que deben hacer partícipes a todas las áreas de investigación científica, académica y social, en pos de ofrecer soluciones al conjunto de la sociedad que garanticen una coexistencia fructífera para todos. Solo de este modo podremos afrontar el reto que se nos presenta y al que no debemos renunciar: al fin y al cabo, es el futuro de la humanidad lo que está en juego.

2. NEUROCIENCIA Y TEORÍA DEL DELITO

2.1. La chispa que desató el incendio: el (re)surgimiento del determinismo

“Podemos definir el delito como la acción u omisión típica, antijurídica, culpable y punible” (Muñoz Conde, 1989, p. 20). Esta definición de la conducta por excelencia del Derecho Penal ha poblado durante décadas los estudios de nuestra disciplina, marcando el camino para discernir aquello que se puede calificar como delito y aquello que no. Sin embargo, la irrupción de la neurociencia ha puesto en jaque la validez de los elementos de esta definición, enfrentando a partidarios de dos corrientes contrapuestas (el determinismo y el indeterminismo) y que amenazan la base de la legitimidad de las sanciones penales: la culpabilidad por el hecho.

En la década de los 80, el neurólogo Benjamin Libet llevó a cabo una serie de experimentos que tambalearían desde entonces el planteamiento clásico acerca de la voluntad, desdibujando la línea entre consciente e inconsciente (Libet *et al.*, 1982, 1983, y Libet, 1985). En estas sesiones, Libet pidió a los sujetos participantes que observasen las manecillas de un reloj girando a gran velocidad y que pulsaran un botón para detenerlas cuando ellos quisieran. Tras ello, les solicitó que anotaran qué hora marcaban las manecillas cuando tomaron la decisión de pararlas apretando el botón. Mediante la colocación de electrodos en el cuero cabelludo, Libet recogió datos de los impulsos producidos por el cerebro y concluyó que ya había actividad en el cerebro (denominada potencial de disponibilidad), dirigida a producir el movimiento de la mano para pulsar el botón antes de que el sujeto sintiese que había querido mover la muñeca. De esta forma, demostró que las decisiones que el ser humano creía haber tomado voluntariamente en realidad habían sido inducidas previamente por el cerebro: dicho de otro modo, nuestras acciones estarían determinadas por factores inconscientes que no podemos controlar¹. Años más tarde, Fried *et al.* (2011) emplearon las mismas técnicas para descubrir que, 1,5 segundos antes de que la persona sintiese conscientemente que quería ejecutar una acción, las neuronas de determinada zona del cerebro incrementaban o disminuían su actividad. Eso significó que era posible predecir cuándo una persona iba a querer algo hasta 700 milisegundos antes de que ella lo deseara conscientemente, basándose en los registros de picos y valles de la

¹ A salvo de una suerte de veto que sí puede controlar voluntariamente el sujeto para abstenerse de realizar una acción.

actividad neuronal que mostraba en sus exámenes con fMRI (resonancia magnética funcional) o EEG (electroencefalograma), con una fiabilidad del 80%.

¿Qué implicaciones tiene esto para el Derecho Penal? Tradicionalmente, el sistema punitivo del que se había valido la justicia se basaba en la culpabilidad por el hecho: un criminal cometía un hecho punible porque tenía la voluntad de hacerlo, y dado que podría haberlo evitado conscientemente, había de ser castigado. Sin embargo, a la luz de las observaciones de Libet, podría concluirse que el delincuente no tiene voz ni voto en la comisión de un delito, pues se habría visto empujado a hacerlo por una decisión inconsciente y propia de su cerebro². Según Roth (2004), las decisiones están prefijadas por procesos neuronales, así que no hay lugar para el acto consciente de voluntad³. Se pone en tela de juicio, por tanto, la capacidad de autodeterminación del ser humano y la vigencia de su libre albedrío, contemplando ahora la posibilidad de cambiar el fundamento de la pena.

Un famoso ejemplo real es el de Phineas Gage⁴, trabajador en la construcción del ferrocarril de Cavendish (Estados Unidos) en 1848. Este joven de 25 años fue atravesado por una barra de metal que salió disparada a causa de una explosión fortuita a corta distancia. Este hierro, de 3 centímetros de grosor y 109 centímetros de longitud, atravesó su cabeza desde la parte inferior de la mejilla izquierda y salió por la parte superior del cráneo, perforando su cerebro⁵. Tras sobrevivir al accidente, fue atendido y eventualmente pudo volver al trabajo, sin afectación en la memoria, inteligencia o habla. Sin embargo, su carácter cambió drásticamente: el antaño hombre afable se convirtió en alguien irresponsable, impaciente e impulsivo, lo que acabó con su trabajo y su vida ejemplar. Unos años después de su muerte, se propuso por parte de los médicos que la barra habría afectado de manera irreversible al lóbulo frontal izquierdo (y, en concreto, a su corteza prefrontal, encargado de

² En famosas palabras de Feijoo Sánchez (2017), poniéndose en el lugar de los neurocientíficos, “no hacemos lo que decidimos, sino que decidimos lo que vamos a hacer de todas maneras” (p. 22).

³ En publicaciones posteriores (Pauen y Roth, 2008, entre otras), este autor suavizó su visión sobre el tema, inclinándose hacia un modelo compatibilista basado en la prevención general positiva, si bien otro sector de la doctrina (Pérez Manzano, 2011) afirma que no es posible sostener un modelo que conjugue prevención general y determinismo.

⁴ La descripción de este archiconocido caso figura en gran parte de la literatura penal sobre neurociencia. Por todos, Feijoo Sánchez (2017).

⁵ Las primeras simulaciones realizadas sobre la cabeza de Gage tras el accidente pueden consultarse en Damasio *et al.* (1994). Los diagramas ilustraron la portada del volumen en que fue publicado el artículo (264, de 20 de mayo).

la gestión de las emociones y el proceso de toma de decisiones). De esta forma, la neurociencia pudo diagnosticar (*post mortem*, claro está) una sociopatía adquirida a causa del accidente que habría eximido a Gage de culpabilidad en caso de haber cometido algún hecho punible⁶.

2.2. Los fundamentos del castigo, en llamas: presupuestos del delito amenazados

Este es precisamente el germen del impulso de las teorías de prevención especial que, al amparo de la neurociencia y sus innovadores instrumentos, buscan la rehabilitación del delincuente para corregir su conducta⁷, pues defienden que el ser humano no pudo escoger otra vía alternativa. A ello responde Feijoo Sánchez (2017) aseverando que “nunca se ha declarado culpable a nadie por haber podido actuar de otra manera” (p. 27). Por su parte, Greco (2022), aun admitiendo que el cerebro procesa más de lo que reconocemos de manera consciente, señala que limitarse a culpar a alguien tan solo por haber podido actuar de otro modo haría caer al sistema en dos polos opuestos. Por una parte, asumir la teoría determinista fuerte haría que ningún delito pudiese ser castigado, pues el ser humano tan solo podría conformarse con lo que su cerebro le ha empujado a hacer. Pérez Manzano (2011) indica que entonces el individuo se “haría acreedor de lástima por las pésimas cartas que le han tocado (...) ni merecería la sanción penal, cuya imposición devendría entonces injusta” (pp. 2-3). Hillenkamp (2005) afirma en este sentido que el determinismo absoluto derrumbaría todo el sistema del Derecho. Al mismo tiempo, adoptar la postura opuesta (es decir, presumir que el ser humano goza de la capacidad de controlar conscientemente la totalidad de las acciones –indeterminismo absoluto–), nos llevaría a establecer una voluntad aleatoria, hasta el punto de no poder exigir tampoco responsabilidad alguna por los actos de uno mismo; al no estar motivada la conducta por nada, la culpabilidad llegaría en forma de “lotería de mala suerte” (Feijoo Sánchez, 2017, p. 45).

Feijoo (2017, 2022), Nombela Otero (2022) o Aguilar Cárceles (2022) defienden la existencia de un doble sustrato biológico y psicológico en la conformación de la voluntad;

⁶ Este sería un ejemplo perfecto para ilustrar la conocida como “paradoja de la culpabilidad”, con la que hay que tener cuidado para no atribuir fácilmente la condición de inimputabilidad. Formulada por Urbaniok *et al.* (2011), esta paradoja consiste en lo siguiente: cuanto más grave es el delito, mayor se considera que es la culpa. Sin embargo, cuanto más grave es el delito cometido, más claramente se trata de poner de relieve (por parte del acusado, sobre todo) la base psicológica de la voluntad del delincuente, para así evitar el castigo.

⁷ A esta materia dedicaremos el grueso de este trabajo.

pues la toma de decisiones no aparece solamente influida por factores neurobiológicos, sino también por otros procedentes del entorno. Lo que Feijoo propone es adoptar una postura de síntesis entre determinismo e indeterminismo.

A pesar de todo, Feijoo Sánchez (2022) subraya que la revolución penal de la que hablábamos no es tan relevante como podría parecer: en realidad, la existencia o no del libre albedrío no amenaza los fundamentos de la pena, sino que tan solo pondría en entredicho la teoría subjetiva de la retribución. Günther (2005) señala el motivo de esta sorprendente afirmación: la culpabilidad que el Derecho Penal utiliza para imputar la comisión de un hecho punible a un sujeto no es en ningún caso una cualidad empírica que se pueda observar en el cerebro. Por el contrario, la teoría del delito basa la atribución de responsabilidad en valoraciones normativas, surgidas a través de un reconocimiento social histórico⁸. Consecuencia de esto es precisamente el hecho de que la sociedad considera culpable a un sujeto no tanto por su estructura cerebral o por sus motivos para consumir el delito (características naturalísticas de la acción), sino más bien porque le parece que ha sobrepasado una línea roja que puede poner en peligro la convivencia si no se sanciona⁹. En la misma línea se pronuncia Pérez Barberá (2011): el dolo, como elemento de la culpabilidad, se adscribe normativamente, por lo que aquí no deben entrar los desarrollos neurocientíficos. Esto hace que se deba castigar al delincuente con independencia de la percepción subjetiva de su conducta o su entorno, teniendo solamente en cuenta las disposiciones jurídicas vigentes en el momento de la comisión del delito y no en momentos anteriores al hecho (como defiende la neurociencia).

Otros autores han interpretado la llegada de la neurociencia de manera distinta, opinando acerca de qué presupuestos del delito son los que se han puesto en duda. Greco (2022) critica severamente que los neurocientíficos hayan apuntado a la culpabilidad, al libre

⁸ Grandes voces autorizadas de la doctrina alemana se han pronunciado en idéntica dirección, como Jakobs (2005). Por su parte, Hassemer (2009) se posiciona en términos similares: al Derecho Penal no le afecta la verdad científica, sino la verdad formal que busca la resolución pacífica de los conflictos.

⁹ Jakobs (1989, 2005) relaciona esto con la teoría comunicativa, consistente en que lo realmente importante en lo que al fundamento de la pena se refiere no son tanto las características naturalísticas, sino la relevancia comunicativa que la sociedad le dé a una acción, dejando a un lado la responsabilidad por erosión del bien jurídico y sustituyéndola por la erosión de la norma. El objetivo de la culpabilidad sería mantener el sistema social, estabilizando la vigencia de la norma (prevención general positiva).

albedrío, sin detenerse a pensar en qué consiste exactamente el libre albedrío¹⁰, que es “algo de lo que se puede hablar, pero en lo que no somos capaces de pensar” (p. 165). Acaba por concluir que la culpabilidad no se basa en el libre albedrío, sino en una realidad “en la que cada uno de nosotros, en principio, es responsable de la forma en que conduce su propia vida (...)” (p. 166). Al referirse así a algo que engloba a todos los seres humanos, Greco indica que los descubrimientos neurocientíficos no son importantes en lo que a culpabilidad se refiere. Sin embargo, sí que admite su posible contribución en la delimitación de la imputabilidad, como capacidad de ser culpable, y de dolo. Respecto del primero de los elementos -amén del uso de herramientas médicas para el diagnóstico de trastornos severos que pudiesen encajar en el supuesto de exención de responsabilidad de “anomalía o alteración psíquica” (artículo 20.1 del Código Penal)- la neurociencia podría detectar la anomalía orgánica en el cerebro, dotando de una base física a las alteraciones psíquicas invisibles sin estos instrumentos¹¹. En cuanto a lo que el dolo se refiere, podría ampliarse su concepto para englobar también a la actividad inconsciente del cerebro y determinar el porqué de la conducta del criminal, “observar lo que ocurre en el interior de la caja negra” (p. 170), tarea en la que puede ser de gran ayuda la neurociencia.

Palma (2022), por su parte, relaciona el cambio en el fundamento de la culpabilidad con la irrupción de los sistemas de Inteligencia Artificial. El determinismo neurocientífico postula que el ser humano no puede escoger sus propias decisiones. A la par, la IA hace lo propio al crear sistemas que pueden decidir autónomamente; el ser humano ya no es la única fuente de responsabilidad, sino que (de existir el libre albedrío) las máquinas pueden emular esa capacidad.

Por otro lado, Alonso Álamo (2022) reconoce la utilidad de la neurociencia a la hora de delimitar los conceptos de falta de acción y falta de culpabilidad. Esta penalista estudia la difuminada línea que hay entre ambos conceptos y la forma en la que puede ayudar la neurociencia a discernir entre los dos. De esta forma, constata que la acción se remite a la

¹⁰ Sobre este extremo, distingue dos corrientes alternativas: un determinismo metodológico representado por autores como Feuerbach (1980), y un agnosticismo basado en la imposibilidad de demostrar empíricamente la existencia del libre albedrío, abanderado por Roxin (1973).

¹¹ Esta utilidad se ve más claramente en el sistema penal alemán, como admite Greco, en el que hay varias categorías de inimputabilidad (artículo 20 del Código Penal): una de ellas, “otros trastornos anímicos graves” (*schwere andere seelische Störung*), comprende casos en los que no hay una base fisiológica que sustente el padecimiento de la anomalía. Las herramientas neurocientíficas pueden servir para hallar esta base de forma definitiva, con las facilidades que ello comportaría para la atribución de inimputabilidad.

voluntad; consecuentemente, pueden enumerarse determinados supuestos en los que no haya responsabilidad penal, pero no por falta de culpabilidad, sino por falta de acción, ya que esos casos no se refieren a la voluntad. Así, es menester distinguir ausencia de acción por no haber voluntariedad y exclusión de la culpabilidad al uso.

De este modo, el tratamiento legal de la falta de acción se centra en tres supuestos: actuación bajo *vis absoluta*, movimientos reflejos y estado de inconsciencia plena. En el primero de los casos, “se requiere la supresión completa de la voluntariedad” (p. 337), rigiendo la máxima latina *non agit sed agitur* (el sujeto no actúa, sino que es actuado). Mir Puig (2015) expone asimismo su visión sobre esta materia de la siguiente manera; si un sujeto es amenazado de forma que no puede resistirse a ejecutar la conducta que alguien externamente le obliga a realizar, no habría acción por falta de voluntad. Distinto sería el caso en que el sujeto pudo resistirse a las amenazas y por tanto, no habría *vis absoluta*: existiría cierta voluntad y con ello no se podría excluir la acción ni la responsabilidad. En segundo lugar, los movimientos reflejos carecen de acción, porque de nuevo falta la voluntariedad. Es en este punto donde las herramientas neurocientíficas pueden sernos de utilidad para determinar la ausencia de acción: como respuesta a un estímulo, el cerebro envía la orden de contracción muscular, por ejemplo, de forma automática, sin que haya participación alguna de la voluntad¹². En otros casos dudosos (en los se puede producir un “desdoblamiento” que Alonso Álamo identifica con una doble posibilidad de atribuir acción o no hacerlo), como los ataques epilépticos, dependerá del control que tenga el sujeto sobre su voluntad; si, aun sufriendo dicha crisis, mantiene sus facultades volitivas, habría acción. Por último, los actos que el sujeto haya efectuado bajo inconsciencia plena tampoco pueden ser tipificados como hecho punible a causa de la inexistencia de acción, con supuestos de desdoblamiento en el caso de narcolepsia o embriaguez letárgica.

Así, en todos estos casos no hay delito por no haber acción, distinguiéndolo de otros hechos en los cuales faltaría la culpabilidad. A la luz de las consideraciones expuestas, Alonso indica que no todas las personas que cometan delitos y padezcan, por ejemplo, una anomalía psíquica grave, deben considerarse inimputables (donde, por tanto, faltaría la culpabilidad), sino que muchas de ellas deben reconducirse a una inexistencia de acción, por no haber ninguna referencia a la voluntad del sujeto en su actuación. Como señala la penalista, “no es

¹² Alonso Álamo diferencia esto de las conductas impulsivas, en las que sí hay cierto elemento volitivo, y por tanto, habría a su vez acción.

que sea incapaz de comprender que lo que se hace se halla prohibido por el ordenamiento jurídico (inimputabilidad), sino que propiamente no actúa porque su acción no le pertenece, no es suya, no es referible a su voluntad” (pp. 343-344). Es en este punto donde los diferentes avances neurocientíficos pueden arrojar luz constatando de manera empírica qué trastornos están exentos del elemento volitivo por parte del sujeto y por tanto, deberían ser considerados como casos de ausencia de acción, como en la enfermedad de Alzheimer o en Trastornos del Espectro Autista¹³.

Busato (2022) estudia también la relación existente entre acción y voluntad, si bien lo hace desde una óptica ciertamente peculiar: la atribución de responsabilidad a la IA. Este autor plantea un debate acerca de regular el estatuto de estos novedosos sistemas (cuestión de la que nos ocuparemos más tarde), criticando que se ha empezado a legislar esta materia centrándose únicamente en las dudas sobre la presencia o no de culpabilidad en sus actos, pasando por alto otros elementos del delito y en particular, la acción. Busato asevera que las máquinas no pueden expresar objetivamente intenciones de efectuar una acción ni tampoco cuentan con un elemento volitivo de venganza, por ejemplo, que pueda servirles de móvil para actuar de un determinado modo. Esta no es sino una aplicación de lo observado por Alonso Álamo (2022): lo que Busato concluye es que la IA no debe ser considerada como incapaz de culpabilidad (inimputable) sino que lo que debe discutirse es su incapacidad de acción, añadiendo sin saberlo un supuesto más a la enumeración de Alonso.

Pérez Manzano (2022), por su parte, analiza el efecto de la neurociencia en la culpabilidad, centrándose en el elemento cognitivo del dolo, redirigido también hacia la voluntad. Para esta autora, el dolo consta de una naturaleza mixta, con elementos normativos (al uso de lo que opinaban Feijoo Sánchez o Pérez Barberá) y otros de carácter fáctico, lo que hace que haya que acreditar los estados mentales. Consiguientemente, sí sería útil el papel de la neurociencia en este extremo (Pérez Manzano, 2011, 2022), para reevaluar el concepto del dolo y ver cómo puede operar a la luz de las apreciaciones médicas. En particular, se propone incorporar las decisiones preconscientes (de las que no nos percatamos, pero que informan nuestra conducta) en el dolo. La penalista defiende que, si un hecho tiene un procesamiento plenamente preconsciente, ello excluiría el dolo; si, por contra, la actividad

¹³ Defiende Alonso Álamo su aplicación también en niños y jóvenes, en los que no solo faltaría la voluntad - Martín López (2022) defiende la inimputabilidad en menores de 14 años, e incluso en menores de mayor edad, sino también la acción.

no fuese totalmente preconscious, sí podría apreciarse¹⁴. En consecuencia, para poder observarlo debe haber cierto control sobre la conducta, pues “requiere la activación de las zonas cerebrales que se vinculan a la consciencia¹⁵” (p. 408), aun asumiendo cierta actividad preconscious tras el telón del escenario. Así, el actuar doloso ya no se limitará a los actos plenamente conscientes, sino que la neurociencia será capaz de ampliarlo a casos en los que parte de ese registro de la realidad se haya llevado a cabo por el preconscious. Por otro lado, Pérez identifica el elemento volitivo del dolo con las funciones de control de esas decisiones tomadas en cooperación con el preconscious, dado que la vigilancia de nuestras acciones ya requiere cierto control consciente sobre ello, cerrando así el círculo en torno a las implicaciones de la neurociencia en la determinación del papel de la consciencia y preconscious en el dolo.

Por último, para Ruggeri (2022) los campos de autodeterminación y libre albedrío son los más fértiles para la discusión de la neurociencia. Admitiendo que la libertad absoluta de decisión no existe, no se puede desdeñar la capacidad que tiene la neurociencia de esclarecer los motivos de las actuaciones de un delincuente. Es especialmente relevante no tanto en sede de imputabilidad, sino más bien para “establecer el nivel de libertad de autodeterminación que realísticamente se le podía exigir al reo en el momento del supuesto delito” (p. 588), con el inconveniente de carácter temporal de no saber exactamente en qué pensaba cuando lo cometió; se pueden observar marcadores de la conducta en su cerebro, si bien resulta algo más difícil probar que ya existían en la comisión del hecho punible. En orden a subsanar esta problemática, este autor italiano subraya los beneficios que tendría usar métodos neurocientíficos también con posterioridad a la comisión del delito, a fin de comprobar su capacidad de entendimiento de la situación, lo que podría servir de ayuda para determinar el nivel de imputabilidad que tenía al realizar el hecho delictivo.

2.3 Y tras la tormenta, llegó la calma: ¿llegó a prender la mecha en realidad?

Hemos comprobado cómo se han incorporado las innovaciones neurocientíficas en la teoría del delito, con la posible modificación de las características del mismo. A modo de

¹⁴ De hecho, señala las dificultades que plantearía asumir la tesis contraria, es decir, considerar suficiente una actividad exclusivamente preconscious para atribuir dolo, como la imposibilidad de distinguir entre dolo e imprudencia.

¹⁵ Pérez Manzano lo identifica como la conexión del circuito neuronal previo con el córtex frontal, en cuyo caso habrá dolo en la conducta.

conclusión, recabaremos la opinión de varios autores sobre la admisibilidad de los presupuestos neurocientíficos en esta parcela del Derecho. Pérez Manzano (2011) se muestra favorable a abrir las puertas a la neurociencia para complementar al Derecho Penal, pues ambas disciplinas tienen en común algo en sus objetos de estudio: la neurociencia estudia el funcionamiento del cerebro en la generación de la conducta humana y el Derecho regula esas conductas. Hassemer (2011) afirma rotundamente que mientras no haya evidencia científica concluyente para dar carpetazo al debate sobre el libre albedrío (“que puede durar décadas o siglos” augura el ilustre penalista alemán (p. 3)), la justicia penal debe seguir juzgando sobre la imputabilidad de los encausados, es decir, que no puede dar por válido el determinismo. Feijoo Sánchez (2022), como ya anticipamos, no vislumbra un cambio radical en el fundamento de la pena bajo el paraguas de la neurociencia, ya que defiende que las explicaciones neurocientíficas no pueden hacernos renunciar a la atribución de responsabilidad basándonos en la culpabilidad por el hecho, esgrimiendo que la teoría comunicativa del delito (basada en la normatividad de la culpabilidad) es un ejemplo de cómo la culpabilidad es un hecho social, compatible con los avances neurocientíficos. Por último, Alonso Álamo (2018) señala que la opción determinista llevaría a no poder fundamentar la responsabilidad penal, pues “en ningún caso, bajo ninguna circunstancia, nunca, somos libres de querer lo que queremos” (p. 14). No obstante, al no haber una base científica que demuestre de manera evidente el determinismo (lo que reconduciría el derecho penal de la culpabilidad al derecho de la prevención o al de la peligrosidad), debemos seguir fundamentando la culpabilidad en la existencia de libre albedrío.

3. NEUROCIENCIA, PROCESO PENAL Y JUSTICIA PREDICTIVA

Como se ha expuesto anteriormente, una de las más novedosas aportaciones de la neurociencia al ámbito penal es la Inteligencia Artificial. En la actualidad, la IA cuenta con tres estrechos vínculos que interesan a nuestra disciplina: la potencial admisión de pruebas neurocientíficas en el proceso penal, obtenidas a través de elementos cada vez más sofisticados –con el *machine learning* a la cabeza como principal exponente–, la sustitución del juez por un algoritmo inteligente capaz de producir una sentencia condenatoria o absolutoria a la luz de las pruebas propuestas y, en particular, la rehabilitación de delincuentes, prevención y predicción de futuros delitos mediante el uso de sistemas de IA.

Si bien centraremos nuestro análisis en el tercero de estos puntos, la estrecha interrelación del Derecho Penal con el Derecho Procesal (y la innegable utilidad de las herramientas que la neurociencia puede ofrecer como respuesta a las violaciones del ordenamiento jurídico) justifica que nos detengamos someramente en las dos primeras facetas.

3.1 Nuevos medios de prueba para nuevos tiempos

La discusión sobre la admisión o no de pruebas neurocientíficas para fundamentar decisiones en procesos penales ha sido ampliamente discutida por la doctrina, que se halla dividida en torno a este debate: así, Porciúncula (2022) se posiciona radicalmente en contra de estos métodos por atentar contra el *nemo tenetur se ipsum accusare* (habitualmente abreviado como *nemo tenetur*). Esta garantía procesal es en esencia la posibilidad con la que cuenta el imputado para no declarar contra sí mismo, que se vería violada, opina Porciúncula, si se somete a exámenes médicos neurocientíficos, en particular el fMRI (imagen por resonancia magnética funcional) o el *brain fingerprinting*. Esta última técnica, entre otras aplicaciones, permite bucear en la memoria del encausado para comprobar si reconoce algo proyectado en pantalla, como un objeto con el que se ha cometido el crimen o un lugar relacionado con la víctima. Según parte de la doctrina neurocientífica, el cerebro produce la conocida como respuesta “ajá” (Farwell, 2012) emitiendo ondas que pueden ser registradas por los agentes

mediante una encefalografía¹⁶. De este modo, el imputado estaría incriminándose a sí mismo de manera inconsciente e irremediable. Merkel (2022) también trae a colación la posible confrontación entre fMRI y *nemo tenetur*, no tanto por el efecto “ajá” incontrolable por el sujeto, sino más bien por la posibilidad de acceder a los recuerdos del examinado, sin llegar a una conclusión clara sobre su admisibilidad.

Otros autores han adoptado una postura algo más moderada, defendiendo que estas pruebas son compatibles con las tradicionales¹⁷. De esta forma, Aguilar Cárceles (2022) acepta su utilidad para comprobar la existencia de culpabilidad en sujetos con patologías mentales en el momento de la comisión del delito. En esta misma línea, es la doctrina italiana –amparándose en el mayor impacto que en la jurisprudencia de los órganos de este país ha tenido la neurociencia- la que más ha defendido su empleo en el ámbito judicial, superados los escepticismos iniciales (Bertolino, 2008). Grandi (2022) adopta una posición de síntesis, combinando estas pruebas como indicios de la culpabilidad del sujeto, pero sin que la decisión del juez pueda fundamentarse exclusivamente en ellas. Por otra parte, Ruggeri (2022) pone el acento en que proporcionan “no solo una garantía de mayor fiabilidad, sino incluso de certeza, en la reconstrucción de los hechos” (p. 593). Cree el penalista transalpino que no distorsionan los juicios de imputación, sino que tan solo delimitan el nivel de autodeterminación que tenía el reo, por lo que son admisibles como complemento a las pruebas tradicionales.

De hecho, Ruggeri discrepa con Porciúncula a la hora de valorar el posible choque del fMRI con el *nemo tenetur*, acogiéndose a la jurisprudencia del TEDH. La Corte de Estrasburgo, en su sentencia de 17 de diciembre de 1996 (Saunders c. Reino Unido¹⁸), declaró que solo las pruebas en las que pueda invocarse el derecho al silencio chocan frontalmente con el *nemo tenetur*; de ello puede desprenderse que “los actos de investigación (...) que no conllevan ninguna participación activa del destinatario, quien simplemente tiene que padecer

¹⁶ Tampoco el *brain fingerprinting* goza aún de plena aceptación en la comunidad científica. Como ejemplo de estudios cuestionando su efectividad, véase Meijer *et al.* (2013).

¹⁷ E incluso imprescindibles: “ya no podemos fabricar un modelo (...) que parta de la base de un sistema sociocognitivo que no contemple bases científicas (...). De eso depende la legitimidad social del derecho penal” (Jiménez Martínez, 2022, p. 487).

¹⁸ El caso relaciona el *nemo tenetur* con el derecho a guardar silencio, si bien el Tribunal aprecia que ello no se extiende al uso de material obtenido del acusado de forma independiente a su voluntad, como análisis de sangre u orina. Por analogía, lo mismo se puede aplicar a los resultados de un fMRI o de un proceso de *brain fingerprinting*.

su realización” (Ruggeri, 2022, p. 606) son perfectamente aceptables como prueba y no entran en conflicto con esta garantía procesal¹⁹.

3.1.1 Su acogida en el ámbito forense

Por último, es menester detenernos en algunos casos en los que se ha puesto de relieve el incipiente desarrollo de las pruebas neurocientíficas en el proceso penal, tanto a nivel nacional como internacional, observando la distinta acogida que han tenido en diferentes ordenamientos.

3.1.1.1 El origen: la neurociencia en la jurisprudencia italiana

La entrada efectiva con éxito de la neurociencia en el proceso penal a nivel europeo se produjo en Italia en el año 2009 –ello explica la atención que la doctrina italiana ha dedicado a la relación neurociencia-derecho penal- con la sentencia 5 del Tribunal de Apelación de Trieste de 18 de septiembre de 2009²⁰. Observa Grandi (2016) que hasta ese momento la tendencia imperante era combinar diagnósticos tradicionales con las novedosas pruebas neurocientíficas. En palabras ciertamente elocuentes de Sartori y Zangrossi (2016), la neurociencia no pretende la sustitución sino el enriquecimiento de la evaluación psiquiátrico-forense tradicional con el fin de aumentar su objetividad y precisión²¹. Como vemos, no es una posición muy alejada de lo que proponen el propio Grandi (2016, 2022) o los autores españoles mencionados anteriormente.

¹⁹ La clave es que el *brain fingerprinting* recoge la información observando las ondas emitidas por el cerebro del sujeto independientemente de su voluntad. No se puede decir lo mismo de otras pruebas en las que se pregunte directamente al sujeto y se trate de detectar si dice la verdad o no, en cuyo caso sí se producirían violaciones del *nemo tenetur*.

²⁰ El conocido como “Caso Bayout” resolvió la controversia sobre un ciudadano argelino que mató con un puñal a un colombiano en la estación de trenes de Udine (Italia) tras ser increpado por este.

²¹ Traducción propia de Grandi (2022, p. 383).

¿Qué importancia tiene por tanto este caso en el devenir del valor probatorio de la neurociencia desde entonces? Por primera vez^{22 23}, la valoración que un juez europeo realiza sobre esta clase de pruebas da un vuelco: su decisión se basa exclusivamente en una prueba neurocientífica, disminuyendo la pena del sujeto por atenuantes de responsabilidad tras ver los informes genéticos que la defensa había puesto a su disposición. En concreto, el juez P. V. Reinotti se inclinó por estos ignorando las conclusiones del perito, las cuales afirmaban que Bayout (el acusado) poseía plenas capacidades de comprensión y conocimiento en el momento de comisión del delito. Así, argumentó a la luz de los informes que en determinados genes del imputado podían observarse alteraciones cromosómicas que lo hacían particularmente reactivo en términos de agresividad –y, consecuentemente vulnerable- ante situaciones inductoras de estrés²⁴, instando al Juez de la Audiencia Preliminar a que corrigiese su decisión inicial, aplicando la reducción máxima (un tercio) de la sanción por inexistencia parcial de imputabilidad²⁵.

Cabe destacar que esta no ha sido la línea dominante de la *Corte di Cassazione* (equivalente a nuestro Tribunal Supremo). Múltiples sentencias se han mostrado algo escépticas con el pleno reconocimiento de las pruebas neurocientíficas²⁶, debido en particular a la inexistencia de acuerdo científico sobre las interrelaciones reales (y sobre todo sobre el grado de las mismas) entre “(...) los aspectos morfológicos y biológicos y el componente voluntario de la conducta²⁷”. Sin embargo, la predisposición inicial de los tribunales a dar a

²² La literatura penal italiana cita además la sentencia 536 del Tribunal de Como, de 20 de mayo de 2011, igualmente pionera y con idénticos resultados.

²³ La trascendencia de este caso fue tal que Feresin (2009) publicó en la prestigiosa revista *Nature* un artículo en que afirmaba que esta era la primera vez que la genética del comportamiento había influido en la decisión de un tribunal europeo.

²⁴ Traducción propia. Fragmento original de la sentencia en Juliá (2020): “(...) determinati ‘geni’ presenti nel patrimonio cromosomico dell’imputato lo renderebbero particolarmente reattivo in termini di aggressività — e, conseguentemente vulnerabile— in presenza di situazioni di stress”.

²⁵ Este caso tuvo especial repercusión mediática en Italia -véase la crónica del diario *La Stampa*- (disponible en línea en <https://www.lastampa.it/cronaca/2009/10/25/news/e-vulnerabile-geneticamente-1.37057128/>) y también en España: véase *El País* (disponible en línea en https://elpais.com/sociedad/2009/11/02/actualidad/1257116411_850215.html)

²⁶ Con la honrosa excepción de la Sentencia 11897 de 18 de mayo de 2018, algo más aperturista hacia la neurociencia en un caso sobre un hombre con conductas agresivas tras sufrir un traumatismo craneal en un accidente de tráfico.

²⁷ Fragmento de la Sentencia 26895 de 12 de junio de 2018 de la *Corte di Cassazione*, traducción propia de Grandi (2022, p. 387). Para más jurisprudencia en la misma línea, véanse Sentencia 43021 de la Corte de 7 de noviembre de 2012 o más recientemente, Sentencia 27129 de 21 de julio de 2016, todas ellas de la Sección 1ª.

estas pruebas un mayor valor probatorio que en otros lugares ha suscitado e impulsado un debate que continúa hasta nuestros días.

3.1.1.2 La prueba neurocientífica en España: un caso paradigmático.

Por último, observaremos el tratamiento de estas pruebas a nivel nacional, analizando las sentencias del caso de Pioz (Guadalajara). El 17 de agosto de 2016, un hombre de nacionalidad brasileña mató a sus tíos y sus dos hijos, descuartizando los cadáveres para ocultarlos. Poco más de un mes después, decidió entregarse a la policía tras haber huido a su país de origen, confesando los crímenes y sometiéndose a juicio por un Jurado Popular. El interés de este caso radica en que fue presentada por la defensa del imputado una neuroimagen de su cerebro (PET-TAC), alegando que padecía un daño cerebral crónico, lo que les llevó a solicitar la eximente incompleta para él, con atenuante por arrebató. Los psicólogos del Instituto de Medicina Legal de Guadalajara concluyeron, por su parte, que actuó en plenas facultades.

La Sentencia original de la AP de Guadalajara (3/2018, de 15 de noviembre) fue apelada ante el TSJCLM y el TS, resultando en una agravación de la pena para el acusado, quedando fijada definitivamente en 25 años de privación de libertad y 3 penas de prisión permanente revisable. Lo que aquí pasaremos a analizar es el tratamiento que los tres órganos dispensaron a la prueba propuesta por la defensa.

El Jurado de la Audiencia Provincial, a tenor de lo observado en la neuroimagen, considera que no se puede acreditar que el acusado padezca o no un daño neurológico suficiente como para no atribuirle responsabilidad, puesto que “tales pruebas no son concluyentes al no haber un diagnóstico por parte de un neurólogo” (Fundamento de Derecho nº 4). Asimismo, descartan que sus comportamientos habituales de impulsividad e ira, típicamente psicopáticos y antisociales, hubiesen sido determinantes para la comisión del delito, puesto que planificó cuidadosamente el asesinato y ocultación de los cuerpos, incluso limpiando la escena del crimen para no dejar rastro: en definitiva, sabía perfectamente lo que estaba haciendo, por lo que no es posible eximirle de responsabilidad por falta de culpabilidad.

El TSJ (sentencia 16/2019, de 13 de junio), a pesar de comenzar citando algunas sentencias que reconocieron con éxito las pruebas neurocientíficas (STS 1079/2009, por poner un ejemplo), opta por seguir la misma línea que la AP, aseverando que el Jurado tan solo ha basado su decisión en otras pruebas igualmente válidas, sin desdeñar el valor que tenían las neurocientíficas –recuérdese que no se limitaron a rechazar el examen del PET-TAC, sino que efectuaron una valoración clara y completa de la neuroimagen, concluyendo que no era suficiente para declarar la no culpabilidad del acusado (Fundamento de Derecho nº 14). El TSJ añade a su vez que, si bien la existencia de trastornos psíquicos puede producir alteraciones en el comportamiento del sujeto, ello no implica que su capacidad cognitiva siempre esté afectada.

El TS (sentencia de 5 de mayo de 2020) subraya la relevancia que tienen las pruebas neurocientíficas en el proceso penal en la actualidad. Sin embargo, rechaza la pretensión del acusado de revocar el fallo de las sentencias reconsiderando su neuroimagen, habida cuenta de que lo que realmente hay que valorar es si los integrantes del Jurado se decantaron de forma suficientemente justificada por aceptar otras pruebas²⁸. El TS se alinea con lo expuesto por el TSJ aclarando que el Jurado “no despreció la existencia de esas imágenes neurológicas. Simplemente las consideró insuficientes para justificar el daño cerebral, al no haber podido contar con las explicaciones de un neurólogo (...)” (Fundamento de Derecho nº 4), desestimando el recurso de la parte apelante y dando por finalizado el proceso.

De todo lo anterior se desprende la adhesión de la jurisprudencia española a la corriente compatibilista, admitiendo la valoración de pruebas neurocientíficas (que ofrecen “incuestionables resultados”, según la sentencia del propio TS) junto con otras de carácter tradicional y optando por basar la decisión en aquellas que se consideren más oportunas en el caso concreto.

²⁸ El proceso llegó al TS aduciéndose una vulneración del derecho a un proceso con todas las garantías.

3.2 De la cualidad de imprescindibilidad del humano en la impartición de justicia y de cómo un robot podría mejorar al juez: la justicia predictiva.

Comenta Chiao (2018) que la justicia predictiva abarca una cuestión empírica (atribuyendo un nivel de riesgo de reincidencia al delincuente) y una cuestión empírico-normativa: qué consideraría el juez que debe ser una sentencia proporcionada, dados unos hechos, pruebas y circunstancias. Se trata de una cuestión “mixta” precisamente porque en la decisión final influyen valoraciones del magistrado sobre los hechos: el derecho es una ciencia social, por lo que no debe limitarse a aplicar una ley de forma literal. En lo que concierne a esta segunda dimensión, la Inteligencia Artificial (como producto estrechamente relacionado con la neurociencia) se ha ido abriendo paso como posible sustituto del juez²⁹.

Sin embargo, a pesar del potencial impacto que la IA pueda tener en el sistema judicial, conviene advertir que la justicia predictiva –o, al menos, el fin que persigue– lleva conviviendo con nosotros más tiempo del que creemos. Von Linstow (1974) menciona un algoritmo creado en Alemania en la época cuya pretensión no era sino obtener una cifra –la pena de prisión adecuada– para castigar a una persona ebria al volante, entre otros casos. La computadora analizaba diversos parámetros, como “nivel de alcohol en sangre” o “distancia recorrida”, que se introducían tras las pertinentes investigaciones. Ponderando cada uno de ellos, la máquina obtenía lo que para ella debía ser la pena proporcional para el sujeto³⁰. Aunque este método no tuvo ningún éxito real en la jurisprudencia alemana –asevera Köberer (1996) que no se acudió a él ni se citó su valor probatorio en una sola sentencia– sí supuso un primer paso para la entrada de la justicia predictiva en el sistema judicial, si bien no era propiamente IA, solo una mera calculadora (Greco, 2020).

²⁹ También del abogado: el diario El Norte de Castilla se ha hecho eco recientemente en una noticia del 12 de marzo de 2023 (Pascual Botrón, 2023) de las pruebas que el Ilustre Colegio de la Abogacía de Valladolid ha realizado con el software ChatGPT para agilizar trámites en los despachos, creando por ejemplo cartas amistosas de requerimiento de pago a un deudor, con resultados satisfactorios. No obstante, aún consideran inviable la sustitución por completo del abogado y del juez: en declaraciones a este mismo periódico, el titular del Juzgado de Instrucción nº1 de Valladolid José María Crespo de Pablo opina que estas “máquinas expendedoras” de sentencias aún no pueden gobernar la impartición de justicia.

³⁰ Esencialmente, este procedimiento no es tan diferente al que emplean herramientas como COMPAS en la actualidad. No obstante, el valor de estas últimas radica en la capacidad de análisis y cruce masivo de datos con los de otros procesos penales que ofrecen.

Todo esto lleva a autores como Grosse-Wilde (2022) a preguntarse si la justicia predictiva es “vino antiguo en botellas nuevas³¹” (p. 451), es decir, si tan solo se trata de un nuevo intento de conseguir el mismo objetivo que hace casi 50 años³², pero disfrazándolo con nuevos métodos (la IA, y en particular, el *machine learning*, un tipo específico de IA al que debe prestarse especial atención).

3.2.1 No es oro todo lo que reluce: críticas a la justicia predictiva

Las principales críticas dirigidas contra la justicia predictiva son tres: su más que posible colisión con los derechos humanos, la falta de transparencia de estos sistemas – extremos que desarrollaremos de modo más profuso en una sección aparte, junto con los desafíos que la neurociencia puede plantear en este sentido- y la dificultad que la IA tiene para distinguir entre correlación y causalidad.

Grosse-Wilde (2022) considera que el talón de Aquiles de la IA no es sino su incapacidad para saber con certeza si un hecho es la causa de otro o si simplemente está relacionado con este último. En esa misma línea, Marcus y Davis (2019) no niegan su asombrosa efectividad a la hora de establecer esas correlaciones de las que habla Grosse-Wilde, pero afirman tajantemente que carecen de lo que los lingüistas llaman composicionalidad: la comprensión del significado de una oración completa como unión del significado de cada una de sus partes^{33 34}. ¿Qué consecuencias tiene esto en la aplicación penal de la justicia predictiva? Por un lado, la IA no será capaz de detectar autónomamente la intencionalidad de un imputado a la hora de consumar un delito (para, en su caso, aplicar o no una atenuación de la pena), sino que obvia el contenido moral de un hecho punible para limitarse a buscar correlaciones en las pruebas (por ejemplo, atestados policiales que recojan el relato de los hechos) de las que se pueda inferir intención, sin que pueda comprender realmente el contexto de los hechos o la faceta psicológica del comportamiento humano.

³¹ Traducción propia. Texto original: *Old wine in new bottles?*

³² Autores españoles como Juliá Pijoan (2022) defienden lo propio, titulado una de sus contribuciones como “Neuropredicción: un Nuevo Ropaje para un Planteamiento Atávico” (p. 549).

³³ Traducción propia. Texto original: *Deep learning is very effective at learning correlations (...). It's missing what linguists call compositionality: a way of constructing the meaning of a complex sentence from the meaning of its parts* (p. 87).

³⁴ Esta carencia no es exclusiva de la IA. Merkel (2022) señala que también es propia de animales ciertamente inteligentes, capaces de traer una pelota concreta cuando se les indica que lo hagan, pero incapaces de comprender qué rol juega esa pelota en el transcurso de un partido de fútbol.

Caro Coria (2022) lo ilustra de un modo brillante en la máxima “la bondad, la justicia o la compasión no son programables” (p. 638).

Por otra parte, que un evento preceda a otro no implica que el primero sea la causa directa de este. Para muestra, un botón: cada vez que llueve, la hierba se moja. Sin embargo, que la hierba esté mojada no siempre implicará que haya llovido (puesto que puede haber sido causado, por ejemplo, por un aspersor). Este archiconocido ejemplo de la literatura sobre problemas de lógica³⁵ nos enseña que una correlación entre dos sucesos (“lluvia” y “hierba mojada”) no siempre implicará causalidad entre ellos. Trasladándolo por analogía a nuestro objeto de estudio, la IA podría malinterpretar una correlación (que, indudablemente, existirá) y tratarla como si fuese una causa, con consecuencias penales relevantes. A este respecto, Grosse-Wilde (2022) llama la atención sobre un caso en Alemania³⁶ cuyos hechos versan del siguiente modo: un traficante de drogas que ha ingerido las mismas para no ser descubierto es detenido y puesto en observación médica hasta que los efectos adversos de las sustancias desaparezcan, con el gasto en sanidad que ello conlleva. Si este supuesto se resolviese empleando la justicia predictiva, el sistema de IA observaría una correlación entre los costes de la sentencia y la culpa –pues, generalmente, los delitos más complejos requieren un mayor esfuerzo económico en recursos para ser perseguidos- proponiendo como resultado una pena más grave de lo que debería. Sin embargo, esos mayores costes de la investigación no cambian en absoluto el grado de culpabilidad del acusado, por lo que no deben repercutirse en la sanción impuesta.

Por el contrario, esta restricción de la IA también podría operar en sentido favorable al acusado. Si se pone en peligro la integridad física de un agente al tratar de huir por la comisión de un delito que nada tiene que ver con ese bien jurídico, la IA no lo considerará agravante de la pena, cuando en muchos casos sí sería constitutivo de ello. Otro ejemplo más: un asesino que actúe con particular ensañamiento³⁷, jugando con las emociones de la víctima momentos antes de cometer el crimen, se vería beneficiado de ser juzgado por una

³⁵ Este es un ejemplo de red bayesiana, muy empleada en modelos matemáticos. Para más información sobre el uso de las redes bayesianas en la optimización de algoritmos de *machine learning*, véase Snoek *et al.* (2012).

³⁶ Resuelto por el *Bundesgerichtshof* (equivalente a nuestro Tribunal Supremo) en Decisión de 16 de enero de 2010.

³⁷ Artículo 139.1.2ª del Código Penal: pena privativa de libertad de 15 a 25 años.

IA que no comprende el papel de los sentimientos y el sufrimiento psicológico en la comisión del hecho punible, considerándolo un acto homicida, con un castigo menor³⁸.

Dejando a un lado los postulados clásicos de la filosofía antigua (Aristóteles ya lo trató en *Segundos Analíticos*), Woodward y Ross (2021) trazan un recorrido histórico acerca del concepto de causalidad para desembocar en el *big data* como “el fin de la ciencia³⁹”. Es a partir del siglo XX cuando autores como Hempel (1942, 1965) desarrollan un modelo deductivo-nomológico relacionando un elemento con lo que denominan capacidad explicativa de otro distinto, de modo que si algo permite aclarar por qué ocurre un evento cuando ya ha acaecido, podemos anticiparlo también antes de que suceda. Este modelo claramente determinista dio paso a otro en el que la inferencia estadística era la protagonista: las causas a estudiar eran aquellas que incrementasen las probabilidades de que acontezca un hecho específico, por lo que se ampliaba el abanico de posibles causas.

Una tercera etapa se basa en la aplicación del análisis multivariante de datos, más sofisticado. Son particularmente relevantes a este respecto las reflexiones de Judea Pearl⁴⁰, firme defensor del determinismo (“no hay libre albedrío (...). La decisión se ha tomado en el cerebro antes”), que abandera el nuevo desafío de la introducción de la neurociencia en la sociedad desdibujando las diferencias entre ser humano e IA (“el hardware es diferente, pero el software es el mismo”). También asume su papel en el Derecho Penal, pudiendo resolver problemas de culpabilidad como “si Juan bloquea la salida de incendios de un edificio y María fallece por no poder acceder a ella, ¿Juan es responsable legal de su muerte, aunque no encendiera el fuego?”

Pearl es el inventor de la “escalera de la causalidad⁴¹”. Este filósofo e ingeniero reconoce que la IA aún se encuentra en el primer escalón de esta pirámide, pues solo “crean asociaciones entre lo que fue observado antes y lo que será observado en el futuro”. Todavía no pueden “predecir acciones que no se han llevado a cabo nunca en las mismas

³⁸ Artículo 138.1 del Código Penal: pena privativa de libertad de 10 a 15 años.

³⁹ Anderson (2008).

⁴⁰ Las declaraciones reflejadas en este trabajo pertenecen a sendas entrevistas con Ana Tagarro (Tagarro, 2022), de *XL Semanal* (véase la bibliografía para más detalles), si bien es cierto que las aportaciones académicas de este pensador al campo de la IA son ingentes. Por ello, nos remitimos a Pearl y Verna (1995), Pearl (2009, 2010) y Pearl *et al.* (2016).

⁴¹ También en Pearl y Mackenzie (2018).

condiciones” (segundo escalón), ni en qué situación estaría algo si otro evento en el pasado hubiese ocurrido (tercer escalón). Pearl ha creado un intrincado modelo matemático que permite precisamente afrontar este último problema, usando lo que Pearl denomina imágenes contrafactuales (contra los datos), pues “los datos, por muchos que sean, no pueden decirnos qué ocurriría en un mundo imaginario”. Pearl admite que la IA aún no ha llegado a este punto: el *deep learning* aún no tiene la respuesta a la pregunta “¿y si hubiera hecho...?”. Cuando se desarrolle por completo, ese sería el fin del método científico de comprobación de la causalidad como lo conocemos.

3.2.2. Después de todo, ¿ha perdido la partida la justicia predictiva?

Miró Llinares y Castro Toledo (2022) no expresan tanto su preocupación por que la neurociencia amenace los pilares fundamentales de la atribución de responsabilidad sobre los que descansa el Derecho Penal, sino más bien por que su generalización nos lleve a decir mal el derecho, o al menos, peor que actualmente. Otros autores se muestran más optimistas poniendo el foco precisamente en este extremo, como Pastor (2022), que sugiere que no hay que tener miedo a cambiar los fundamentos del Derecho, si ello sirve para hacer justicia de modo distinto, pero mejor. En similares términos se pronuncia Weinberger (2018): “la IA nos obliga a revisar nuestra idea de justicia”; en definitiva, de interpretar el Derecho con las mismas metas, pero de forma diferente y mejor.

En este sentido, Pastor (2022) contrapone la llegada de la IA al sistema penal a un envío de tecnología actual a los juicios de la Antigua Grecia, pronosticando que se servirían de estos medios (más eficientes) para apoyarse en el procedimiento, pero el núcleo del proceso seguiría siendo el mismo. En opinión de este autor, el pensamiento generalizado del grueso de juristas y no juristas es que la introducción de la IA tendría los mismos efectos (auxiliar a las partes para lograr mayor eficiencia), cuando no debería ser así; la IA cambiará radicalmente el sistema penal, consiguiendo que sea “algo mucho mejor” (p. 538).

Para ilustrar su postura, llama la atención sobre un ejemplo de procedimiento sancionador que, si bien obvia todas las exigencias de un proceso al uso, está plenamente aceptado por la población por ser más eficaz a pesar de hacer justicia ‘de modo diferente’. Los radares⁴² detectan la velocidad a la que se desplazan los vehículos mediante un sistema

⁴² Véase también Haissiner y Pastor (2019).

de ondas en un segmento concreto de la vía; si superan un límite prefijado⁴³, toman una fotografía del vehículo y emiten una sanción⁴⁴ que es comunicada al infractor. Así, Pastor indica que el cinemómetro genera instantáneamente “un proceso de primera instancia completo (...). Es ella sola [la máquina] denunciante, policía, testigo, fiscal, tribunal sentenciador y (...) juez de ejecución (...)” aunque queden al margen “separación de funciones entre requerir y decidir, división de etapas entre investigar y juzgar, exclusión como juzgadores de denunciantes, testigos y peritos, procedimiento oral y público (...)” (pp. 541-542). A pesar de ello, el público lo asume como una forma más efectiva de controlar los excesos de velocidad, pues sería impensable situar a un agente en la margen de la vía para que midiese las velocidades y detuviese a los conductores que superan el límite establecido. Aunque los radares no incluyen (al menos por ahora) IA, no hay por qué temer, en opinión de Pastor, el impacto de la IA en la forma de impartir justicia a la luz de los argumentos expuestos.

Para cerrar este apartado sobre la IA y su posible compatibilidad con el Derecho Penal, concluiremos con una elocuente cita de Melanie Mitchell (Roberts, 2020), que resume en cierto modo el parecer de muchos compatibilistas para reflexión del lector:

“(...) La IA funciona cada vez mejor (...). Sin embargo, cuando las personas ven un sistema de IA como Siri, por ejemplo, tienden a antropomorfizarlo. Tiene un nombre, tiene una voz: casi tiene personalidad. Solemos darle mayor mérito del que merece por pensar o ser inteligente o compasivo, y esa es una reacción muy humana⁴⁵”.

⁴³ Sin olvidar los márgenes de error de los radares fijos y móviles.

⁴⁴ Dejamos a un lado la eventual responsabilidad penal del conductor por un delito contra la seguridad vial del Capítulo IV del Título XVII del Código Penal (arts. 379-385 ter).

⁴⁵ Traducción propia. Texto original: *(...) the better AI works (...). But also, I think people often, when we see an AI system like Siri, for example, we tend to anthropomorphize it. It has a name, it has a voice: it almost has a personality. We tend to give it more credit than it actually deserves for thinking or being intelligent or understanding, and that's a very human reaction.*

4. NEUROCIENCIA Y POLÍTICA CRIMINAL

El ámbito penal de mayor proyección para las diferentes herramientas que la neurociencia nos proporciona es quizás a su vez el de mayor aplicación práctica en la actualidad, con sistemas plenamente operativos en lugares como Estados Unidos o España. Estamos hablando de la Criminología y, en particular, de la Política Criminal, donde estos instrumentos pueden emplearse para, entre otras cosas, evaluar el riesgo de reincidencia de un delincuente o aplicar medidas de seguridad apoyándose en los pronósticos de estas máquinas.

4.1 Un inesperado y poderoso aliado en la prevención del delito

Con carácter previo a detallar el abanico de oportunidades que nos brinda la neurociencia en este campo, cabe detenerse a analizar el fundamento de la pena y el cambio que ha supuesto en este sentido la irrupción de la neurociencia en el sistema penal. En sus orígenes, el Derecho Penal fijaba las penas basándose en el principio de proporcionalidad: ya en el Derecho Romano⁴⁶ la Ley del Talión (ojo por ojo) se erigió como máxima expresión de este criterio, que defendía que el delincuente debía ser castigado con una pena equivalente y proporcional al daño realizado. Su imposición tenía esencialmente un fin de prevención general, orientado a advertir a la población general de las consecuencias que podía tener el quebrantamiento de la ley.

La aparición de la Criminología a mediados del siglo XIX centra su análisis en la rehabilitación del reo para evitar que vuelva a delinquir (prevención especial), lo que en términos agregados comportaría una reducción generalizada de la delincuencia. Rodríguez (2022), analizando la obra de Röder (1870), rechaza su teoría de que el fin del correccionalismo fuese rehabilitar el libre albedrío, cuya existencia o no ya hemos desarrollado en el capítulo anterior. Hoy en día, la neurociencia ha ampliado las miras del Derecho Penal en esta dirección: a partir de ahora es posible la cooperación entre juristas y médicos para decidir una pena en base a este factor de reincidencia, empleando novedosas herramientas para inclinar la balanza hacia el lado de la prevención especial.

⁴⁶ Podemos remontarnos incluso a momentos anteriores: el Código babilónico de Hammurabi, promulgado en el lejano 1755 antes de Cristo, ya contenía normas punitivas que seguían este criterio.

En la actualidad, numerosos autores abogan por adoptar una posición de síntesis: Feijoo Sánchez (2022) afirma de manera rotunda que el Derecho Penal no puede ni debe limitarse a la prevención especial; en definitiva, se trata de “potenciar las posibilidades preventivo-especiales en una pena impuesta por motivos de prevención general” (Feijoo Sánchez, 2017, p. 33⁴⁷). De hecho, Rodríguez (2022) indica que la propia inocuización de un delincuente (tomando medidas concretas y adaptadas al sujeto para evitar su reincidencia) como expresión de la prevención especial tiene a la vez efectos de prevención general, puesto que los potenciales delincuentes han podido observar qué consecuencias tiene la comisión de ese hecho punible, pudiendo disuadirles de la consumación de un delito en el futuro. Según Feijoo Sánchez (2017), la solución consistirá en aplicar al sujeto una pena proporcional al hecho con complementos dirigidos a la prevención.

Es en este último extremo donde las herramientas neurocientíficas pueden ser de gran ayuda para reforzar el sistema penal, evaluando la situación mental del sujeto y ofreciendo un seguimiento tanto durante el cumplimiento de su condena como tras su liberación. No obstante, las contribuciones de la neurociencia no quedan ahí; los informes médicos pueden aconsejar la sustitución de la pena por unas medidas de seguridad amparándose en los resultados arrojados por diversas pruebas neurocientíficas, haciéndola completamente innecesaria. Burns y Swerdlow (2003) mencionan el caso de un hombre que desarrolla intereses pedófilos, razón por la cual es condenado por acoso a niñas de corta edad. Tras practicársele un escáner cerebral a causa de los agudos dolores de cabeza que sufría, descubren un tumor del que es operado, momento a partir del cual su tendencia pedófila se reduce. Un tiempo después, vuelven las jaquecas y observan cómo su tumor ha crecido, y con él sus intereses pedófilos vuelven a aflorar. En realidad, la aportación de la neurociencia fue crucial, ya que se observa la falta de culpabilidad del sujeto: sus comportamientos estaban directamente relacionados con la aparición del tumor, pues desaparecieron tras ser extirpado y resurgieron junto con él: en resumen, lo razonable sería procesar al sujeto aplicando medidas de seguridad y no una pena.

Distinta es la cuestión acerca de sobre quién pesa la responsabilidad en caso de reincidencia de un condenado a medidas de seguridad aplicadas en base a resultados obtenidos con técnicas neurocientíficas. Según Rodríguez (2022), no puede hacerse

⁴⁷ En referencia a la opinión -algo más moderada con el tiempo- de Roth sobre el debate acerca del determinismo (Roth y Merkel, 2008).

responsable de la comisión del delito al sujeto reincidente, pues fueron los informes médicos los que aconsejaron la sustitución de la pena por una medida concreta señalando los pasos a seguir para su tratamiento, dejando la puerta abierta al debate sobre este extremo⁴⁸. Este penalista brasileño concluye que aún son necesarias ambas facetas preventivo-general y preventivo-especial de la pena, pues solo garantizando la no-repetición del sujeto se podría prescindir de la prevención general.

4.2 Distintos caminos, un mismo propósito: herramientas neurocientíficas en la rehabilitación del delincuente

¿Cómo puede contribuir la neurociencia a la aplicación de la consecuencia penal más apropiada? Con carácter previo a la emisión de la sentencia, se puede efectuar un diagnóstico estrictamente médico del sujeto, para poder aplicar medidas de seguridad tendentes a su corrección. Rosburg *et. al.* (2020) observan cómo es posible someter a un detenido por indicios de pedofilia a un procedimiento de neuroimagen por fMRI, en el cual se le van mostrando imágenes de adultos y niños alternativamente y se mide su respuesta corporal a dichos estímulos, registrándose un incremento anormal de oxígeno en la amígdala -parte del cerebro clave en la gestión emocional y producción de testosterona- en los sujetos cuando se les mostraban imágenes de menores⁴⁹. De esta forma, es posible emitir un diagnóstico fiable que permita adoptar las medidas más oportunas.

Naturalmente, una de las posibilidades que la neurociencia ofrece para la rehabilitación de estos delincuentes es ceñirse a un tratamiento estrictamente médico. Por ejemplo, es posible suministrar fármacos con principios activos que actúen directamente sobre la función

⁴⁸ Se hace una mención expresa a la posibilidad de responsabilizar a los médicos que avalaron el tratamiento alternativo a la pena, o bien al Estado como ente que admitió las conclusiones de los informes neurocientíficos. Un argumento de peso para descartar la primera alternativa tiene que ver con la propia naturaleza de las obligaciones contraídas por estos profesionales, que son de medios y no de resultados. Esta cuestión ha sido ampliamente estudiada por civilistas como Domínguez Luelmo (2007). Los tratamientos médicos individualizados no siempre tienen la misma eficacia en todos los pacientes, y en muchas ocasiones no tienen éxito, con consecuencias adversas para el sujeto que pueden suponer incluso la muerte. Por este motivo, el médico no está sujeto a la obligación de curar al paciente, sino de hacer todo lo que esté en sus manos para conseguirlo. Esta tesis está refrendada por la reiterada doctrina del Tribunal Supremo al respecto (véase, por todas, STS 6/2008, de 16 de enero de 2008, en la que se rechaza la petición de indemnización por fallecimiento de una mujer, habida cuenta del buen proceder del equipo médico). Por ello, resulta dudosa la atribución de responsabilidad a este colectivo por los infructuosos resultados de un tratamiento neurocientífico.

⁴⁹ Esta clase de pruebas no está exenta de cierta polémica entre los penalistas y procesalistas, por su posible contradicción con el principio *nemo tenetur se ipsum accusare* (nadie tiene por qué declararse a sí mismo culpable). A todas estas cuestiones nos hemos referido ya en el capítulo anterior.

cerebral (Glenn y Raine, 2013), a fin de inhibir determinadas conductas y potenciar otras, tendentes a la reducción de los impulsos delictivos del sujeto. En esta línea, estudios como los de Pappadopoulos *et al.* (2006) subrayan la eficacia de antidepresivos para reducir la agresividad en niños y adolescentes, mientras que New *et al.* (2004) prescriben inhibidores de serotonina para incrementar la metabolización de la glucosa en la corteza orbitofrontal y reducir la agresividad de adultos⁵⁰. Otra alternativa, orientada a condenados por delitos sexuales, es someterlos a un tratamiento de castración química que reduzca el deseo sexual de estos sujetos (Lösel y Schmucker, 2005)⁵¹.

¿Qué métodos se utilizan en la conocida como Neurocriminología para efectuar todas estas valoraciones? En primer lugar, es posible elaborar un perfil neurobiológico asociado a la existencia o no de determinados rasgos en el sujeto en cuestión, denominados biomarcadores, entre los que se incluyen factores físicos como la anatomía del cerebro (con posibles anomalías en determinadas zonas relacionadas con la sociabilidad del sujeto⁵²) o la secuenciación particular de determinados genes con una composición peculiar en los cromosomas. Estos exámenes deben combinarse con análisis psicosociales que permitan indagar en profundidad en su percepción de lo que le rodea, tratando de comprender por qué se comporta de este modo. Mediante la práctica de pruebas para la averiguación de esos caracteres presentes en el sujeto, podemos crear un perfil dinámico (González Tapia, 2022) con una faceta retrospectiva, que logre explicar por qué se produjo el hecho punible (lo más difícil de conseguir para relacionarlo con la imputabilidad del sujeto en el momento de comisión del delito, según González) y otra prospectiva, para tratar de anticipar la conducta futura del sujeto.

⁵⁰ En relación a la serotonina, Glenn y Raine (2013) comentan dos debilidades: por un lado, la existencia de estudios (véase Coccaro, 1989) que sugieren precisamente lo contrario, es decir, que los bajos niveles de serotonina son la causa de la agresividad, y por otra parte, la incógnita acerca del efecto de estos medicamentos a largo plazo.

⁵¹ A pesar de su posible colisión con los derechos fundamentales, la castración química está aceptada en 9 estados de Estados Unidos a abril de 2023: California, Florida, Iowa, Georgia, Luisiana, Montana, Oregón, Texas, Wisconsin y Alabama.

⁵² A modo de ejemplo, Kolla *et al.* (2017) indican algunos biomarcadores asociados a la psicopatía, como anomalías en la forma de la amígdala y la corteza orbitofrontal o polimorfismos de la monoamino oxidasa-A, un gen relacionado con los comportamientos violentos, que tenía como resultado una menor respuesta al miedo.

La elaboración de perfiles neurobiológicos ha sido duramente criticada por varios sectores de la doctrina. Por un lado, Juliá Pijoan (2022) censura la contradicción inherente de este procedimiento, pues los perfiles no consideran la capacidad de cambio del cerebro a lo largo de los años. Los biomarcadores se refieren a características biológicas del sujeto que forman parte de él e incrementan su probabilidad de cometer delitos. Algunos de ellos tienen una naturaleza altamente heredable⁵³: Raine (1993) realiza un estudio en el que analiza el comportamiento de gemelos univitelinos (procedentes del mismo óvulo) separados a temprana edad, con unos resultados aplastantes: se observó que, si uno de ellos mostraba signos biológicos de conducta antisocial, el otro también lo haría en un 50% de los casos estudiados⁵⁴. Sin embargo, otros factores se ven claramente modificados por circunstancias del entorno del sujeto en su etapa de desarrollo—Romeo Casabona (2022) llega a afirmar que la personalidad del ser humano no la determinan sus genes, sino el entorno, y otros autores respaldan esta tesis—. Por lo tanto, se elaboran perfiles considerando como dados unos rasgos que no varían con el tiempo. Sin embargo, a lo largo de su vida el cerebro va creando nuevas conexiones en su interior que modifican el entramado neuronal, lo que ha llevado a los científicos a atribuirle la característica de neuroplasticidad (Sale *et al.*, 2014), gracias a la cual este órgano puede adaptarse a las circunstancias incorporando a la memoria los resultados del aprendizaje.

Otra de sus críticas es la asunción de la existencia de un cerebro normal, a partir de cuyas desviaciones se pueden empezar a plantear casos anómalos que deben ser tratados. Esto aviva un debate acerca de qué características debe tener un cerebro para considerarse normal, surgiendo el problema de la poca variabilidad de los sujetos en muestras de estudios destinados a ofrecer una respuesta a semejante cuestión. El penalista catalán se fija en Henrich *et al.* (2010) para afirmar que casi todos los sujetos que participaron en la creación del modelo de cerebro normal pertenecían a la sociedad occidental, educada, industrializada, rica y democrática, por lo que los resultados serían poco representativos.

⁵³ Muy consistentes, además, con independencia del género o la etnia (Ferguson, 2010).

⁵⁴ Esto no significa que el autor británico ignore o minusvalore el valor de los factores sociales y medioambientales en las estimaciones de peligrosidad y reincidencia, sino todo lo contrario: en el año 2011, Raine (2011) presentó los resultados de un estudio en el que se analizó el historial y tendencia a delinquir de jóvenes de 23 años que habían sido sometidos a un examen psicobiológico a los 3 años, criados en distintos ambientes. Raine demostró que el empleo únicamente de variables sociales -como el nivel socioeconómico de la familia o el trato de los padres hacia los niños- para predecir la criminalidad futura de estos ofrecía un nivel de precisión del 73% (comparándolo con los datos de criminalidad reales de estos jóvenes). Sin embargo, al combinarlo con antecedentes de carácter biológico, la fiabilidad ascendió hasta el 88,5%. González Tapia (2022) también defiende la compatibilización de ambos criterios.

Consiguientemente, la inclusión de un delincuente en un perfil tipo estaría injustificada por los sesgos que puede arrastrar la creación del modelo y por la maleabilidad de su cerebro, que puede modificarse hasta parecerse a uno normal, en cuyo caso debería salir de ese perfil, con las dificultades que ello conlleva para su reubicación, en particular si se le ha adjudicado un tratamiento basándose en ello. Redondo Illescas (2008) ofrece una solución a este problema: clasificar los factores de riesgo (o los biomarcadores que predecirán la reincidencia) a la hora de la evaluación en históricos (aquellos que no se pueden modificar, especialmente útiles en la prevención a largo plazo) y dinámicos, que son cambiantes y de corta duración.

Por otro lado, Juliá Pijoan (2022) critica que la mera observación de un determinado rasgo en el cerebro del sujeto lleve aparejada la atribución de la condición de delincuente, cuando no necesariamente es la causa de ello, puesto que puede haber personas con esa misma anomalía que no hayan delinquido nunca. Esta confusión entre causa y correlación ha abierto un amplio debate en cuanto al empleo de la IA en el sistema judicial, aspecto que ya hemos desarrollado.

Por último, la elaboración de perfiles puede verse sesgada de manera inconsciente debido a las propias exigencias del método estadístico que rige todo su proceso de creación (Miró Llinars, 2018, y Martínez Garay, 2022). Habida cuenta de la heterogeneidad de la sociedad y del colectivo de delincuentes y la presencia de los mismos en diferentes proporciones, es difícil esperar que todos los perfiles figuren igualmente representados. De esta forma, puede haber algunos infrarrepresentados y otros sobrerrepresentados (Martínez Garay, 2022), en atención a su mayor o menor preponderancia en los dos ámbitos mencionados. Sumando este factor al hecho de que los datos sobre criminalidad proporcionados por entidades policiales no pueden ser plenamente fiables (bien porque haya delincuentes que logren evadir la acción de la justicia o porque se cometan delitos sin que nadie llegue a saber de su existencia), se podría distorsionar la validez de estos perfiles criminológicos.

Otra posibilidad, de corte más psicológico, es introducir al delincuente en programas cognitivo-conductuales⁵⁵. Este tratamiento parte de la consideración de que una persona con

⁵⁵ Véase, por todos, Pearson *et al.* (2002).

biomarcadores típicos de conducta antisocial es más proclive a dejarse llevar por ellos en una situación de estrés, lo que podría desembocar en una conducta impulsiva que lleve a la comisión de un hecho punible. Mediante la participación en terapias dirigidas a aumentar el autocontrol o desarrollar habilidades sociales y emocionales, es posible combatir estos factores que predisponen al individuo hacia una conducta antijurídica.

Una tercera herramienta que se puede emplear con fines preventivos es la IA y otros mecanismos análogos. Es lo que Chiao (2018) categoriza como la parte empírica de la justicia predictiva⁵⁶, y lo que Miró Llinares (2018) denomina Inteligencia Artificial Policial⁵⁷. Para este último autor, las aportaciones de la neurociencia al ámbito preventivo penal trascienden la elaboración de perfiles que tratábamos anteriormente (eso sí, sin desdeñar su importancia) para auxiliar a los cuerpos policiales en la propia persecución del delito. En este sentido, contamos hoy en día con una cantidad masiva de datos brutos que las fuerzas y cuerpos de seguridad pueden usar para anticipar la comisión de un delito, desde imágenes tomadas por cámaras de videovigilancia hasta la geolocalización emitida por dispositivos móviles. Sin embargo, a la recogida y conservación de todos estos datos le acompaña la adecuada selección, procesamiento, filtrado y tratamiento de los mismos, lo que se plantea como una tarea ciertamente abrumadora para los agentes. Es ahí donde la IA, en particular los algoritmos de *big data*, pueden entrar en juego para aprovechar el potencial que pueden ofrecer y anticiparse a la comisión de un hecho punible con mayor vigilancia en la zona en la que se piensa que puede ocurrir algo⁵⁸.

Por último, Miró Llinares (2018) señala otro ámbito (quizás menos evidente a simple vista) en el que puede utilizarse el *big data* por la policía: los ciberdelitos, tanto en la faceta de su persecución, con la identificación de vulnerabilidades en las barreras de protección como los antivirus, como de su prevención. Por poner un ejemplo, el cruce de mensajes en redes sociales y otros datos puede detener la radicalización de un sujeto antes de que sea demasiado tarde.

⁵⁶ La parte empírico-normativa se refiere al uso de IA dentro del propio proceso penal, a lo que ya dedicamos un espacio.

⁵⁷ En contraposición a la Inteligencia Artificial Judicial, que se identifica con lo que los autores suelen considerar justicia predictiva en sentido estricto.

⁵⁸ E igualmente predecir el lugar donde un delincuente va a reincidir: Rossmo (1995) plantea un algoritmo capaz de indicar en qué zona geográfica va a atacar un asesino en serie basándose en las localizaciones de delitos anteriores.

4.3 La Inteligencia Artificial en acción: evaluaciones sistemáticas de riesgo

Como indicábamos al inicio de este capítulo, este ámbito preventivo de la neurociencia ya se ha generalizado en la práctica del día a día. Al contrario que la justicia predictiva - entendida como el uso de IA para auxiliar (¿e incluso sustituir?) las tareas de un juez- llegando a producir una sentencia absolutoria o condenatoria basándose en las pruebas presentadas⁵⁹, que aún no está plenamente desarrollada, las herramientas preventivas dirigidas a la evaluación del riesgo de reincidencia y a la rehabilitación del delincuente sí gozan de cierta aceptación en nuestros ordenamientos. De esta forma, pasaremos a analizar el uso de varios instrumentos predictivos, tanto a nivel nacional como internacional, y dirigidos a distintos ámbitos, como el procesal penal, policial y penitenciario.

4.3.1 Riesgo de reincidencia y proceso penal: COMPAS

COMPAS (siglas de *Correctional Offender Management Profiling for Alternative Sanctions*) es un sistema desarrollado en 1998 por la empresa norteamericana Northpointe (ahora con el nombre de Equivant), usado desde entonces en Estados Unidos para evaluar el riesgo de reincidencia de un encausado⁶⁰, en pos de aplicar la medida sancionadora más apropiada para su caso, o bien de un condenado, para plantear la forma de ejecución de la pena o incluso imponer medidas de vigilancia tras cumplir condena. Este algoritmo procesa información detallada acerca de las circunstancias personales del sujeto, como la edad, el empleo, su residencia o su historial delictivo y, cruzándolo con datos recogidos anteriormente mediante un análisis ponderado de las variables, se obtiene un informe que asesora al juez sobre diversos extremos, como por qué ha llegado a cometer el delito o el riesgo de reincidencia del inculpaado.

COMPAS está formado por dos tipos de herramientas (Guía del Facultativo para el uso de COMPAS, 2019), a saber, las escalas de necesidades y las escalas de riesgo. Las primeras se limitan a describir al individuo, sin la intención de predecir su conducta futura, basándose en la medición de diversos factores criminógenos. Como colofón, clasificarán al imputado en una de 8 tipologías distintas para asesorar al juez sobre las medidas que se deben

⁵⁹ Ámbito al que ya hemos dedicado un espacio.

⁶⁰ Parte de la doctrina identifica a COMPAS como IA. Sin embargo, sus características son más bien propias de una herramienta de valoración del riesgo (Miró Llinares y Castro Toledo, 2022).

tomar para su rehabilitación. Por su parte, las escalas de riesgo pretenden diferenciar qué encausados tienen probabilidad real de reincidir y cuáles no.

Existen cuatro escalas de riesgo diferentes: la escala de libertad antes del juicio (*Pretrial Release Risk Scale*) pretende averiguar si un encausado puesto en libertad provisional hasta la citación para el juicio tiene un alto riesgo de fuga -poniendo en peligro la celebración del mismo- o incluso de reincidir durante ese periodo. Nótese que el algoritmo no sugiere directamente la imposición de una medida cautelar determinada para asegurar su presencia en el juicio; sencillamente se limita a ofrecer evidencia empírica del riesgo, quedando en manos del juez la decisión final de qué hacer a la luz de los resultados. El segundo modelo es una escala de reincidencia general (*General Recidivism Risk Scale*), que usa información de carácter estático (historial delictivo, edad del primer arresto...) ⁶¹ para predecir la reincidencia en los próximos 2 años desde su sometimiento a los análisis de COMPAS. Una tercera escala es la de reincidencia violenta (*Violent Recidivism Risk Scale*), similar a la segunda, pero ponderando criterios relacionados concretamente con la violencia y no tanto con la reincidencia genérica. Equivant menciona una cuarta, la pantalla de riesgo de reincidencia (*Recidivism Risk Screen*), una suerte de versión simplificada de las dos últimas escalas para la sustitución por medidas correccionales en encausados por delitos menos graves.

Por otro lado, las escalas de necesidades criminógenas (*Criminogenic Need Scales*), cuyos resultados se emplean en el asesoramiento sobre el tratamiento más apropiado para cada delincuente, se basan en criterios como el comportamiento cognitivo, la amistad con otros criminales, su involucración en el delito, su nivel de violencia, problemas financieros o incluso su tiempo libre. El tratamiento concreto de cada variable se recoge en otro documento de la compañía, fechado en 2009, que incluye preguntas realizadas a los delincuentes e interpretaciones de los posibles resultados.

En este sentido, COMPAS reseña la importancia de crear lo que ellos llaman “tipologías”, es decir, categorías de individuos basadas en sus patrones de comportamiento con una serie de pasos a seguir de cara a su correcta rehabilitación en la sociedad. Aun asumiendo que cada persona responderá de manera diferente a un mismo tratamiento y que nadie encajará a la perfección en los estándares de un tipo, los creadores de COMPAS lo

⁶¹ Los propios diseñadores reconocen la validez de los indicadores dinámicos, admitiendo la necesidad de combinar ambos métodos para obtener resultados más fiables, aunque en COMPAS predominen claramente los indicadores estáticos.

consideran un punto de partida que después puede adaptarse a las circunstancias particulares de cada sujeto, y nunca como una clasificación de riesgo en sí misma. En el caso de los hombres, existen 8 categorías distintas: drogadictos crónicos (que suelen ser imputados por delitos no violentos), situacional de bajo riesgo (con buenas perspectivas económicas, habitualmente realizan delitos de acoso o agresión doméstica), alcohólicos crónicos (los de mayor edad, relacionados con delitos de violencia doméstica y delitos contra la seguridad vial), marginados sociales (con circunstancias socioeconómicas adversas, delinquen de forma habitual, recomendándose tratamientos psicológicos), los criminalmente versátiles (jóvenes y asociados a bandas callejeras o pandillas), los aislados socialmente y adictos a sustancias a largo plazo (muy proclives a la comisión de delitos), individuos de riesgo alto, serio y versátil (con una tendencia crónica a cometer delitos graves, como robos y homicidios), y los de bajo riesgo “accidental” (una subcategoría de los de bajo riesgo, comprendiendo a aquellos que caen en la criminalidad fruto de una circunstancia casual, con un perfil social normal y encausados por delitos menos graves).

Por su parte, las tipologías femeninas son personas con problemas en relación a las drogas y tendencia antisocial, procedencia de familias desestructuradas (con traumas asociados a un mal cuidado por parte de los padres, que suelen ejecutar delitos leves y no violentos), adicción a sustancias como el alcohol (de mayor nivel económico, proclives a los delitos contra la seguridad vial), mujeres mayores, pobres y marginadas (con delitos cometidos para subsistir económicamente), mujeres jóvenes y que no han disfrutado de un sistema educativo (que comienzan a delinquir de manera temprana, a veces con conductas violentas), delincuentes reincidentes con largos historiales delictivos (con dos subcategorías: aquellos afectados por múltiples factores sociales y psicológicos de forma que aumenta el riesgo, y otra modalidad con las mismas causas pero presentadas de modo más intenso) y por último, las delincuentes tardías, de mayor nivel socioeconómico y con menores factores de riesgo.

El avezado lector habrá detectado entre estos párrafos varias coincidencias con lo expuesto en la sección anterior. Efectivamente, el funcionamiento de COMPAS se basa en la elaboración de perfiles (sus “tipologías”), basándose en criterios estáticos y dinámicos (que, sin ser biomarcadores como tales al no centrarse solamente en aspectos biológicos del sujeto, sí se pueden asimilar a ellos por cumplir la misma función). En este sentido, la propia compañía incide en su Guía en que la valoración del riesgo trata de predecir el

comportamiento de grupos, no a título individual: la puntuación obtenida (con el consiguiente nivel de reincidencia atribuido) se estima basándose en resultados obtenidos anteriormente por sujetos con características similares a las del encausado.

Esto no es una cuestión baladí: por una parte, se garantiza un análisis de cierto rigor, puesto que el destino de la persona no está sometido únicamente a sus características o comportamientos (de hecho, el delito por el que se le juzga no se incluye en la valoración del riesgo por parte del sistema, teniendo solamente en cuenta el historial delictivo a estos efectos). Sin embargo, esta categorización de individuos en grupos con características comunes ha llevado a numerosos autores a criticar esta herramienta por sus sesgos hacia colectivos minoritarios, en particular hacia los de raza negra (Angwin *et al.*, 2016). No será este el lugar donde demos respuesta a las críticas que han suscitado las herramientas neurocientíficas tanto en el ámbito preventivo-penal como en el procesal, sino que lo reservaremos para más adelante. Baste decir, por ahora, que la doctrina⁶² pone en duda la eficacia de estos sistemas por los sesgos en su programación o su falta de transparencia: son públicos los datos que recoge el algoritmo y la forma de tratarlos, pero la interpretación y procesamiento de los datos quedan protegidos por el secreto industrial. A pesar de las explicaciones dadas por la empresa⁶³, con estudios avalando la fiabilidad tanto de su habilidad predictiva del riesgo de reincidencia (Singh, 2013) como de la utilidad de los indicadores descriptivos (Brennan *et al.*, 2009), estas no convencen a un sector ciertamente numeroso de la doctrina (Martínez Garay, 2018, o González Tapia, 2022).

4.3.2 *¿Predecir la escena del crimen antes de que ocurra? Geolitica*

Otro de los sistemas empleados en el continente norteamericano es Geolitica (antes conocido como PredPol⁶⁴) diseñado en 2012 para satisfacer las necesidades de las distintas unidades de policía a lo largo y ancho de Estados Unidos y otros países. Al contrario que COMPAS, cuyo uso predominante es en el seno del proceso penal, Geolitica (combinación

⁶² También la jurisprudencia se ha pronunciado sobre este aspecto, como veremos posteriormente con el caso *State vs. Loomis* de Wisconsin.

⁶³ La publicación de la Guía del Facultativo se produjo como respuesta de Equivant a las críticas sobre la falta de transparencia de COMPAS, en un intento por no dañar su imagen.

⁶⁴ El cambio de nombre surgió precisamente como intento de reposicionamiento de marca desde la faceta predictiva hasta la evaluación de riesgo.

de *Geographical Analytics*) es de especial utilidad para detectar en qué zonas geográficas y momentos del día hay una mayor probabilidad de que se cometa un crimen específico, pudiendo asignar una mayor vigilancia como medida de disuasión y así prevenir su comisión. Su fin último es reducir la criminalidad y mejorar la imagen de los servicios de seguridad pública de cara a la sociedad. Para ello se sirve fundamentalmente de herramientas de geolocalización de agentes y delitos, así como de datos sobre crímenes anteriores.

La compañía basa su algoritmo en tres propiedades: transparencia, rendición de cuentas y efectividad⁶⁵. En lo que a la transparencia se refiere, Geolítica proporciona informes que recogen interacciones entre todo tipo de datos cruzados, como tipos de crímenes y distritos, y que se muestran en gráficos sencillos de entender para el público general. Algunos de los informes son mapas de calor de la actividad delictiva –filtrando según la hora del día o el tipo de delito y pudiendo ver qué zonas están sobrevigiladas y cuáles infravigiladas-, informes de los agentes -controlando el movimiento individual de los mismos también en los cambios de turno-, análisis de puntos calientes (detectando a qué áreas dedican más atención los oficiales), informes sobre misiones de las patrullas (para aumentar la confianza de la sociedad en el sistema) e informes usando COMPSTAT, que realiza comparaciones entre los datos registrados a nivel histórico. La rendición de cuentas implica vigilar en qué partes de la ciudad están trabajando los policías; para ello, se usan “cajas de patrulla”, zonas delimitadas por calles y plazas en las que se distribuye a los agentes según el criterio escogido (mayor o menor seguridad, por ejemplo) o según necesidades puntuales (como un desfile que reúna grandes aglomeraciones), y se mide cuánto tiempo pasa cada agente dentro de su caja asignada. Por último, la efectividad se consigue alcanzando los objetivos, en este caso la reducción de la criminalidad, lo que según Geolítica hace que los policías tengan una actitud proactiva en lugar de reactiva, redundando en una mayor confianza hacia ellos.

El algoritmo de Geolítica trata de estudiar los patrones delictivos a través de un modelo que estudia tres conductas de los criminales. En primer lugar, el riesgo de victimización repetida consiste en estudiar cómo aumenta la probabilidad de que un delincuente vuelva al mismo lugar a cometer un delito. El equipo de Geolítica ha descubierto que la probabilidad de reincidencia es más alta en un lugar en el que el delincuente ya ha perpetrado un delito en

⁶⁵ No deja de resultar llamativo que la compañía defienda de forma directa la transparencia de su algoritmo: ello seguramente es debido a las críticas que estos sistemas cosechan en torno a su condición de caja negra, que oculta su verdadero funcionamiento.

una víctima anterior, puesto que ya conoce el entorno y los posibles riesgos de que fracase. En segundo lugar, se estudia el riesgo de victimización cuasirrepetida, referida a un incremento en la probabilidad de comisión de un hecho punible en un lugar cercano o con similitudes a la anterior escena del crimen o sobre una persona con características parecidas a las de la víctima. Por último, analizan la búsqueda local: según este algoritmo, los criminales suelen cometer delitos cerca del centro de actividades donde desarrollan su vida diaria, por lo que los delitos tienden a concentrarse geográficamente.

La fórmula empleada por este sistema recoge datos históricos acerca de los crímenes que ya se hayan consumado y otros registros de las unidades de policía. En concreto, se toman cinco variables de todos los incidentes registrados para poner en marcha la predicción del riesgo por el algoritmo: identificador del incidente (como códigos usados por la comisaría para clasificar cada expediente), tipo de crimen (descripción que en la unidad se utiliza para referirse al hecho punible en cuestión), localización del incidente (con coordenadas de latitud y longitud o con la dirección), hora de inicio y fin del incidente⁶⁶ (o momento aproximado, con la reserva de excluir del análisis aquellos delitos con un abanico de más de 72 horas entre ambos, para lograr mayor precisión en las evaluaciones) y el registro modificado de fecha y hora del incidente (opcional, para los casos en que se haya reclasificado el crimen en los archivos de la unidad).

La acogida de la policía predictiva (y en concreto, de Geolítica) por juristas y expertos ha sido diversa. Por un lado, algunos se muestran a favor de su uso, ya que es particularmente eficaz a nivel empírico, en el ámbito de la detección de puntos calientes. Por ejemplo, Levine *et al.* (2017) llevaron a cabo un estudio en Nueva York de 24 semanas de duración en el que comprobaron los resultados muy positivos de estos *software*⁶⁷, con una reducción del índice de criminalidad del 6%⁶⁸. Sin embargo, la dinámica de funcionamiento de Geolítica ha levantado cierta suspicacia entre la doctrina. Así, Benjamin (2019) ha reprochado que Geolítica sea una fábrica de crímenes, ya que al indicar qué zonas tienen mayor probabilidad

⁶⁶ La hora del fin del incidente se usa como criterio subsidiario, si no hay certeza de cuándo ha terminado de producirse un crimen (en el caso de unos dueños que descubren que han entrado a robar en su casa mientras no estaban, pero no saben exactamente cuándo se produjo la salida de los delincuentes, se usaría la hora en que descubrieron esa situación).

⁶⁷ El estudio se ejecutó empleando más herramientas de policía predictiva, como ShotSpotter.

⁶⁸ Con resultados muy similares (reducción del 7,4%), pero en los departamentos de policía de Los Ángeles (California), y Kent (Reino Unido), véase el estudio de Mohler *et al.* (2015).

de acoger la escena de un delito, los agentes enviados allí para su supervisión estarán predispuestos a encontrarlos, por lo que serán más susceptibles de actuar de un modo concreto y que así surjan conductas que sean castigadas⁶⁹. Por otro lado, Sankin *et al.* (2021) señalan los sesgos de este algoritmo hacia distritos poblados por personas de raza negra y latinoamericana.

4.3.3 El último eslabón de la cadena: de cómo auxiliar a las víctimas. VioGén.

Regresando a nuestras fronteras, en el año 2007 entró en funcionamiento en nuestro país (con la excepción de Cataluña y País Vasco) el Sistema de Seguimiento Integral en los Casos de Violencia de Género, popularmente conocido como VioGén, cuya gestión es dependiente de la Secretaría de Estado de Seguridad del Ministerio del Interior. Este sistema fue establecido en virtud de lo dispuesto en la Ley Orgánica 1/2004, de 28 de diciembre, de Medidas de Protección Integral contra la Violencia de Género, para cumplir con lo dispuesto en los artículos 31.3 y 32.1 de la Ley⁷⁰, con el fin de defender a las víctimas, reuniendo en una única base de datos toda la información recabada por la Policía, efectuando predicciones de riesgo e imponiendo medidas de seguridad dirigidas a la protección de la persona⁷¹. Este procedimiento está regido por el empleo de algoritmos actuariales capaces de estimar la probabilidad de reincidencia de un individuo en base a la información proporcionada por la propia víctima, y trasciende los límites del proceso penal. COMPAS entra en juego sobre todo en la fase de enjuiciamiento, pero VioGén se anticipa, actuando ya en la de instrucción.

El sistema se pone en marcha a través de la denuncia de la mujer o de cualquier otra persona de hechos susceptibles de tipificación como violencia de género. En ese momento, se realiza un cuestionario a la víctima, valorando diferentes aspectos socioeconómicos de carácter general y otros propiamente vinculados con la comisión del hecho punible (factores históricos y dinámicos), como el grado de vulnerabilidad de la víctima o la situación del

⁶⁹ Un ejemplo de lo que en psicología se conoce como profecía autocumplida.

⁷⁰ Art. 31.3: *La actuación de las Fuerzas y Cuerpos de Seguridad habrá de tener en cuenta el Protocolo de Actuación de las Fuerzas y Cuerpos de Seguridad y de Coordinación con los Órganos Judiciales para la protección de la violencia doméstica y de género.* Art. 32: *Los poderes públicos elaborarán planes de colaboración que garanticen la ordenación de sus actuaciones en la prevención, asistencia y persecución de los actos de violencia de género, que deberán implicar a las administraciones sanitarias, la Administración de Justicia, las Fuerzas y Cuerpos de Seguridad y los servicios sociales y organismos de igualdad.*

⁷¹ Que, como es lógico, pueden compatibilizarse con medidas cautelares personales como órdenes de alejamiento, tomadas ya en el seno de un proceso penal.

presunto agresor. A partir de entonces, y tras la debida introducción y cruce de los datos en el sistema algorítmico, VioGén genera dos estimaciones: una sobre el riesgo de reincidencia y otra sobre el riesgo de homicidio. Una vez con los datos en la mano, los agentes de la Policía evalúan la situación, pudiendo modificar el nivel de riesgo al alza si lo estiman oportuno.

Cabe efectuar dos consideraciones en este punto. Por un lado, y al igual que ocurría con COMPAS, la estimación estructurada de riesgo sirve únicamente como propuesta sobre la que el juez (o la policía) pueden apoyarse para tomar sus decisiones; rige plenamente la tesis que defiende gran parte de la doctrina consistente en que sea el ser humano el que tenga la última palabra acerca de la aceptación de estas pruebas (Caro Coria, 2022 o Romeo Casabona, 2022, por citar tan solo dos ejemplos). Por otra parte, los policías tienen la capacidad de modificar el nivel de riesgo atribuido a ese delincuente, lo cual concuerda perfectamente con el primero de los puntos. Sin embargo, no deja de resultar curioso que solamente puedan hacerlo al alza (es decir, para endurecer las medidas de seguridad a imponer al encausado). Podría entenderse esta decisión como una forma de proteger más a las víctimas si fuera necesario; no obstante, ello implica admitir que el algoritmo puede arrojar resultados erróneos. Si eso es así, puede suponerse razonablemente que pueda equivocarse también en sentido contrario⁷² (obteniendo niveles de riesgo desproporcionadamente altos). A juicio de quien escribe, resulta incoherente que se asuma un fallo a la baja (pudiendo corregirlo al alza), pero que sin embargo no se pueda modificar a la baja un nivel de riesgo demasiado alto por un fallo del algoritmo⁷³.

Tras dicha evaluación definitiva, se encuadra el caso en uno de cinco perfiles de riesgo distintos en función de la gravedad de la situación. En primer lugar, el riesgo no apreciado se reserva para los casos en los que los resultados del análisis no son suficientes como para establecer medidas concretas sobre la víctima o el agresor, en cuyo caso la policía se limita a indicar a la mujer los cauces de los que se puede valer para continuar defendiendo sus intereses y los recursos de apoyo que puede tener a su disposición. El siguiente nivel es de riesgo bajo, en el cual se proporciona un contacto a la mujer y se hacen llamadas de forma ocasional para comprobar su situación. Además, se puede proceder a la incautación de las

⁷² O quizás no: esta parece otra expresión más de los problemas de falta de transparencia y sesgos cognitivos que lastran el uso de estos sistemas de prevención.

⁷³ A decir verdad, esa modificación *a posteriori* por la Policía se da únicamente en el 5% de los diagnósticos (Zurita Bayona, 2014), lo que indica que las predicciones suelen tener una alta fiabilidad.

armas del atacante si las tuviera en su posesión. El nivel de riesgo medio lleva aparejados controles ocasionales, con el posible ingreso en centro de acogida de la mujer y la geolocalización del agresor. El riesgo alto cuenta con medidas algo más drásticas, como el cambio de domicilio de la víctima, controles frecuentes en su ámbito laboral y familiar y otros de carácter aleatorio al inculpatado. Por último, el riesgo extremo se corresponde con los casos de mayor gravedad, con vigilancia policial permanente a la víctima y control sobre el encausado.

Una vez iniciado el proceso penal y recabadas las medidas judiciales que se hayan tomado, se idea un Plan de Seguridad Personalizado contando con la participación de la víctima. Este PSP incluye medidas y recomendaciones como llevar consigo el teléfono móvil manteniendo siempre la localización activada o apuntarse a cursos de defensa personal, así como crear un plan de protección (o incluso de escape, en caso de necesidad) dentro de la vivienda o en el puesto de trabajo. En principio, la protección continúa mientras el nivel de riesgo siga existiendo. Según se vaya reduciendo por circunstancias naturales (como que la víctima haya cambiado de domicilio o que el agresor se encuentre en prisión), las medidas adoptadas se irán relajando hasta que dejen de ser necesarias, momento en el que el caso pasará a ser “inactivo”.

Por último, es menester hacer una breve referencia al tratamiento de datos de víctima y agresor. Toda la información proporcionada por la mujer y la Policía se conserva en la base de datos de VioGén. Sin embargo, tal y como especifica el Ministerio del Interior, el acusado que haya sido absuelto en el marco del proceso penal iniciado a través de la denuncia puede pedir la supresión de los datos del sistema siempre y cuando exista sobreseimiento definitivo o sentencia absolutoria firme y se hayan cancelado los antecedentes penales, reservándose el Ministerio el derecho de denegarlo si continúan siendo necesarios para proteger a la víctima o prevenir la comisión de nuevos delitos.

Mensualmente, se publican estadísticas sobre los casos que investiga VioGén. En el informe disponible sobre el mes de marzo (31 de marzo de 2023), hay un total 738 611 casos abiertos en nuestro país, de los cuales 76 404 se encuentran en estado activo. La mayoría de ellos cuentan con niveles de riesgo no apreciado o bajo (63 983), y tan solo 21 corresponden a casos de nivel extremo. Similares tendencias se observan en nuestra Comunidad Autónoma: 33 536 casos (3 638 activos) con 3 009 de riesgo no apreciado o riesgo bajo y únicamente 1

caso extremo, en la provincia de Salamanca, correspondiente a una víctima en la franja de edad de 18 a 30 años.

¿Qué podemos decir sobre el éxito de este sistema? López Ossorio *et al.* (2016) analizaron la eficacia predictiva de las valoraciones de VioGén comparándolas con los episodios de reincidencia denunciados, llegando a la conclusión de que las estimaciones son adecuadas, si bien señalan que las de corto plazo suelen tener más fiabilidad que las de largo plazo (lo que por otro lado es bastante lógico, dado que cuanto más tiempo pase desde la medición, más fácil es que tanto la víctima como el delincuente hayan rehecho su vida). López Ossorio *et al.* (2019) efectúan un examen similar y destacan la capacidad de discriminación y calidad de los parámetros empleados por VioGén. Por su parte, González Álvarez y Garrido Antón (2015) llevaron a cabo un estudio estadístico de la satisfacción de las beneficiarias de VioGén, con resultados muy positivos en relación a la eficacia de las medidas: la media es de un 8 sobre 10 (a pesar del reducido tamaño de la muestra, que admiten los autores). González Álvarez *et al.* (2018) proponen extender este sistema, a la luz de su efectividad, a otros ámbitos como el abuso y abandono de personas mayores o el tráfico de seres humanos.

Por último, efectuaremos un conciso análisis sobre el papel que juega VioGén en la parcela jurisprudencial española, fijándonos en la Sentencia de la Audiencia Provincial de León 683/2022 (Sección 3ª), de 1 de julio de 2022, que resuelve sobre un recurso de apelación planteado frente a un auto que denegaba la solicitud de una mujer consistente en la implantación de un dispositivo de geolocalización a un acusado por delitos de violencia de género consumados en La Bañeza. La apelante aduce que VioGén había arrojado un nivel de riesgo alto, habida cuenta de las lesiones recibidas y recogidas por atestado policial (“fractura de los huesos nasales, además de golpe en región frontal”), así como la omisión de socorro por parte de su agresor. Así, basándose en lo dispuesto en la Instrucción 4/2019 de la Secretaría de Estado de Seguridad –que insta a las Fuerzas y Cuerpos de Seguridad del Estado a pedir la medida discutida en casos de riesgo medio, alto y extremo-, solicita la rectificación de este extremo. Finalmente, la Sala estima su recurso argumentando que las lesiones son de suficiente gravedad como para, atendiendo a los informes favorables a la adopción de esta medida, colocar un dispositivo de rastreo al presunto agresor.

4.3.4 La evaluación del riesgo al otro lado de los barrotes: RisCanvi

También en nuestro país se encuentra RisCanvi, un algoritmo dependiente de la Dirección General de Asuntos Penitenciarios del Departamento de Justicia de la *Generalitat* de Cataluña. Fue implantado en el año 2007 en estrecha colaboración con el Grupo de Estudios Avanzados en Violencia de la Universidad de Barcelona (capitanado por el profesor Antonio Andrés Pueyo) para mejorar la gestión del sistema penitenciario catalán, emitiendo distintas predicciones sobre violencia y riesgo de reincidencia en los reos de las prisiones⁷⁴, empezando a aplicarse sistemáticamente en 2008.

Las escalas RisCanvi, de clara vocación preventivo-especial, tienen como objetivo la detección del riesgo sobre 5 tipos de comportamiento diferenciados. En primer lugar, la violencia autodirigida se refiere a agresiones que la persona dirige contra ella misma, incluyendo autolesiones leves y graves e intentos de suicidio. En siguiente lugar, la violencia intrainstitucional comprende las mismas conductas que la violencia autodirigida, si bien el destinatario de las mismas son otros reclusos o funcionarios de prisión que trabajen en el propio centro penitenciario. Para que una conducta se encuentre incluida en esta categoría, se requiere que sea objetivamente capaz de causar daño a uno o varios individuos⁷⁵, aunque abarca también las amenazas o intentos de daño no consumados y la destrucción de objetos. En tercer lugar, la reincidencia general engloba la comisión de un hecho punible que desemboque en una condena o una medida de seguridad alternativa. Se exige que este delito se haya cometido tras ser puesto en libertad por cualquier motivo (cumplimiento de condena, permiso penitenciario o tercer grado, entre otros), por lo que nunca serían susceptibles de esta calificación la reincidencia por conductas realizadas durante la estancia dentro del establecimiento penal. En siguiente lugar, la reincidencia violenta incluye la realización de las mismas conductas que la general y bajo las mismas circunstancias (siempre fuera de prisión), pero con un requisito adicional: el hecho punible debe estar revestido de una particular violencia. Por último, RisCanvi trata de predecir el posible quebrantamiento de condena (que es un delito en sí mismo, tipificado en el artículo 468 del Código Penal), con comportamientos como no retorno de un permiso, fuga o violación de las normas del centro penitenciario.

⁷⁴ Un sistema similar es HART (*Harm Assessment Risk Tool*), usado por la Policía de Durham (Reino Unido) para derivar en su caso al convicto al programa de rehabilitación Checkpoint.

⁷⁵ Traducción propia. Texto original: *és la conducta que objectivament és capaç de causar dany a un o varis individus.*

Para ello, se hace uso de un sistema formado por una lista de factores indicativos de un posible comportamiento violento futuro que, combinados a través de un algoritmo, arrojan un resultado final que el juez evalúa, por ejemplo, a la hora de conceder terceros grados. Estos factores se agrupan en cuatro categorías diferentes: criminales y penitenciarios (como el historial delictivo o problemas de conducta), personales y sociofamiliares (recursos económicos o apoyo familiar) clínicos y de personalidad (consumo de sustancias tóxicas o conductas suicidas) y otros como el sexo, edad o nacionalidad.

RisCanvi cuenta con dos escalas diferentes de evaluación. Por un lado, RisCanvi Screening (RisCanvi-S) analiza 10 factores de riesgo en todos los presos de manera sistemática para clasificarlos en dos categorías, bajo riesgo y alto riesgo, para cada uno de los criterios que mide el sistema, pudiendo aconsejar una valoración posterior por parte de RisCanvi Completa si se considera que el sujeto pertenece a un grupo especial. RisCanvi-C efectúa un análisis de mayor calado, empleando 43 factores, con tres categorías de riesgo: bajo, medio y alto. En el caso de RisCanvi-C, los técnicos podrán corregir (al alza o a la baja) la valoración de manera justificada.

Tras la aplicación de la escala correspondiente, puede ocurrir que el sistema considere que el sujeto reúne ciertas peculiaridades que recomienden un análisis todavía más exhaustivo por parte de otras herramientas, como *Historical Clinical Risk – 20* (HCR-20) o *Psychopathy Check List – Revised* (PCL – R). Así, RisCanvi ha ideado seis posibles “grupos especiales”, revestidos de una especial tendencia hacia el comportamiento violento, donde incluir a estos reos para indagar de manera más profunda en su situación: VIOFAM (para agresores violentos domésticos con víctima distinta a la pareja), VIOPAR (agresores de pareja o expareja, excluyendo los delitos sexuales), VIOSEX (agresores condenados por delitos sexuales violentos), VIOCRO (delincuentes crónicos), VIOTRAS (delincuentes violentos que en el momento de la agresión estaban afectados por trastornos mentales severos, excluyendo la psicopatía) y PSICOPAT (con comportamiento violento precoz y susceptibles de sufrir un trastorno psicópata).

En cuanto a su funcionamiento, cabe destacar varias diferencias fundamentales respecto a otros algoritmos analizados previamente, como COMPAS o VioGén. En primer lugar, RisCanvi es un sistema de uso exclusivamente penitenciario: solo se pone en marcha con la

entrada del reo en prisión, y no antes -para asesorar a las partes o proteger a la víctima durante el desarrollo del juicio-. Además, su uso se extiende a todos los penados que ingresen en las cárceles de Cataluña, y no solo a los que hayan cometido delitos de carácter sexual. Por otro lado, el sistema se basa en una reevaluación permanente del preso, pues los análisis caducan a los 6 meses de su realización, efectuando uno nuevo a los efectos de observar la evolución del riesgo⁷⁶ y actualizar la situación del reo a las circunstancias de cada momento. Por último, el método de recogida de datos no se basa directamente en cuestionarios⁷⁷, como sí ocurría en los otros sistemas, sino que se protocoliza directamente a través del algoritmo que reúne y cruza información de todo tipo de fuentes (desde la ficha policial hasta la propia sentencia, pasando por la preceptiva entrevista al ingresar en el establecimiento penitenciario) que aporte evidencias que puedan servir para el diagnóstico. A partir de ese análisis, se identifica qué elementos han podido causar la comisión del delito y qué se puede hacer para evitar la reincidencia en el futuro, tomando las medidas oportunas de manera individualizada para cada uno.

RisCanvi cuenta con dos procedimientos diferentes a seguir según el sujeto que se esté analizando. El primer proceso se reserva para los internos en prisión preventiva, y comprende la evaluación a través de distintas fases. En primer lugar, se utiliza RisCanvi Screening para hacer un primer cribado en el que únicamente se analiza el riesgo de reincidencia (o de comisión) en violencia intrainstitucional y violencia autodirigida. El plazo máximo para cerrar esta etapa es de 4 semanas desde que el condenado entre en prisión. En función de los resultados obtenidos con RisCanvi-S, se abren dos posibilidades diferentes; si se categoriza al sujeto como de riesgo bajo, el proceso finaliza ahí, hasta que en 6 meses se vuelva a efectuar un nuevo diagnóstico, ya como parte del segundo colectivo (población penada) o no, si finalmente la prisión preventiva se ha levantado sin condena. También es posible realizarlo antes de 6 meses en caso de circunstancias críticas (cambio de actitud repentino, intento de suicidio, o huelga de hambre, entre otras). Por el contrario, una calificación de riesgo alto lleva a una segunda evaluación usando la escala más compleja –

⁷⁶ Como mera anécdota, Carlos Soler –en su momento, Subdirector General de Programas de Rehabilitación y Sanidad de la *Generalitat* - (2019) comenta que precisamente este fue el germen del nombre RisCanvi: un juego de palabras entre *risc* – riesgo en catalán – y *canvi* –cambio- para expresar la capacidad de cambio en el riesgo de reincidencia de los penados y la adaptación constante a las circunstancias.

⁷⁷ Andrés Pueyo (2019), en declaraciones a *Els Matins*, tras obtener el Grupo de Estudios Avanzados en Violencia el Premio Antoni Caparrós otorgado por la Fundació Bosch i Gimpera por su labor en el desarrollo de RisCanvi.

RisCanvi-C- que atribuirá de manera más concreta la atribución de riesgo bajo, medio o alto, y cuyos resultados deben ofrecerse en 4 semanas desde la decisión de emplear esta escala, con la posibilidad de ajustar el nivel de riesgo justificadamente⁷⁸. Desde ese momento, y con independencia del resultado obtenido, los exámenes se sucederán cada 6 meses o antes en caso de necesidad.

El segundo protocolo se aplica a dos tipos de sujetos (la población penada por sentencia firme y los penados por un delito violento), evaluándose las cinco conductas: violencias institucional y autodirigida, reincidencias general y violenta y riesgo de quebrantamiento de condena. La administración a los penados por delitos no violentos comienza con la aplicación de RisCanvi-S. Si el riesgo es bajo, se procederá a la reevaluación en 6 meses o circunstancia excepcional, mientras que si el riesgo es alto o se ha hecho una recomendación por pertenencia a un grupo especial, se pasará al segundo análisis por RisCanvi-C. En el primero de los casos, la evaluación de riesgo bajo, medio o alto comporta la revisión en 6 meses o caso de urgencia. Si se ha producido una recomendación, la valoración se llevará a cabo usando los algoritmos específicos según el grupo especial en que hayan sido encuadrados. Por su parte, los condenados por delitos violentos obvian el trámite de la revisión con RisCanvi Screening, evaluándose directamente a través de Riscanvi Completa, con el consiguiente resultado de riesgo bajo, medio o alto (y con el mismo régimen de revisión que en los casos anteriores) y la posible valoración posterior con escalas específicas si sus características son compatibles con los estándares de los grupos especiales (plazo máximo de 4 semanas para completar el análisis).

En cuanto a los resultados que ofrece RisCanvi, Férez Mangas y Andrés Pueyo (2018) se muestran algo optimistas, indicando que RisCanvi (especialmente RisCanvi Completa) ofrece predicciones más fiables que la Tabla de Variables de Riesgo, una de las primeras herramientas de análisis actuarial en España, y en particular en el ámbito del quebrantamiento de condena. Varios estudios (Férez Mangas, 2017a y Férez Mangas, 2017b) afirman que su eficacia predictiva es de aproximadamente el 60% en RisCanvi-S y del 69,9% para RisCanvi-

⁷⁸ De nuevo, el ser humano tiene la última palabra (al menos en RisCanvi-C).

C, por lo que aconsejan realizar algún ajuste (especialmente en la versión Screening) para así optimizar sus pronósticos⁷⁹.

RisCanvi también ha irrumpido con fuerza en la práctica forense como un elemento de cierta eficacia probatoria, con numerosas sentencias que lo consideran en la fundamentación de su fallo. Para ilustrar este punto, nos detendremos brevemente en la Sentencia de la Audiencia Provincial de Barcelona 353/2021 (Sección 21^a), de 28 de enero de 2021, que resuelve un recurso de apelación sobre un auto del Juzgado de Vigilancia Penitenciaria que denegaba un permiso penitenciario a un recluso de la cárcel de Lledoners. El auto apelado, aun admitiendo la “buena evolución del interno”, señala que la naturaleza del delito cometido (de índole sexual), la no compleción de un programa de tratamiento y en particular, el hecho de que de acuerdo con RisCanvi, el penado tenía un riesgo de quebrantamiento de condena medio, no era procedente la concesión del permiso. Por otro lado, la parte apelante esgrime que el penado tiene arraigo en España y que el programa de rehabilitación se suspendió de forma sobrevenida por razones ajenas al interesado (derivado de la crisis sanitaria de 2020). Por último, señala el informe favorable de la Junta de Tratamiento de Lledoners basándose en criterios objetivos, en particular que RisCanvi arrojaba un riesgo bajo en el resto de variables a medir (siendo el quebrantamiento de condena la única excepción con riesgo medio). A la luz de todas estas consideraciones, la Sala decidió estimar el recurso de apelación atendiendo al derecho fundamental del penado y autorizando el permiso penitenciario.

En conclusión, la política criminal se está sirviendo de numerosos instrumentos (plenamente integrados en el sistema penal), a fin de avanzar en los objetivos preventivo-especiales que caracterizan una parte de nuestra disciplina, incorporando herramientas que pueden ser de gran ayuda para que “el problema de la prevención del delito se solucione en el ámbito tecnológico” (Romero Flores, 2015, p. 336).

⁷⁹ A pesar de todo, aún hay dudas sobre su efectividad, pues medios como *La Vanguardia* señalan que hasta un 82% de los presos que se consideraron peligrosos no volvieron a delinquir (con las debidas reservas por proceder esta información del ámbito periodístico y no estrictamente jurídico o académico).

5. DESAFÍOS ÉTICO-PENALES DE LA NEUROCIENCIA

Teniendo en cuenta lo que al conjunto del sistema penal pueden aportar los avances en neurociencia e IA, deben tenerse presente los desafíos que traen consigo, desde una posible codificación de nuevos tipos penales asociados a las vulneraciones de nuevos bienes jurídicos emergentes hasta la atribución de responsabilidad penal a estos productos neurocientíficos (en particular, de la IA) o la colisión de estas tecnologías con los derechos humanos planteada desde un punto de vista ético-moral⁸⁰.

En primer lugar, cabe efectuar un análisis acerca de la posible inclusión en nuestro Código Penal de delitos cuya responsabilidad pueda atribuirse a la IA. En este sentido, es menester diferenciar dos facetas distintas: por un lado, la responsabilidad de la IA como ente autónomo, y por otro, la que pueda inferirse de la programación del propio sistema.

5.1 ¿La IA, responsable de hechos punibles? Estado actual y obstáculos para su regulación

En relación con el primer extremo, se ha planteado la tipificación de delitos causados por inteligencia artificial. Esta no es sino una expresión más del proceso de antropomorfización que el ser humano ha otorgado a la IA, en la línea de lo que mencionaba Mitchell (2020), y que ya hemos mencionado en este trabajo. Los constantes –y cada vez más acelerados y bruscos- cambios de tendencia en la sociedad actual⁸¹ aconsejan una acorde actualización del Derecho Penal. Así lo aprecia Palma (2022), que, adhiriéndose a la tesis de Mitchell, se pronuncia en los siguientes términos: “el proyecto del Derecho (...) también puede ofrecer soluciones y límites significativos y no ser tan solo un artefacto vuelto obsoleto frente a nuevas posibilidades científicas” (p. 185).

De hecho, esta revisión ya se ha llevado a cabo en las últimas décadas con la introducción de tipos penales relacionados con las nuevas tecnologías. Sirvan a modo de ejemplo los artículos 197 bis, 197 ter, 249, 264 y 264 bis. Los artículos 197 bis y 197 ter

⁸⁰ Muy recientemente, Geoffrey Hinton, considerado por muchos uno de los pioneros de la IA, renunció a su cargo de Vicepresidente de Ingeniería en Google constatando que estos sistemas nos superarán en capacidades intelectuales y advirtiendo del aciago futuro que nos espera si no actuamos de manera urgente.

⁸¹ Algunos, como el director de la entidad financiera BlackRock, Larry Fink, incluso vaticinan el fin de la globalización (2022) en una carta dirigida a los accionistas de la compañía.

fueron introducidos en nuestro Código Penal por la Ley Orgánica 1/2015, de 30 de marzo, y castigan con pena de prisión de 6 meses a 2 años a quien “por cualquier medio o procedimiento, vulnerando las medidas de seguridad establecidas para impedirlo, y sin estar debidamente autorizado, acceda o facilite a otro el acceso al conjunto o una parte de un sistema de información o se mantenga en él en contra de la voluntad de quien tenga el legítimo derecho a excluirlo”, así como a quien “intercepte transmisiones no públicas de datos informáticos que se produzcan desde, hacia o dentro de un sistema de información (...)”, en cuyo caso la pena de prisión se reduce a un abanico de 3 meses a 2 años y es sustituible por multa de 3 a 12 meses. El artículo 197 ter castiga a quien provea a terceros de un programa informático u otros códigos para acceder al sistema, es decir, al *hacker* que tenga los conocimientos técnicos necesarios para crear un virus informático capaz de acceder al sistema. El artículo 249, por su parte, se modificó en 2022 (Ley Orgánica 14/2022, de 22 de diciembre), e incluye entre los autores de delitos de estafa a quien consiga una “transferencia no consentida de cualquier activo patrimonial en perjuicio de otro” usando medios informáticos, en referencia a los piratas que operen fraudulentamente con fondos de cuentas bancarias ajenas. Por último, el artículo 264 castiga, en definitiva, a quien dañe o altere datos de sistemas informáticos ajenos, siendo modificado también en 2015. Es posible advertir, por tanto, que surgen o se modifican tipos penales para proteger bienes jurídicos (la intimidad y la propia imagen, en el caso de los artículos 197 bis y 197 ter⁸², o la integridad de los activos propiedad de una persona) que antes no podían vulnerarse a través de medios inexistentes en la época.

El punto de conexión más evidente entre la neurociencia y el Derecho Penal en el ordenamiento jurídico español se encuentra en el Título V del Libro II del Código Penal vigente (Delitos relativos a la Manipulación Genética) que abarca los artículos 159 a 162. En lo que a nuestro ámbito de estudio interesa, centraremos nuestro análisis en los dos primeros preceptos. El artículo 159.1 castiga con prisión de 2 a 6 años e inhabilitación especial de desempeño de cargo público de 7 a 10 años a quienes “con finalidad distinta a la eliminación o disminución de taras o enfermedades graves, manipulen genes humanos de manera que se altere el genotipo”. Observamos aquí una diferencia fundamental con otros ordenamientos de nuestro entorno: mientras que la alteración del genoma para eliminar defectos del ser humano es una gran preocupación entre los juristas (Romeo Casabona, 2022) -llegando a

⁸² Refrendado a su vez por el encaje sistemático de los propios delitos, situados en el Título X del Código Penal (Delitos contra la intimidad, el derecho a la propia imagen y la inviolabilidad del domicilio).

estar taxativamente prohibida la selección genética en países como Suiza- la redacción de este precepto invita a entender que esta práctica no está penalizada en España, lo cual bien merece una revisión de acuerdo con las tendencias actuales tanto doctrinales como legislativas. El artículo 159.2 tipifica una modalidad atenuada de alteración del genotipo, en este caso por imprudencia grave, en cuyo caso la pena de prisión se sustituye por una de multa de 6 a 15 meses e inhabilitación de 1 a 3 años.

Más consonante con las tendencias actuales se encuentra el artículo 160, y en particular sus apartados 2 y 3, introducidos con la modificación del Código del año 2003 a través de la Ley Orgánica 15/2003, de 25 de noviembre, y en vigor desde el 1 de octubre de 2004. Estos preceptos tipifican la fecundación de óvulos humanos para fines distintos a la procreación y la clonación de seres humanos para seleccionar la raza, castigándolos con pena de prisión de 1 a 5 años e inhabilitación especial de 6 a 10 años. La redacción de este artículo no es sino la respuesta del ordenamiento a desafíos como el transhumanismo, sancionando la experimentación con óvulos o embriones en etapa germinal y la clonación de seres humanos.

¿Qué dificultades tiene la neurociencia que lastran la creación de tipos penales adecuados a las necesidades actuales? En primer lugar, y al contrario que los medios informáticos al uso, los avances neurocientíficos aún no han llegado a su apogeo: a día de hoy, aún se está experimentando con ellos para descubrir qué son capaces de hacer autónomamente y qué cotas pueden alcanzar. Dado este estadio temprano de la técnica, los debates éticos –aun siendo imprescindibles⁸³- posiblemente no cuenten con todos los datos necesarios para obtener conclusiones realmente determinantes que puedan plasmarse en textos legislativos⁸⁴.

Otra de las razones que motiva esta tardanza es la calificación que se les pueda dar a estos nuevos entes. Pacheco Rodríguez (2022) indica el desafío del Derecho de la “atribución de la condición de persona a seres racionales no humanos y por tanto, sujetos titulares de derechos” (p. 766). Los seres humanos son responsables por sus acciones; tienen, por tanto,

⁸³ Comentan Miró Llinares (2018) y Muñoz Ortega (2022) que precisamente este es el momento propicio para indagar en las implicaciones de la neurociencia para el derecho y la sociedad y así estar preparados antes de que se desarrolle del todo y sea demasiado tarde para enfrentarse a ello.

⁸⁴ Con todo ello, algunos países ya han incluido en sus instrumentos jurídicos disposiciones encaminadas a la protección de nuevos derechos, como veremos posteriormente.

obligaciones, pero esas obligaciones siempre vienen acompañadas por derechos de los que son titulares. Es en este punto donde se trunca el planteamiento de los legisladores: si se pretende atribuir responsabilidad directa a estas entidades, deben otorgárseles correspondientemente unos derechos, lo que plantearía dificultades en relación a los límites de la convivencia con los seres humanos.

Busato (2022) suma a estos obstáculos uno adicional, relativo al planteamiento dogmático sobre la capacidad de acción y la voluntad. El propio término “Inteligencia Artificial” lleva a equívoco, pues estas máquinas no son inteligentes como tal: la inteligencia implica, además de una capacidad de lógica o comprensión, una inteligencia emocional, componente volitivo que brilla por su ausencia en la anatomía de la IA. Ello se ilustra con un sencillo ejemplo. La responsabilidad por un homicidio causado en accidente por un coche dirigido autónomamente nunca podría ser atribuida a la IA que lo controla, por lo siguiente: para que haya acción, debe haber intencionalidad objetiva de hacerlo, pero una IA es incapaz de expresar objetivamente intenciones; ella no quiso matar, falta la voluntad para ello.

Romeo Casabona (2022) asume el reto que tiene por delante el Derecho Penal de proteger al ciudadano de “nuevas formas de agresión” (p. 78) causadas por la IA, y asevera que el sistema lo ha resuelto de manera ciertamente eficaz. De este modo, en lugar de cohibir el desarrollo e investigación de estas tecnologías, se ha apostado por afrontar la situación desde una óptica preventiva conservando el Derecho Penal su característica subsidiariedad – es decir, su entrada en juego únicamente en los casos en que es estrictamente necesario-. Romeo cree que la controversia anterior debería resolverse otorgándoles la naturaleza de bien jurídico en sí mismo, protegiendo la consideración social de la entidad. Para este autor, la posible responsabilidad jurídica de los sistemas de IA debe resolverse teniendo siempre en mente que el ser humano es quien se encuentra ulteriormente detrás de su programación, imponiendo deberes de prevención a estos para evitar consecuencias adversas derivadas de su comportamiento.

5.2 ¿Clara como el cristal? Responsabilidad del programador y falta de transparencia

Esta reflexión nos lleva a desarrollar el segundo de los puntos: la responsabilidad aplicable no a las máquinas en sí mismas, sino a los programadores y personas encargadas de

su construcción. El mayor obstáculo en este sentido es el que para muchos autores es el talón de Aquiles de los sistemas de neurociencia e IA como sustitutivos del juez en la justicia predictiva, y que ya anticipamos previamente: la falta de transparencia de estos sistemas. La dificultad para saber cómo está programada una máquina para decidir, cómo piensa y por qué piensa así –algo que en ocasiones ni siquiera el programador conoce (Busato, 2022), ya que solo sientan las bases de la máquina, que después aprende autónomamente (de ahí el término *machine learning*)- se presenta como una auténtica traba en el camino hacia la regulación de esta innovación llamada a jugar un rol muy importante en el devenir de nuestra sociedad. Para De Sousa (2020), la responsabilidad debería recaer sobre el programador como culpable por un delito de peligro abstracto por haber creado la IA a pesar de conocer los riesgos de su posible “enajenación”, pero nunca porque hubiese cometido el delito por sí mismo.

Tanto en el ámbito judicial como en el estrictamente criminal⁸⁵, los sistemas de *machine learning* se ven lastrados por esta falta de transparencia⁸⁶. En el primero de los casos, una máquina capaz de analizar datos para arrojar una sentencia absolutoria o condenatoria podría arrastrar sesgos cognitivos del propio programador, reflejados en su código fuente. Consecuentemente, esos resultados podrían ser discriminatorios con determinados colectivos, según la forma en que su creador haya enseñado a estos sistemas a interpretar estos datos. Tiarks (2021) afirma que estos dos factores –falta de transparencia y existencia de sesgos- pueden causar un impacto ciertamente negativo en la legitimidad de las sentencias y, ulteriormente, podrían minar la confianza de la población en todo el sistema penal y judicial. En este sentido, esta autora inglesa señala la problemática de considerar que la legitimidad de una sentencia suele ser mayor cuanto más cercana esté la decisión tomada a la visión de justicia de la sociedad. Por lo tanto, la forma de alentar a la población a seguir lo dispuesto en los distintos instrumentos legales debe ser el meticuloso seguimiento de un proceso penal justo, con todas las garantías. Según Tiarks, dos de los indicadores que el público identifica en un juicio justo son la imparcialidad del juzgador -o, lo que es lo mismo, el razonamiento de las decisiones, de forma transparente- y la falta de prejuicios al tomarlas. Precisamente ello es lo que lastra el uso de algoritmos en el sistema de justicia penal, por lo

⁸⁵ E incluso en más ámbitos, como el financiero o el sanitario (Caro Coria, 2022).

⁸⁶ Esta crítica trasciende el ámbito legal: el conocido filósofo Popper (1971) ya negó valor probatorio real a un sistema inductivo que se base en simples correlaciones.

que concluye que no debería perseguirse una expansión en su uso habida cuenta de que exacerbarían los sesgos y reducirían la transparencia existente, cayendo la legitimidad de las sentencias y de confianza en el sistema al que todo ordenamiento jurídico debería aspirar.

5.2.1 *¿Los sistemas de evaluación de riesgo, a prueba? El caso State vs. Loomis*

Esta cuestión ya ha sido tratada jurisprudencialmente en diversas ocasiones. Como ejemplo, dedicaremos un breve espacio al análisis del Caso *State v. Loomis* (resolución 881N.W.2d749) sentencia emitida por el Tribunal Supremo de Wisconsin (Estados Unidos) el 13 de julio de 2016. Eric Loomis fue detenido y acusado de haber participado en un tiroteo en el municipio de La Crosse en el año 2013. Como parte del juicio, fue sometido a un análisis por COMPAS (cuyo funcionamiento ya desgranamos en el capítulo correspondiente), para obtener un nivel de riesgo de reincidencia que el juez tuvo a su disposición a la hora de fijar la pena en 6 años de prisión y 5 años de vigilancia posterior, llegando a obviar la posibilidad de suspender la pena por el alto riesgo de reincidencia arrojado por COMPAS. Tras desoír su solicitud de revisión de la pena, Loomis recurrió la sentencia alegando que el Tribunal había vulnerado su derecho a un proceso con todas las garantías, ya que asumía que su condena se había establecido atendiendo a los datos producidos por un algoritmo oscuro cuyo funcionamiento e información no era clara, y que incluso lo discriminó por ser varón.

El Tribunal rechaza el recurso del penado⁸⁷, fundamentando su decisión en el hecho de que había tenido acceso a la información que había obtenido (Fundamento Jurídico 46⁸⁸). Además, argumenta el Tribunal que COMPAS obtuvo esa información de servidores como historiales criminales o información aportada por el propio Loomis (FJ 55). Asimismo, deniega la revisión o eliminación del factor género en el análisis de COMPAS, ya que estadísticamente son los hombres los que mayor riesgo de reincidencia tienen; consiguientemente, modificar este parámetro conllevaría una reducción del nivel de

⁸⁷ No sin los votos particulares de dos de los siete magistrados que componen el Tribunal, Patience D. Roggensack y Shirley S. Abrahamson, que están de acuerdo con la decisión del Tribunal de rechazar los recursos, pero no con los motivos esgrimidos para ello.

⁸⁸ Dada la ingente información procedente de los Fundamentos Jurídicos de la sentencia original, me limito a incluir traducciones y adaptaciones propias sin detallar los textos originales en inglés. Para más información, véase la sentencia, disponible en línea en: <https://law.justia.com/cases/wisconsin/supreme-court/2016/2015ap000157-cr.html>

confianza de la máquina, por lo que arrojaría resultados menos fiables en procesos venideros (FJ 83).

Sin embargo, lo que realmente sobresale de esta sentencia es la fijación de un determinado ámbito en el que pueden emplearse estas herramientas para valorar el riesgo de reincidencia del reo. En su FJ 88, el Tribunal adjudica a COMPAS la tarea de asesorar al juez en la posibilidad de sustituir pena de prisión por otra que no lo sea en el caso de delincuentes con bajo riesgo de reincidencia, tener en cuenta posibles refuerzos en la vigilancia del sujeto dentro de la comunidad o las condiciones de dicha supervisión, pero nunca para decidir la pena (FJ 93)⁸⁹. Además, se deben incluir diversos pronunciamientos y garantías junto con el informe de COMPAS en su remisión al Tribunal, como que no debe usarse como fundamento único de la pena en la fase procesal de dictado de sentencia, la invocación expresa del secreto de empresa para evitar la promulgación de la forma concreta de efectuar los cálculos o el uso expreso de COMPAS para determinar grupos de riesgo alto y no individuos concretos, así como la constatación de la existencia de estudios que arrojan dudas sobre su efectividad (FFJJ 98-101).

A pesar de lo novedoso de esta sentencia en limitar el ámbito de aplicación de COMPAS, la resolución del Tribunal de Wisconsin no está exenta de detractores. Autores como Romeo Casabona (2018), Martínez Garay (2018) o Liu *et al.* (2019) coinciden en que resulta algo contradictorio que el Tribunal admita la utilización de informes de COMPAS solo si vienen precedidos por una advertencia de su falta de fiabilidad, lo que implica precisamente desconfiar del algoritmo. Recordemos la opinión de Martínez Garay (2022) al respecto:

“Quizá en ámbitos como la traducción automática (...) podamos preferir un algoritmo más exacto, aunque no sepamos cómo funciona. Pero cuando se trata de aplicar y ejecutar sanciones penales, es esencial que las decisiones contengan una motivación y que esta sea comprensible para el ser humano” (p. 499).

⁸⁹ Esto lo confirma Scott L. Horne, juez que dictó la sentencia apelada (FJ 106), que asevera que la pena habría sido exactamente la misma aunque no se hubiese llevado a cabo el análisis con COMPAS.

Los dos últimos trabajos arriba mencionados también critican la posible existencia de sesgos en la programación del algoritmo. De hecho, ambos mencionan las críticas a COMPAS por arrojar resultados discriminatorios contra la etnia afroamericana; Liu *et al.* (2019) recuerdan el derecho del acusado a “interrogar” al algoritmo a la luz de los errores de apreciación de COMPAS en una muestra de delincuentes no reincidentes a los que COMPAS sí atribuyó un riesgo alto.

Por último, critican la falta de transparencia del sistema –Martínez Garay (2018) llega a afirmar taxativamente que “los tribunales no deberían admitir valoraciones de riesgo hechas con algoritmos secretos” (p. 499)-, abriendo Liu *et al.* (2019) la veda a publicar los secretos industriales que conforman la llamada “caja negra legal” (p. 135), que hace que ni siquiera los Tribunales estén completamente seguros de cómo funciona COMPAS, no sin antes advertir de la más que probable protesta por parte de las empresas que diseñaron el servicio por la pérdida de ingresos que ello conllevaría. También mencionan la caja negra técnica⁹⁰, puesto que el *machine learning* aprende autónomamente de modos que ni los propios programadores pueden anticipar.

Por su parte, Martínez Garay (2022) dedica una parte de su análisis al segundo de los ámbitos tratados a lo largo del presente trabajo: la adjudicación de un determinado nivel de riesgo de reincidencia a un delincuente, con la correspondiente aplicación de las medidas de seguridad oportunas. Según esta autora, los defensores de la introducción de estos algoritmos en dichas valoraciones alegan una mayor objetividad de los datos –pues los seres humanos estarían en cierto modo contaminados por los sesgos- y una (paradójicamente) mayor transparencia –aduciendo que el cerebro humano puede racionalizar decisiones de forma inconsciente-. Martínez se opone a ambos argumentos; en primer lugar, y como ya hemos anticipado, los propios algoritmos pueden contar con sesgos introducidos por sus programadores (que pueden inducirles a valorar de manera más severa algunas características del condenado, por ejemplo). Por el contrario, el juez sí es capaz de verbalizar las razones por las que se ha pronunciado en una determinada dirección en una sentencia (de hecho, esto se ha erigido precisamente como una garantía procesal fundamental, el derecho a una resolución judicial motivada); si el órgano no efectúa esa valoración atendiendo a todas las pruebas, podría ser una valoración incorrecta, pero no poco transparente.

⁹⁰ Similar a lo observado en Martínez Garay (2022).

Resulta además paradójico que el algoritmo se considere transparente cuando su programación está protegida por el secreto industrial y, por tanto, no tiene por qué ser divulgada. A ello se suma la cada vez mayor sofisticación del código de estos aparatos, lo que da la razón a los detractores de la falta de transparencia de estos sistemas. Rudin (2019) advierte que ello es debido a un incentivo para proteger aún más si cabe esta inagotable fuente de ingresos: si un algoritmo fuese sencillo, otras compañías podrían adquirir su licencia y analizarlo con la conocida como ingeniería inversa (tratar de obtener la fórmula de un producto final analizando su composición y así producirlo). Una vez diesen con la clave de funcionamiento del algoritmo, podrían emular su producción y la empresa original que lo hubiese desarrollado inicialmente perdería su propia ventaja competitiva.

Otro sector de la doctrina se ha mostrado algo más permisivo con el uso de Inteligencia Artificial. Caro Coria (2022) acusa la falta de auditorías a estos sistemas para arrojar luz sobre su funcionamiento y asume la existencia de sesgos cognitivos en la caja negra que es el *machine learning*. No obstante, admite que, durante el confinamiento mundial que tuvo lugar a partir del año 2020, los distintos sistemas de IA permitieron, por un lado, rastrear y vigilar la evolución de la situación epidemiológica (geolocalizando, por ejemplo, a los infectados en determinados países) y abrir nuevos escenarios de cara a la suavización de las medidas tomadas: de lo contrario, “solo nos quedaba el encierro, las cuarentenas, los toques de queda, (...), para decirlo líricamente, ‘viendo la vida pasar’” (p. 646). A pesar de todo, sigue insistiendo en la necesidad de compatibilizarlo con los derechos fundamentales del individuo para limitar el margen de actuación del Estado que use estas tecnologías.

Por último, cabe hacer una breve referencia a un segmento en pleno desarrollo que se halla estrechamente relacionado con la IA y que ataja, según autores como Navarro Cardoso (2022), el problema de la falta de transparencia de estos aparatos: las cadenas de bloques, más conocidas con el anglicismo *blockchain*, y en particular su inserción en programas de cumplimiento penal con fines preventivos. El *blockchain* es una base de datos descentralizada y encriptada, de modo que nadie puede cambiar su contenido si no es con la autorización de la mayoría de usuarios. En opinión de este autor, la característica gestión conjunta del *blockchain* precisamente aporta mayor transparencia, ya que “todos los usuarios pueden acceder a la práctica totalidad de los datos” (p. 677). Debido a ello, la implementación de estos sistemas en el ámbito de la prevención penal se postula como una posible solución para contribuir a alcanzar la tan deseada transparencia exigida por los penalistas para la plena

aceptación de la IA y otros instrumentos neurocientíficos en el proceso decisorio político-criminal.

5.3 ¿Y nuestro papel en esta Revolución? Relación de la neurociencia con los Derechos Humanos⁹¹

Para culminar este capítulo y el grueso de este trabajo, analizaremos el choque que los diversos instrumentos neurocientíficos pueden tener con los derechos humanos, aspecto que ya hemos ido anticipando en cierto modo a lo largo de la redacción del presente escrito⁹², señalando además la consolidación de nuevos neuroderechos vigentes en distintos ordenamientos jurídicos.

En primer lugar, la neurociencia al servicio de la política criminal puede constituir una violación al derecho a la privacidad e intimidad. Como hemos mencionado a lo largo de este capítulo, técnicas como el fMRI, PET y el *brain fingerprinting*⁹³ son capaces de aseverar cuándo un cerebro humano reconoce un objeto a través de la detección del estímulo en sus ondas cerebrales. De este modo, los recuerdos del encausado pueden quedar expuestos al público de manera irremediable, con la vulneración de esos derechos⁹⁴. Otro instrumento que puede vulnerar potencialmente este derecho a la intimidad es el uso de *nudges*, bastante más intrusivos: un ejemplo es el uso de sensores oculares en monitores que controlan la dirección de la mirada al navegar por una página web⁹⁵, de modo que es posible registrar a qué parte de la pantalla dedica el usuario más atención y así poder inducirle a tomar determinadas decisiones basándose en estos datos⁹⁶ (Pärnamets *et al.*, 2015).

⁹¹ Sobre este tenor, véase el amplio estudio de Villamarín López (2014).

⁹² Con carácter preliminar, conviene advertir que las pruebas neurocientíficas podrían entrar en conflicto con la integridad mental pero no con la corporal, puesto que la actividad neuronal se relaciona con la personalidad o los sentimientos (Merkel, 2022).

⁹³ La doctrina también ha discutido la posible injerencia de técnicas como la hipnosis en este derecho. Dada su absoluta falta de rigor científico, no parece oportuno ni apropiado dedicarle un espacio en este trabajo.

⁹⁴ Además del *nemo tenetur*.

⁹⁵ Aunque no se trate de un instrumento propiamente neurocientífico, su uso podría extenderse fácilmente al ámbito penal, controlando por ejemplo la atención que un encausado por pederastia dedica a observar fotografías de niños mezcladas con otras de adultos.

⁹⁶ En realidad, las páginas web ya aplican distintos sistemas para optimizar y personalizar sus contenidos en tiempo real; por ejemplo, si el cursor de un ordenador permanece sobre un anuncio situado al margen de la página durante un tiempo prolongado, podría inferirse cierto interés del usuario por su contenido.

Estas vulneraciones de la intimidad hunden sus raíces en la monitorización de constantes anatómicas del propio cuerpo humano: presión arterial, impulsos cerebrales u otras de similar índole. Sin embargo, un sector de la doctrina defiende que la vulneración del derecho a la intimidad debe extenderse al registro del contenido de teléfonos móviles y otros aparatos que contengan información privada del investigado⁹⁷, muy habitual en diligencias previas en un proceso penal. Esta tesis de la que Porciúncula (2022) se hace eco fue postulada por primera vez por Clark y Chalmers (1998), que acuñaron el término *Extended Mind Theory* o Teoría de la Mente Extendida, para explicar que la mente del ser humano no se limita a lo que pueda conservarse en el cerebro, sino que debe extenderse a entidades extracorporales, basándose en la influencia que los factores externos pueden tener en la composición de nuestros recuerdos. Así, mencionan estos autores, un paciente de Alzheimer podría llevar consigo una libreta en la que apuntaría todo para no olvidarse de ello, y la libreta como tal formaría parte efectiva de su memoria. Por consiguiente, desarrollan el llamado principio de paridad: si algo se considera como cognitivo por haber ocurrido en la mente, debería considerarse cognitivo si ha sido realizado en una entidad externa. Partiendo de esta premisa y trasladándolo por analogía a nuestro campo, Porciúncula (2022) desarrolla el principio de paridad legal: toda evidencia obtenida a través de la tecnología de la mente extendida debe tener el mismo tratamiento que la evidencia neurocientífica detallada en el párrafo anterior. De este modo, el acceso a los archivos de un teléfono podría ser considerado una intrusión ilegítima en la memoria no biológica del sujeto con idénticas consecuencias al acceso a la mente biológica, situación ante la que el interesado podría invocar su derecho a la intimidad y una violación del *nemo tenetur*.

A continuación, analizaremos las posibles colisiones de la neurociencia con el derecho a la libre determinación de la personalidad. Algunas técnicas neurocientíficas empleadas en el sistema penal inciden directamente sobre la superficie del cerebro, pudiendo llegar a modificar las conexiones neuronales que se forman a lo largo de nuestra vida. En el ámbito preventivo, ha llegado a discutirse la procedencia o no de la supresión de memorias

Consecuentemente, la página ofrecerá más anuncios relacionados con el primero la próxima vez que visite su dominio.

⁹⁷ Aun discrepando radicalmente con la tesis de estos autores por la relevancia que tiene el registro de aparatos móviles en el esclarecimiento de la comisión de un hecho punible y los problemas que acarrearía equipararlo a una monitorización directa del cerebro, reconozco que su planteamiento es suficientemente novedoso como para tenerlo en cuenta de cara a una posible adición o reformulación de nuevos derechos en nuestros ordenamientos jurídicos, razón por la cual se ha considerado oportuno incluirlo en este apartado.

mediante técnicas como la optogenética⁹⁸ (Muñoz Ortega, 2022) en delincuentes con un largo historial delictivo provocado por experiencias traumáticas en la infancia⁹⁹, por ejemplo. De igual modo, la optogenética (que inicialmente fue ideada como solución para problemas oculares) podría servir en el ámbito penal para recuperar recuerdos perdidos o reforzar los recuerdos positivos (Adamczyk y Zawadzki, 2020). Naturalmente, esta neurotecnología con rasgos propios de una película de ciencia ficción¹⁰⁰ se encuentra aún en sus primeras etapas de desarrollo, pero los penalistas ya se han adelantado para advertir de los riesgos que comporta la posibilidad de actuar sobre los recuerdos de un delincuente, ya que podría atentar contra el derecho a la libre determinación de la personalidad, pues estaría perdiendo parte de la que adquirió a través de sus vivencias e induciéndolo a pensar de otra manera.

Otra vía neurocientífica que puede entrar en conflicto con este derecho es la implantación de dispositivos de decodificación de la actividad cerebral (BCI o *brain computer interfaces*), que permiten conectar el cerebro con un aparato externo que reproduzca lo que la persona quiere decir a través de la actividad neuronal. Ello es de gran utilidad para personas con dificultades en la movilidad o en la comunicación¹⁰¹ (como, en su momento, Stephen Hawking), pero puede causar graves daños a sus derechos. Dada la conexión de estas interfaces a Internet para poder transmitir la interpretación del contenido desde el cerebro hasta un sintetizador de voz (este es tan solo un ejemplo de sus muchas aplicaciones), Ienca y Andorno (2022) se preguntan si estas BCI podrían ser *hackeadas* externamente, pudiendo manipular directamente la actividad neuronal de la persona. Consecuentemente, se podría llegar a “eliminar el control del usuario sobre la aplicación, o incluso sustituir el control voluntario de la víctima” (p. 708). En un futuro no tan lejano, el delincuente podría controlar remotamente una silla de ruedas teledirigida directamente por su usuario a través de BCI, en contra de su voluntad. Estos autores advierten sobre una vulneración, además del derecho a la privacidad mental, a la integridad física y mental. Merkel (2022) apunta la posibilidad de

⁹⁸ Podría valorarse también la relación de la optogenética con el derecho a la integridad física, pues se requiere practicar una craneotomía en el sujeto como paso previo a la implantación del dispositivo en el cerebro.

⁹⁹ Asociación sobradamente demostrada por muchos autores. Por todos, Martín López (2022).

¹⁰⁰ Aunque con resultados prometedores (si bien muy escasos) en el ámbito médico: un hombre ciego de 58 años aquejado de retinitis pigmentaria ha recuperado la vista parcialmente gracias a un tratamiento optogenético (Sahel *et al.*, 2021).

¹⁰¹ Aunque no solo para ellos: Caro Coria (2022) contempla la posibilidad de aumentar la capacidad cerebral mediante BCI como Neuralink.

que, empleando esta clase de instrumentos, pueda llegar a castigarse la intención de hacer algo ya grabada en el cerebro y captada por estos sistemas. Sin embargo, termina por descartarlo argumentando que la BCI (como cualquier instrumento de IA) no entiende el contexto en que se desarrollan las palabras, sino que tan solo se limita a reproducirlos: en definitiva, lo precriminal no tiene contenido social.

En siguiente lugar, Ienca y Andorno (2022) mencionan los mecanismos de estimulación transcraneal de corriente directa (tDCS)¹⁰², técnica consistente en la aplicación de estímulos eléctricos sobre el cerebro a través de electrodos alternando beneficiosamente la actividad del cerebro con fines terapéuticos. Estos tratamientos aún en desarrollo también pueden provocar cambios en la personalidad, con la consecuente amenaza para el derecho a la libre determinación de la personalidad. En este sentido, estudios como los de Houeto *et al.* (2002) o Sensi *et al.* (2004) refieren cambios en el comportamiento sexual o agresividad de los pacientes sometidos a tDCS.

Por último, nos limitaremos a mencionar dos sectores tan mediáticos como controvertidos, que se encuentran en una fase (todavía) más temprana de desarrollo que los anteriores, pero cuya colisión con los derechos humanos fundamentales puede ser aún más flagrante. Por un lado, el transhumanismo (Pacheco Rodríguez, 2022), es decir, el biomejoramiento de la estructura biológica del ser humano mediante la implantación de tecnologías en su propio organismo, a fin de reducir enfermedades y mejorar la calidad de vida. Por último, la alteración del genoma humano en su etapa germinal (Romeo Casabona, 2022), con la posibilidad de elegir determinados rasgos por sus progenitores, con las dudosas implicaciones morales que ello conlleva.

5.3.1 Nuevos avances, nuevos derechos: la perspectiva de Ienca y Andorno (2022)

¿Hay algo que podamos hacer ante esta marea de descubrimientos neurocientíficos con claros desafíos a nivel ético-penal? Al fin y al cabo, “se trataría de proteger aquello que nos hace únicos” (Demetrio Crespo, 2017, p. 112). Este debate se ha abierto a nivel doctrinal, con autores como Ienca y Andorno (2022) formulando una serie de nuevos derechos concretos que en ocasiones se han visto reflejados en ordenamientos jurídicos vigentes en la actualidad. Con carácter previo a su enumeración, estos autores

¹⁰² También de estimulación magnética transcraneal (TMS) y estimulación cerebral profunda (DBS).

advierten de la necesidad de evitar lo que ellos llaman la inflación de derechos, “que es la objetable tendencia a etiquetar como derecho humano todo lo que es moralmente deseable” (p. 704), aportando pruebas concretas de por qué son necesarios.

En primer lugar, mencionan el derecho a la libertad cognitiva, con una doble dimensión positiva y negativa de hacer uso de las neurotecnologías y, en particular, a abstenerse de ello (Bublitz, 2013). Este derecho opera como sustrato del resto, y se asemeja a la libertad de pensamiento (Sententia, 2004). En segundo lugar, el derecho a la privacidad mental se encuentra estrechamente relacionado con el *big data*, pues protege al ser humano de la recolección y uso masivo de datos (por ejemplo, a través de *cookies*) procedentes directamente del cerebro por parte de las compañías que oferten productos neurocientíficos, como los BCI. En definitiva, se trata de encontrar un amparo legal para proteger a los ciudadanos de la lectura indiscriminada de los datos alojados en la mente, y proteger simultáneamente la información en la dimensión mental y la digital. En siguiente lugar, el derecho a la integridad mental puede ser vulnerado por las incisiones directas en el cerebro causadas por los BCI o incluso por las craneotomías empleadas en optogenética que, como consecuencia, pueden producir alteraciones en la memoria. Por último, se menciona el derecho a la continuidad psicológica, o derecho a conservar los rasgos identitarios de uno mismo tras una intervención neurológica. Actúa como una suerte de contrapeso con las técnicas de tDCS, combatiendo los cambios de personalidad o el no reconocerse a uno mismo tras la cirugía (Schüpbach *et al.*, 2006).

5.3.2 Un rayo de esperanza: mirada al derecho comparado

A pesar de lo novedoso de todos estos planteamientos, su tratamiento doctrinal ya se ha traducido en regulaciones de diferentes ordenamientos jurídicos que no por adolecer de un desarrollo completo merecen menos atención. Esta es una decisión aplaudida por, entre otros, Romeo Casabona (2022) y Caro Coria (2022), que defienden la necesidad de convertir recomendaciones (*soft law*) en disposiciones legales respaldadas por un régimen penal y administrativo sancionador en caso de su incumplimiento (*hard law*), como paso previo a una autorregulación del sector mediante medidas de *compliance*. Este último lo compara con la regulación de la libre competencia, que solo se solucionó con el intervencionismo estatal.

De este modo, nos centraremos en tres casos paradigmáticos reguladores de neuroderechos, nuevas tendencias neurocientíficas (en concreto, la ingeniería genética) y el estatuto de la IA (con posible atribución de derechos y obligaciones), no sin antes aclarar que aún no existe nada parecido en nuestro país; lo más cercano que tenemos es una Carta de Derechos Digitales, sin carácter normativo, emitida por el Gobierno el 13 de julio de 2021¹⁰³. Su carácter no vinculante aconseja que nos detengamos en derechos contenidos en una norma que sí esté en vigor. Resulta notorio destacar que en ocasiones estas regulaciones se adoptan a nivel constitucional, lo que indica la intención de los legisladores de que informen todo el ordenamiento jurídico.

La regulación pionera de los neuroderechos le corresponde a Chile, que en el año 2021 modificó el artículo 19 de su Constitución. El 25 de octubre de ese año, se promulgó en el Diario Oficial de la República de Chile la modificación de la Carta Magna por la ley 21383, que introdujo un párrafo nuevo en la redacción del artículo 19.1, que consagra los derechos “a la vida, integridad física y psíquica de la persona”. En particular, la Norma Fundamental recoge desde ese momento que la ley deberá regular la aplicación de los avances científicos en las personas, “debiendo resguardar especialmente la actividad cerebral, así como la información proveniente de ella”. Esto no es sino un ejemplo del papel de la doctrina en la codificación del Derecho, ya que entran en juego los derechos de continuidad psicológica, privacidad mental e integridad mental de Ienca y Andorno (2022) pues se trata tanto de proteger la información sita en el cerebro como de evitar su divulgación o modificación.

En siguiente lugar, dos países europeos han dedicado sendos apartados de su Constitución a tratar la cuestión de la biotecnología y otros instrumentos neurocientíficos, en particular los relativos a la genética: nos estamos refiriendo a Portugal y Suiza.

La nación helvética sitúa la regulación de estas materias dentro de la Sección 8^a (Vivienda, Empleo, Seguridad Social y Sanidad) del Capítulo II del Título III, sobre los Poderes de la Confederación Helvética, de los cantones y de las comunas. En primer lugar,

¹⁰³ Sus puntos 21.6 y 27 tratan, entre otras cosas, el respeto a la “dignidad, la libre autodeterminación individual, la intimidad y la integridad de las personas” en las investigaciones neurocientíficas y garantiza “el control de cada persona sobre su propia identidad, autodeterminación individual, confidencialidad (...) y seguridad de los datos obtenidos de sus procesos cerebrales”, así como fijar el compromiso de “regular el uso de interfaces persona-máquina susceptibles de afectar a la integridad física o psíquica” (es decir, de las BCI).

incluye la investigación sobre seres humanos en el artículo 118b¹⁰⁴, titulado precisamente de este modo, desde su introducción el 7 de marzo de 2010 por consulta popular en la Carta Magna, vigente de 1999. Este precepto recoge el deber del Estado de legislar sobre esta cuestión para proteger la dignidad y privacidad del ser humano, garantizando la prestación de consentimiento libre e informado. A continuación, el artículo 119 se ocupa de la medicina reproductiva y biogenética en relación a seres humanos, de nuevo atribuyendo al Estado el deber de proteger los derechos a la dignidad, privacidad y familia, y evitando el uso fraudulento de estas tecnologías. En este sentido, se prohíbe tajantemente la clonación y manipulación de células humanas y embriones y la hibridación con material de otros seres no humanos. El apartado c, incluido en el año 2015, señala que la reproducción asistida nunca podrá usarse para concebir un niño con características concretas (prohibiendo, por tanto, la alteración del genoma humano que mencionaba Romeo Casabona (2022)). Resulta curioso, no obstante, la posibilidad de hacer público el material genético de la persona (siempre con consentimiento de la misma), y el acceso a los datos relativos a sus ascendientes. Cierra la Sección 8ª el artículo 120, de carácter transitorio, según la propia Carta Magna, que regula la tecnología genética no humana. De nuevo, se consagra el derecho a la dignidad del ser humano y la seguridad del medioambiente ordenando al Estado el desarrollo de una legislación sobre el uso de material genético de animales, plantas y otros organismos.

Por su parte, el tratamiento de la neurociencia en Portugal se encuentra en el artículo 26 de la Constitución de 1976, encajado sistemáticamente en el Capítulo I (Derechos, Libertades y Garantías Personales) del Título II, junto al derecho a la vida o a la integridad personal. El artículo 26.1¹⁰⁵ consagra los clásicos derechos a la identidad personal, privacidad o libre autodeterminación de la personalidad. Sin embargo, es el apartado 3 el que, en términos similares a la Constitución Federal Suiza (si bien de forma algo más escueta), encarga al Estado la garantía de la dignidad personal y la identidad genética del ser humano, en todo lo relacionado con la creación y uso de estas tecnologías en la investigación científica.

¹⁰⁴ Adaptaciones propias. Texto original (en alemán) disponible en <https://www.fedlex.admin.ch/eli/cc/1999/404/de>. Texto disponible en inglés en el mismo enlace.

¹⁰⁵ Adaptaciones propias. Texto original (en portugués) disponible en: <https://www.parlamento.pt/Legislacao/Paginas/ConstituicaoRepublicaPortuguesa.aspx>. Texto disponible en inglés en: <https://www.parlamento.pt/sites/EN/Parliament/Documents/Constitution7th.pdf>

Por último, el Parlamento Europeo emitió el 16 de febrero de 2017 una Resolución –acompañada de una Carta sobre Robótica y un Código de Conducta Ética para los programadores- dirigida a la Comisión para que cree un cuerpo de normas civiles sobre robótica, llegando a dar un estatuto a la IA con posible atribución de responsabilidad¹⁰⁶. El Parlamento justifica la necesidad de regular esta materia atendiendo, entre otras cosas, a la influencia que la IA tendrá sobre sectores como la sanidad y el empleo y a la protección irrenunciable que los ordenamientos deben dispensar a la ciudadanía sobre los derechos a la dignidad, autonomía y autodeterminación del ser humano (incluso en las tecnologías que “mejoren a los seres humanos”, como el transhumanismo -letra O de la Introducción-, contemplando que “la Inteligencia Artificial llegue a superar la capacidad intelectual humana” –letra P-), o el uso creciente de estas tecnologías en ámbitos como la justicia predictiva –letra Q-, y concluyendo que los robots supondrán la siguiente revolución industrial –letra B-. De hecho, el Parlamento se muestra favorable a emplear ingeniería inversa para fomentar la innovación tecnológica en su punto 22, tratando de abrir la caja negra técnica que mencionaba Rudin (2019). En esta misma línea, se propone que el código fuente de los sistemas esté disponible para poder investigar los daños causados por estos robots.

De esta forma, este organismo aboga por crear una definición propia que englobe a todos estos sistemas y un registro de aquellos que se encuentren en circulación. Resulta interesante destacar que tres de los puntos (12, 14 y 20) de la Resolución se dedican al principio de transparencia, privacidad y protección de datos, aspectos que preocupan a los detractores de la justicia predictiva o al legislador en materia de neuroderechos. Tras enumerar una serie de ámbitos en los que estos instrumentos jugarán un papel vital en la sociedad (transporte autónomo, medicina, o empleo) pasa a detallar la necesidad de regular la responsabilidad de estas máquinas. Resulta notorio destacar el posicionamiento en torno a la responsabilidad objetiva, que “únicamente exige probar que se ha producido un daño o perjuicio y el establecimiento de un nexo causal entre el funcionamiento perjudicial del robot y los daños (...)” (punto 54). Por lo tanto, se está tratando de aligerar la carga de la prueba, facilitando al damnificado el acceso a la justicia por daños causados por los robots¹⁰⁷. Por

¹⁰⁶ En 2021, el Parlamento Europeo adoptó también una Propuesta de Reglamento para promulgar una Ley de Inteligencia Artificial, con modificaciones introducidas muy recientemente (10 de mayo de 2023) por las Comisiones de Mercado Interior y Libertades Civiles del Parlamento para abarcar también a los modelos de base generativa (*chatbots* como ChatGPT).

¹⁰⁷ La Resolución únicamente se refiere a los daños por responsabilidad civil (pues consideran más apropiado empezar por ahí –letra Y de la introducción-), si bien cabe esperar una futura interpelación a la Comisión en idéntico instrumento para legislar sobre responsabilidad penal.

último, se planta la semilla para “crear a largo plazo una personalidad jurídica específica para los robots, de forma que (...) puedan ser considerados personas electrónicas responsables (...)”, en definitiva, para regular la responsabilidad de estas entidades.

En conclusión, los países y alianzas internacionales han sabido recoger el guante de la doctrina de inmediato y se han decidido a proteger nuevos bienes jurídicos con recomendaciones o modificaciones recientes de sus instrumentos legislativos, consagrando nuevos derechos y debatiendo el castigo de conductas (otorgando responsabilidad penal a programadores o investigadores, por ejemplo) para afrontar el desafío que estas tecnologías han planteado y no verse superados por la situación. La audacia de estos Estados se puede resumir de forma tan breve como acertada acudiendo al refranero español: *Renovarse o Morir*.

6. ¿QUÉ PODEMOS APRENDER DE TODO ESTO?

CONCLUSIONES

Para finalizar este trabajo, presentamos aquí un decálogo que nos permita extraer conclusiones sobre los estrechos vínculos de la neurociencia y el Derecho a modo de recapitulación, tratando de ofrecer respuestas a las incógnitas que surgieron con carácter previo al análisis efectuado a lo largo de este escrito.

1. En el estado actual de la técnica, la neurociencia no va a alterar los postulados clásicos de la Teoría del Delito. A pesar de los estudios concluidos por los neurólogos sobre la voluntad y la discusión que ello ha suscitado entre la doctrina penalista (con cierta aceptación de un trasfondo biológico en nuestras decisiones), lo cierto es que los avances no son (todavía) suficientemente concluyentes en lo que al libre albedrío se refiere para cambiar el sistema punitivo basado en la culpabilidad por el hecho. La búsqueda de una conducta típica, antijurídica, culpable y punible, rasgos que se predicen de todo delito, sigue plenamente vigente.
2. Con mayor o menor fortuna, la admisión de los resultados de exploraciones neurocientíficas como medio de prueba en el proceso penal es ya una realidad, si bien aún no ha alcanzado su punto álgido. Su utilidad no se limita únicamente a sustentar la argumentación de una de las partes, sino que enriquece el debate judicial y contribuye a obtener una respuesta más precisa sobre la evaluación de las circunstancias del delito.
3. También la Inteligencia Artificial se ha abierto camino en el ámbito judicial como una alternativa al juez tradicional, tomando en consideración todas las circunstancias del caso y las pruebas presentadas para dictar una sentencia proporcional. No obstante, su uso se ve lastrado por las críticas que arrastra, como sus dificultades para distinguir entre correlación y causalidad o la falta de transparencia, lo que podría distorsionar la validez del juicio emitido.
4. El punto de conexión más evidente de la neurociencia y el Derecho Penal, con prolíficas sinergias entre ambas disciplinas, se halla en la Política Criminal. Los diferentes instrumentos neurocientíficos a nuestra disposición potencian el carácter

preventivo-especial de la pena, y pueden auxiliarnos para aplicar medidas de seguridad apropiadas al caso en pos de una reducción generalizada de la criminalidad.

5. Entre las herramientas de las que nos podemos servir para lograr este objetivo no solo se incluyen los tratamientos médicos prescritos, sino otras de carácter psicológico y tecnológico. Destaca entre todos estos métodos la elaboración de perfiles haciendo uso de biomarcadores biológicos y de conducta en el sujeto, que permitan explicar su comportamiento y, ulteriormente, posibiliten una rehabilitación más eficaz de la persona.
6. Pese a su extendido uso, los perfiles neurobiológicos han promovido cierta polémica entre los penalistas; entre las críticas más habituales se encuentran el peso desproporcionado concedido a los factores biológico y genético como indicadores de una mayor criminalidad en detrimento del social y afectivo y los sesgos a la hora de crear un cerebro modelo y elaborar los perfiles.
7. Los algoritmos y la IA basada en *big data* juegan un papel crucial en la Política Criminal, con sistemas empleados en la actualidad en todo el mundo con resultados satisfactorios. Sus aplicaciones son de lo más variadas y comprenden la predicción del riesgo de reincidencia del delincuente, el asesoramiento a unidades policiales para prevenir la comisión de delitos de manera eficaz, la defensa de la víctima a través de medidas de protección disuasivas y la dispensa de un tratamiento más adecuado a los reos en el ámbito penitenciario, entre otras.
8. El Derecho Penal debe afrontar el desafío de la irrupción de la neurociencia y la IA en nuestro mundo, acometiendo una reforma de las disposiciones legales acorde a las circunstancias. De particular interés resultan la atribución de responsabilidad penal a estos sistemas (o al programador que los diseña) y la definición de su estatuto jurídico, así como la creación de nuevos bienes jurídicos a proteger a través de la ley.
9. No debemos olvidarnos del impacto que el desarrollo de estas tendencias puede tener en los Derechos Humanos. Amén de su colisión con derechos clásicos como la intimidad o la libre determinación de la personalidad, cabe preguntarse si deben crearse nuevos derechos que, en consonancia con la política legislativa

preponderante, sirvan para proteger al ser humano de las posibles amenazas que pueden romper la armonía de nuestro mundo.

10. Algunos ordenamientos jurídicos ya incluyen en sus instrumentos legales diferentes disposiciones que consagran nuevos derechos y regulan el uso de estas tecnologías. Aunque queda mucho camino por recorrer, se observa ya un incipiente interés de los gobernantes en regular la incorporación de estas innovaciones que prometen jugar un rol central en el devenir de nuestra sociedad.

7. BIBLIOGRAFÍA

CITAS Y REFERENCIAS BIBLIOGRÁFICAS¹⁰⁸

Adamczyk, A. K. y Zawadzki, P. (2020). The memory-modifying potential of optogenetics and the need for neuroethics, *NanoEthics*, 14, 207-225. DOI: 10.1007/s11569-020-00377-1

Aguilar Cárceles, M. M. (2022). Neurociencia, culpabilidad y patología mental. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (241-253). Tirant lo Blanch.

Alonso Álamo, M. (2018). Culpabilidad jurídico-penal, motivos y emociones. La existencia enigmática de la libertad de la voluntad. *Revista penal México*, 14-15, 9-45.

Alonso Álamo, M. (2022). Acción y Derecho Penal: problemas de delimitación entre capacidad de acción y capacidad de culpabilidad. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (329-347). Tirant lo Blanch.

Anderson, C. (2008). The end of theory: the data deluge makes the scientific method obsolete, *Wired Magazine*, 16(7). <https://www.wired.com/2008/06/pb-theory/> (consultado por última vez el 1 de mayo de 2023)

Andrés Pueyo, A. (2019). RisCanvi: preveure el risc de tornar a delinquir. *Els Matins*, 18 de diciembre de 2019. Fragmento disponible en línea en: <https://www.ccma.cat/tv3/alcanta/els-matins/riscanvi-peveure-el-risc-de-tornar-a-delinquir/video/5986786/> Programa completo: <https://www.ccma.cat/tv3/alcanta/els-matins/els-matins-18122019/video/5986806/> (consultados por última vez el 1 de mayo de 2023)

¹⁰⁸ Las citas y referencias del presente trabajo se han redactado siguiendo el estilo de la American Psychological Association (APA), en su séptima y última edición hasta la fecha (2019). Criterios disponibles en línea en: <https://apastyle.apa.org/> y <https://apastyle.apa.org/style-grammar-guidelines/references/examples>, entre otros.

Angwin, J., Larson, J., Mattu, S. y Kirchner, L. (2016). Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks. *ProPublica*. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (consultado por última vez el 1 de mayo de 2023)

Benjamin, R. (2019). *Race after technology: abolitionist tools for the New Jim Code*. Polity.

Bertolino, M. (2008). Il "breve" cammino del vizio di mente. Un ritorno al paradigma organicistico? *Criminalia*, 3, 325-346.

Brennan, T., Dieterich, W. y Ehret, B. (2009). Evaluating the predictive validity of the COMPAS Risk and Needs Assessment System. *Criminal Justice and Behavior*, 36(1), 21-40. DOI: 10.1177/0093854808326545

Bublitz, J. C. (2013). My mind is mine!? Cognitive liberty as a legal concept. En E. Hildt y A. G. Franke (eds.), *Cognitive Enhancement: an interdisciplinary perspective* (233-264). Springer. DOI: 10.1007/978-94-007-6253-4_19

Busato, P. C. (2022). De máquinas y seres vivos: ¿quién actúa en los resultados delictivos derivados de decisiones cibernéticas? En E. Demetrio Crespo (Ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (349-376). Tirant lo Blanch.

Burns, J. M. y Swerdlow, R. H. (2003). Right orbitofrontal tumor with pedophilia symptom and constructional apraxia sign. *Archives of Neurology*, 60(3), 437-440. DOI: 10.1001/archneur.60.3.437

Caro Coria, D. C. (2022). Compliance, neurociencias e Inteligencia Artificial. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (633-655). Tirant lo Blanch.

Chiao, V. (2018). Predicting proportionality: the case for algorithmic sentencing. *Criminal Justice Ethics*, 37(3), 238-261. DOI: 10.1080/0731129X.2018.1552359

- Clark, A. y Chalmers, D. (1998). The Extended Mind. *Analysis*, 58(1), 7-19.
- Coccaro, E. F. (1989). Central serotonin and impulsive aggression. *The British Journal of Psychiatry*, 155(S8), 52-62. DOI: 10.1192/S0007125000291769
- Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M. y Damasio, A. R. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science*, 264(5162), 1102-1105. DOI: 10.1126/science.8178168
- De Sousa, S. A. (2020). “Não fui eu, foi a máquina”: teoria do crime, responsabilidade e Inteligência Artificial. En A. Miranda Rodrigues (ed.), *A Inteligência Artificial no Direito Penal* (59-94). Almedina.
- Demetrio Crespo, E. (2017). *Fragmentos sobre neurociencias y Derecho Penal*. B de F.
- Domínguez Luelmo, A. A. (2007). *Derecho sanitario y responsabilidad médica. Comentarios a la Ley 41/2002, de 24 de noviembre, sobre derechos de paciente, información y documentación clínica*. Lex Nova.
- Farwell, L. A. (2012). Brain fingerprinting: a comprehensive tutorial review of detection of concealed information with event-related brain potentials. *Cognitive Neurodynamics*, 6(2), 115-154. DOI: 10.1007/s11571-012-9192-2
- Feijoo Sánchez, B. J. (2017). *Derecho Penal, Neurociencias y Bien Jurídico*. Ediciones Olejnik.
- Feijoo Sánchez, B. J. (2022). Neurociencias y Derecho Penal de un sistema de libertades. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (127-145). Tirant lo Blanch.
- Ferguson, C. J. (2010). Genetic contributions to antisocial personality and behavior: a meta-analytic review from an evolutionary perspective. *The Journal of Social Psychology*, 150, 160–180. DOI: 10.1080/00224540903366503

Feresin, E. (2009). Lighter sentence for murderer with “bad genes”. *Nature*. 30 de octubre de 2009. DOI: 10.1038/news.2009.1050

Férez Mangas, D. (2017a). *Eficàcia del RisCanvi Complet en la predicció del trencament de permís de sortida*. Centro de Estudios Jurídicos y Formación Especializada. <https://cejfe.gencat.cat/web/.content/home/recerca/catalog/crono/2017/EficaciaRisCanvi.pdf> (consultado por última vez el 1 de mayo de 2023)

Férez Mangas, D. (2017b). *Factores de riesgo y predicción del quebrantamiento de los permisos penitenciarios de salida* [Tesis doctoral, Universidad de Barcelona]. Dipòsit Digital <https://diposit.ub.edu/dspace/handle/2445/155369>

Férez Mangas, D. y Andrés Pueyo, A. (2018). Eficacia predictiva en la valoración del riesgo de quebrantamiento de permisos penitenciarios. *La Ley Penal: revista de Derecho Penal, Procesal y Penitenciario*, 134, 1-14.

Feuerbach, P. J. A. (1800). *Revision der Grundsätze und Grundbegriffe des positiven peinlichen Rechts (Zweiter Teil)*. Henning.

Fink, L. D. (2022). *To our shareholders*. BlackRock (24-03-2022). <https://www.blackrock.com/corporate/investor-relations/larry-fink-chairmans-letter> (consultado por última vez el 13 de abril de 2023)

Fried, I., Mukamel, R. y Kreiman, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron*, 69(3), 548-562. DOI: 10.1016/j.neuron.2010.11.045

Glenn, A. L. y Raine, A. (2013). Neurocriminology: implications for the punishment, prediction and prevention of criminal behaviour. *Nature Reviews Neuroscience*, 15(1), 54-63. DOI: 10.1038/nrn3640

González Álvarez, J. L. y Garrido Antón, M. J. (2015). Satisfacción de las víctimas de violencia de género con la actuación policial en España. Validación del sistema VioGén. *Anuario de Psicología Jurídica*, 25, 29-38.

González Álvarez, J. L., López Ossorio, J. J., Urruela Cortés, C. y Rodríguez Díaz, M. (2018). Integral monitoring system in cases of gender violence. VioGén system. *Behavior and Law Journal*, 4(1), 29-40. DOI: 10.47442/blj.v4.i1.56

González Tapia, M. I. (2022). Perfil neuropsicológico del delincuente violento: implicaciones político-criminales. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (255-271). Tirant lo Blanch.

Grandi, C. (2016). *Neuroscienze e responsabilità penale. Soluzioni nuove per problema antichi?* Giappichelli Editore.

Grandi, C. (2022). El papel de la neurociencia en el juicio de la imputabilidad: el debate teórico y las consecuencias prácticas en la experiencia italiana. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (377-391). Tirant lo Blanch.

Greco, L. (2020). Richterliche Macht ohne richterliche Verantwortung: warum es den Roboter-Richter nicht geben darf. *Rechtswissenschaft*, 11(1), 29-62. DOI: 10.5771/1868-8098-2020-1-29

Greco, L. (2022). ¿Qué pueden aprender los penalistas de los neurocientíficos? En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (161-172). Tirant lo Blanch.

Grosse-Wilde, T. (2022). <<Predictive Sentencing>>: just a hype or a revolution? En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (449-472). Tirant lo Blanch.

Günther, K. (2005). *Schuld und kommunikative Freiheit*. Vittorio Klostermann.

Haissiner, M. y Pastor, D. R. (2019). *Neurociencia, Tecnologías Disruptivas y Tribunales Digitales*. Hammurabi.

Hassemer, W. (2009). Grenzen des Wissens im Strafprozess. Neuvermessung durch die empirischen *Wissenschaften vom Menschen?* *Zeitschrift für die gesamte Strafrechtswissenschaft*, 121(4), 829-859. DOI: 10.1515/ZSTW.2009.829

Hassemer, W. (2011). Neurociencias y culpabilidad en Derecho Penal. *InDret: Revista para el Análisis del Derecho*, 2, 1-15.

Hempel, C. (1942). The function of general laws in History. *Journal of Philosophy*, 39(2), 35-48.

Hempel, C. (1965). *Aspects of scientific explanation and other essays in the Philosophy of Science*. New York Free Press.

Henrich, J., Heine, S. J. y Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2-3), 61-83. DOI: <https://doi.org/10.1017/S0140525X0999152X>

Hillenkamp, T. (2005). Strafrecht ohne Willensfreiheit? Eine Antwort auf die Hirnforschung. *JuristenZeitung*, 60(7), 313-320.

Houeto, J. L., Mesnage, V., Mallet, L., Pillon, B., Gargiulo, M., Tezenas du Moncel, S., Bonnet, A. M., Pidoux, B., Dormont, D., Cornu, P. y Agid, Y. (2002). Behavioural disorders, Parkinson's Disease and subthalamic stimulation. *Journal of Neurology, Neurosurgery & Psychiatry*, 72(6), 701-707. DOI: <https://doi.org/10.1136/jnnp.72.6.701>

Ienca, M. y Andorno, R. (2022). Hacia nuevos Derechos Humanos en la era de la neurociencia y la neurotecnología. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (701-712). Tirant lo Blanch.

Jakobs, G. (1989). Über die Behandlung von Wollensfehlern und von Wissensfehlern. *Zeitschrift für die gesamte Strafrechtswissenschaft*, 101(3), 516-537. DOI: 10.1515/zstw.1989.101.3.516

Jakobs, G. (2005). Individuum und Person. Strafrechtliche Zurechnung und die Ergebnisse moderner Hirnforschung. *Zeitschrift für die gesamte Strafrechtswissenschaft*, 117(2), 247-266. DOI: <https://doi.org/10.1515/zstw.2005.117.2.247>

Jiménez Martínez, C. (2022). Irrupción de las pruebas de neuroimagen en España. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (473-488). Tirant lo Blanch.

Juliá Pijoan, M. (2020). *Proceso penal y (neuro)ciencia: una interacción desorientada*. Marcial Pons.

Juliá Pijoan, M. (2022). Neuropredicción: un nuevo ropaje para un planteamiento atávico. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (549-563). Tirant lo Blanch.

Köberer, W. (1996). *Index non calculat – Über die Unmöglichkeit, Strafzumessung sozialwissenschaftlich-mathematisch zu rationalisieren*. Peter Lang.

Kolla, N., Patel, R., Meyer, J. H. y Chakravarty, M. M. (2017). Association of monoamine oxidase-A genetic variants and amygdala morphology in violent offenders with antisocial personality disorder and high psychopathic traits. *Scientific Reports*, 7(1), 1-13. DOI: [10.1038/s41598-017-08351-w](https://doi.org/10.1038/s41598-017-08351-w)

Levine, E. S., Tisch, J., Tasso, A., y Joy, M. (2017). The New York City police department's domain awareness system. *Interfaces*, 47(1), 70–84. DOI: [10.1287/inte.2016.0860](https://doi.org/10.1287/inte.2016.0860)

Libet, B., Wright, E. W. y Gleason, C. A. (1982). Readiness-potentials preceding unrestricted "spontaneous" vs. pre-planned voluntary acts. *Electroencephalography & Clinical Neurophysiology*, 54(3), 322–335. DOI: [https://doi.org/10.1016/0013-4694\(82\)90181-X](https://doi.org/10.1016/0013-4694(82)90181-X)

Libet, B., Gleason, C. A., Wright, E. W. y Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (Readiness-Potential): the unconscious initiation of a freely voluntary act. *Brain*, 106(3), 623-642. DOI: [10.1093/brain/106.3.623](https://doi.org/10.1093/brain/106.3.623)

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8(4), 529 – 566. DOI:

<https://doi.org/10.1017/S0140525X00044903>

López Ossorio, J. J., González Álvarez, J. L. y Andrés Pueyo, A. (2016). Eficacia predictiva de la valoración policial del riesgo de la violencia de género. *Psychosocial Intervention*, 25(1), 1-7. DOI: <https://doi.org/10.1016/j.psi.2015.10.002>

López Ossorio, J. J., González Álvarez, J. L., Muñoz Vicente, J. M., Urruela Cortés, C. y Andrés Pueyo, A. (2019). Validation and calibration of the spanish police intimate partner violence risk assessment system (VioGén). *Journal of Police and Criminal Psychology*, 34(1), 1-11. DOI: 10.1007/s11896-019-09322-9

Lösel, F. y Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: a comprehensive meta-analysis. *Journal of Experimental Criminology*, 1(1), 117-146. DOI: 10.1007/s11292-004-6466-7

Liu, H. W., Lin, C. F. y Chen, Y. J. (2019). Beyond State v Loomis: Artificial Intelligence, government algorithmization and accountability. *International Journal of Law and Information Technology*, 27(2), 122-141. DOI: 10.1093/ijlit/eaz001

Marcus, G. y Davis, E. (2019). *Rebooting AI – building Artificial Intelligence we can trust*. Pantheon.

Martín López, M. T. (2022). Neurodesarrollo infantil: conducta antisocial y criminalidad. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (285-313). Tirant lo Blanch.

Martínez Garay, L. (2018). Peligrosidad, algoritmos y due process: el caso State v Loomis. *Revista de Derecho Penal y Criminología*, 20, 485-502. DOI: 10.5944/rdpc.20.2018.26484

Martínez Garay, L. (2022). ¿Ciencia o alquimia? Algoritmos y transparencia en la valoración del riesgo de reincidencia. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García

de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (489-505). Tirant lo Blanch.

Meijer, E. H., Ben-Shakhar, G., Verschuere, B. y Donchin, E. (2013). A comment on Farwell (2012): brain fingerprinting: a comprehensive tutorial review of detection of concealed information with event-related brain potentials. *Cognitive Neurodynamics*, 7(2), 155-158. DOI: 10.1007/s11571-012-9217-x

Merkel, G. (2022). Criminal law and brain decoding. In the need of a right to mental integrity? En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (733-746). Tirant lo Blanch.

Mir Puig, S. (2015). *Derecho Penal. Parte General*. Reppertor.

Miró Llinares, F. (2018). Inteligencia Artificial y justicia penal: más allá de los resultados lesivos causados por robots. *Revista de Derecho Penal y Criminología*, 20, 87-130. DOI: 10.5944/rdpc.20.2018.26446

Miró Llinares, F. y Castro Toledo, F. J. (2022). ¿Correlación no implica causalidad? El valor de las predicciones algorítmicas en el sistema penal a propósito del debate epistemológico sobre el <<Fin de la Teoría>>. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (507-529). Tirant lo Blanch.

Mohler, G. O., Short, M. B., Malinowski, S., Johnson, M., Tita, G. E., Bertozzi, A. L., y Brantingham, P. J. (2015). Randomized controlled field trials of predictive policing. *Journal of the American Statistical Association*, 110(512), 1399–1411. DOI: 10.1080/01621459.2015.1077710

Muñoz Conde, F. (1989). *Teoría General del Delito*. Tirant lo Blanch.

Muñoz Ortega, J. M. (2022). La Optogenética: fundamentos y retos ético-legales. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.),

Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial (747-758). Tirant lo Blanch.

Navarro Cardoso, F. (2022). Blockchain, Smart Contract y Compliance: Anotaciones para el Derecho Penal y Procesal de la Persona Jurídica. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (673-698). Tirant lo Blanch.

New, A. S., Buchsbaum, M. S., Hazlett, E. A., Goodman, M., Königsberg, H. W., Lo, J., Iskander, L., Newmark, R. E., Brand, J., O'Flynn, K. y Siever, L. J. (2004). Fluoxetine increases relative metabolic rate in prefrontal cortex in impulsive aggression. *Psychopharmacology*, 176(3-4), 451-458. DOI: 10.1007/s00213-004-1913-8

Nombela Otero, C. (2022). ¿Existe un algoritmo que explique el comportamiento humano? En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (315-326). Tirant lo Blanch.

Pacheco Rodríguez, M. A. (2022). Transhumanismo: desafíos éticos y jurídicos (Breve Inventario de conceptos y problemas). En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (759-769). Tirant lo Blanch.

Palma, M. F. (2022). El Derecho como Universo de Sentido para la Interfaz Cerebro e Inteligencia Artificial. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (173-186). Tirant lo Blanch.

Pappadopulos, E., Woolston, S., Chait, A., Perkins, M., Connor, D. F. y Jensen, P. S. (2006). Pharmacotherapy of aggression in children and adolescents: Efficacy and effect size. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 15(1), 27-39.

Pärnamets, P., Johansson, P., Hall, L., Balkenius, C., Spivey, M. J. y Richardson, D. C. (2015). Biasing Moral Decisions by exploiting the Dynamics of Eye Gaze. *Proceedings of the National Academy of Sciences*, 112(13), 4170-4175. DOI: 10.1073/pnas.1415250112

Pascual Botrán, M. J. (12 de marzo, 2023). Los abogados de Valladolid ensayan con la inteligencia artificial para agilizar tareas. *El Norte de Castilla*. <https://www.elnortedecastilla.es/valladolid/abogados-valladolid-ensayan-20230312192800-nt.html>

Pastor, D. R. (2022). ¿Sueña el Sistema Penal con Jueces Electrónicos? En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (531-548). Tirant lo Blanch.

Pauen, M. y Roth, G. (2008). *Freiheit, Schuld und Verantwortung. Grundzüge einer naturalistischer Theorie der Willensfreiheit*. Suhrkamp Verlag.

Pearl, J. y Verma, T. S. (1995). A Theory of Inferred Causation. *Studies in Logic and the Foundations of Mathematics*, 134(1), 789-811. DOI: 10.1016/S0049-237X(06)80074-1

Pearl, J. (2009). *Causality: Models, Reasoning and Inference*. Cambridge University Press.

Pearl, J. (2010). An Introduction to Causal Inference. *The International Journal of Biostatistics*. 6(2), 7, DOI: 10.2202/1557-4679.1203

Pearl, J., Glymour, M. y Jewell, N. P. (2016). *Causal Inference in Statistics: A Primer*. Wiley.

Pearl, J. y Mackenzie, D. (2018). *The Book of Why: the new Science of Cause and Effect*. Basic Books.

Pearson, F. S., Lipton, D. S., Cleland, C. M. y Yee, D. (2002). The Effects of Behavioral/Cognitive-Behavioral Programs on Recidivism. *Crime and Delinquency*, 48(3), 476-496. DOI: 10.1177/001112870204800306

Pérez Barberá, G. (2011). *El dolo eventual. Hacia el abandono de la idea de dolo como estado mental*. Hammurabi.

Pérez Manzano, M. (2011). Fundamento y fines del Derecho penal. Una revisión a la luz de las aportaciones de la neurociencia. *InDret: Revista para el Análisis del Derecho*, 2, 1-40.

Pérez Manzano, M. (2022). Reflexiones sobre el concepto de dolo a la luz de los desarrollos de la Neurociencia Cognitiva. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (393-416). Tirant lo Blanch.

Popper, K. R. (1971). *La Lógica de la Investigación Científica*. Tecnos.

Porciúncula, J. C. (2022). Technologies of the Extended Mind and Nemo Tenetur Principle. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (565-577). Tirant lo Blanch.

Raine, A. (1993). *The psychopathology of crime: criminal behavior as a clinical disorder*. Academic Press.

Raine, A. (17-21 de febrero, 2011). *Nature, nurture and antisocial behavior: biological and biosocial research on crime*. [Comunicación en Congreso]. Annual Meeting of the American Association for the Advancement of Science, Washington.

Redondo Illescas, S. (2008). Individuos, sociedades y oportunidades en la explicación y prevención del delito: Modelo del Triple Riesgo Delictivo (TRD). *Revista Española de Investigación Criminológica*, 7(6), 1-53. DOI: 10.24310/Boletin-criminologico.2008.v14i.8754

Roberts, R. (Presentador) (2020). Melanie Mitchell On Artificial Intelligence. *Econtalk Podcast*, <https://www.econtalk.org/melanie-mitchell-on-artificial-intelligence/> (consultado por última vez el 1 de mayo de 2023)

Röder, K. D. A. (1870). *Las Doctrinas Fundamentales reinantes sobre el Delito y la Pena en sus interiores contradicciones: ensayo crítico preparatorio para la renovación del derecho penal* (trad. Francisco Giner). Ediciones Revista General de Legislación y Jurisprudencia.

Rodríguez, V. G. (2022). Correccionalismo y no-repetición: el papel de las neurociencias en un derecho penal sin dolor. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García

de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (199-220). Tirant lo Blanch.

Romeo Casabona, C. M. (2018). Riesgo, procedimientos actuariales basados en inteligencia artificial y medidas de seguridad. *Revista de Derecho, Empresa y Sociedad*, 13, 39-55.

Romeo Casabona, C. M. (2022). Avances Científicos y Derecho Penal. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (77-107). Tirant lo Blanch.

Romero Flores, B. (2015). Las neurociencias frente a la función de la pena. *Anuario de Derecho Penal y Ciencias Penales*, 68(1), 335-357.

Rosburg, T., Pflueger, M. O., Mokros, A., Boillat, C., Deuring, G., Spielmann, T. y Graf, M. (2020). Indirect and neuropsychological indicators of pedophilia. *Sexual Abuse: A Journal of Research and Treatment*, 33(1), 1-27. DOI: 10.1177/1079063220931822

Rossmo, D. K. (1995). *Profiling: Target Patterns of Serial Murders*. [Tesis Doctoral, Simon Fraser University]. Simon Fraser University Library. <https://summit.sfu.ca/item/6820>

Roth, G. (2004). Worüber dürfen Hirnforscher reden - und in welcher Weise? *Deutsche Zeitschrift für Philosophie*, 52(2), 223-234. DOI: <https://doi.org/10.1524/dzph.2004.52.2.223>

Roth, G. y Merkel, G. (2008). Bestrafung oder Therapie? – Möglichkeiten und Grenzen staatlicher Sanktion unter Berücksichtigung der Hirnforschung. En *Hirnforschung – Chancen und Risiken für das Recht*, 21. Rechtswissenschaftliche Fakultät der Universität Zürich.

Roxin, C. (1973). *Strafrechtliche Grundlagenprobleme*. De Gruyter. DOI: <https://doi.org/10.1515/9783110890525>

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1, 206-215. DOI: 10.1038/s42256-019-0048-x

Ruggeri, S. (2022). Investigaciones Neurocientíficas, Prueba Penal y Nuevas Dinámicas de Determinación de los Hechos. El Impacto de las Neurociencias sobre las Garantías Individuales y el Respeto de los Estándares de Fairness en el Juicio Criminal. En E. Demetrio Crespo (ed.), M. De la Cuerda Martín y F. García de la Torre García (coords.), *Derecho Penal y Comportamiento Humano. Avances desde la Neurociencia y la Inteligencia Artificial* (579-615). Tirant lo Blanch.

Sahel, J. A., Boulanger-Scemama, E., Pagot, C., Arleo, A., Galluppi, F., Martel, J. N., Degli Esposti, S., Delaux, A., De Saint Aubert, J. B., De Montleau, C., Gutman, E., Audo, I., Duebel, J., Picaud, S., Dalkara, D., Blouin, L., Taiel, M. y Roska, B. (2021). Partial recovery of visual function in a blind patient after optogenetic therapy. *Nature Medicine*, 27(7), 1223-1229. DOI: 10.1038/s41591-021-01351-4

Sale, A., Berardi, N. y Maffei, L. (2014). Environment and Brain Plasticity: Towards an Endogenous Pharmacotherapy. *Physiological Reviews*, 94(1), 189-234. DOI: 10.1152/physrev.00036.2012

Sankin, A., Mehrotra, D., Mattu, S., Cameron, D., Gilbertson, A., Lempres, D. y Lash, J. (2021). Crime Prediction Software Promised to Be Free of Biases. New Data Shows It Perpetuates Them. *The Markup* y *Gizmodo*. <https://gizmodo.com/crime-prediction-software-promised-to-be-free-of-biases-1848138977> y <https://themarkup.org/prediction-bias/2021/12/02/crime-prediction-software-promised-to-be-free-of-biases-new-data-shows-it-perpetuates-them> (consultados por última vez el 1 de mayo de 2023)

Sartori, G. y Zangrossi, A. (2016). Neuroscienze forensi. *Giornale Italiano di Psicologia*, 4, 689-711.

Schüpbach, M., Gargiulo, M., Welter, M. L., Mallet, L., Béhar, C., Houeto, J. L., Maltête, D., Mesnage, V. y Agid, Y. (2006). Neurosurgery in Parkinson disease – A distressed mind in a repaired body? *Neurology*, 66(12), 1811-1816. DOI: 10.1212/01.wnl.0000234880.51322.16

Sensi, M., Eleopra, R., Cavallo, M. A., Sette, E., Milani, P., Quatrone, R., Capone, J. G., Tugnoli, V., Tola, M. R., Granieri, E. y Data, P. G. (2004). Explosive-aggressive behavior

related to bilateral subthalamic stimulation. *Parkinsonism & Related Disorders*, 10(4), 247-251. DOI: 10.1016/j.parkreldis.2004.01.007

Sententia, W. (2004). Neuroethical considerations: cognitive liberty and converging technologies for improving human cognition. *Annals of the New York Academy of Sciences*, 1013(1), 221-228. DOI: 10.1196/annals.1305.014

Singh, J. P. (2013). Predictive validity performance indicators in violence risk assessment: A methodological primer. *Behavioral Sciences and the Law*, 31(1), 8-22. DOI: 10.1002/bsl.2052

Snoek, J., Larochelle, H. y Adams, R. P. (2012). Practical Bayesian Optimization of Machine Learning Algorithms. *Advances in Neural Information Processing Systems*, 4, 2951-2959.

Soler Iglesias, C. V. (2019). RisCanvi: Preveure el risc de tornar a delinquir. *Els Matins*, 18 de diciembre de 2019. Fragmento disponible en línea en: <https://www.ccma.cat/tv3/alacarta/els-matins/riscanvi-peveure-el-risc-de-tornar-a-delinquir/video/5986786/> Programa completo: <https://www.ccma.cat/tv3/alacarta/els-matins/els-matins-18122019/video/5986806/> (consultados por última vez el 1 de mayo de 2023)

Tagarro, A. (23 de abril, 2022). Judea Pearl, el padre de la Inteligencia Artificial: “Los robots hablarán entre ellos, tendrán deseos, voluntad propia...no sé qué le sorprende de esto” *XL Semanal*. <https://www.elcorreo.com/xlsemanal/personajes/robots-inteligentes-premio-bbva-fronteras-conocimiento-humanos-judea-pearl-ingeniero-filosofo.html>

Tagarro, A. (23 de abril, 2022). La Escalera de la Inteligencia Artificial según el matemático Judea Pearl. *XL Semanal*. <https://www.elcorreo.com/xlsemanal/ciencia/robots-judea-pearl-inteligencia-artificial.html>

Tiarks, E. (2021). The impact of Algorithms on Legitimacy in Sentencing. *Journal of Law, Technology and Trust*, 2(1), 1-23. DOI: 10.19164/jltt.v2i1.1151

Urbaniok, F., Laubacher, A., Hardegger, J., Rosseger, A., Endrass, J. y Moskvitin, K. (2011). Neurobiological Determinism: Human Freedom of Choice and Criminal Responsibility.

International Journal of Offender Therapy and Comparative Criminology, 56(2), 174-190. DOI: 10.1177/0306624X10395474

Villamarín López, M. L. (2014). *Neurociencia y detección de la verdad y del engaño en el proceso penal: el uso del escáner cerebral (fMRI) y del brainfingerprinting (p300)*. Marcial Pons.

Von Linstow, B. (1974). *Berechenbares Strafmaß. Eine Neue Methode der Strafzumessung am Beispiel wichtiger Verkehrsdelikte*. J. Schweitzer Verlag.

Weinberger, D. (2018). La Inteligencia Artificial nos obliga a revisar nuestra Idea de Justicia. *El País* (27 de octubre de 2018). https://elpais.com/retina/2018/10/26/tendencias/1540549298_079415.html (consultado por última vez el 1 de mayo de 2023)

Woodward, J. y Ross, L. (2021). Scientific Explanation. *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/sum2021/entries/scientific-explanation>

Zurita Bayona, J. (2014). *Violencia contra la mujer. Marco histórico evolutivo y predicción del nivel de riesgo* [Tesis doctoral, Universidad Autónoma de Madrid]. E-Archivo (Biblioteca Universidad Autónoma). <https://repositorio.uam.es/handle/10486/661810>

JURISPRUDENCIA:

Tribunal Europeo de Derechos Humanos (Gran Sala). Caso Saunders c. Reino Unido. Sentencia 19187/91 de 17 de diciembre de 1996.

Supreme Court of Wisconsin. Caso State v. Loomis. Resolución 881N.W.2d749, de 13 de julio de 2016.

Bundesgerichtshof. Decisión de 16 de enero de 2010.

Corte di Cassazione (Sección 1ª). Sentencia 43021 de 7 de noviembre de 2012.

Corte di Cassazione (Sección 1ª). Sentencia 27129 de 21 de julio de 2016.

Corte di Cassazione (Sección 1ª). Sentencia 11897 de 18 de mayo de 2018.

Corte di Cassazione (Sección 1ª). Sentencia 26895 de 12 de junio de 2018.

Corte d'Assise d'Appello di Trieste. Sentencia 5 de 18 de septiembre de 2009.

Tribunale di Como. Sentencia 536, de 20 de mayo de 2011.

Tribunal Supremo (Sala de lo Civil, Sección 1ª). Sentencia de 16 de enero de 2008.

Tribunal Supremo (Sala de lo Penal, Sección 1ª). Sentencia de 5 de mayo de 2020.

Tribunal Superior de Justicia de Castilla-La Mancha (Sala de lo Civil y Penal). Sentencia 16/2019, de 13 de junio.

Audiencia Provincial de Guadalajara (Sección 1ª), Tribunal del Jurado. Sentencia 3/2018, de 15 de noviembre.

Audiencia Provincial de Barcelona (Sección 21ª). Sentencia 353/2021, de 28 de enero.

Audiencia Provincial de León (Sección 3ª). Sentencia 683/2022, de 1 de julio.

INSTRUMENTOS LEGALES:

Resolución del Parlamento Europeo de 16 de febrero de 2017. Disponible en línea en:

<https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52017IP0051&from=EN>

Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión, de 21 de abril de 2021. Disponible en línea en:

https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0008.02/DOC_1&format=PDF

Constitución Política de la República de Chile (versión vigente desde el 28 de abril de 2021 hasta el 24 de octubre de 2021, consultado por última vez el 19 de abril de 2023). Disponible en línea en: <https://www.bcn.cl/leychile/navegar?idNorma=242302&idVersion=2021-04-28>

Constitución Política de la República de Chile (versión vigente a 19 de abril de 2023, consultado por última vez el 19 de abril de 2023). Disponible en línea en: <https://www.bcn.cl/leychile/navegar?idNorma=242302&idVersion=2023-02-03&idParte=>

Bundesverfassung der Schweizerischen Eidgenossenschaft (versión vigente a 19 de abril de 2023, consultado por última vez el 19 de abril de 2023). Disponible en línea en: <https://www.fedlex.admin.ch/eli/cc/1999/404/de>

Constituição da República Portuguesa (versión vigente, consultado por última vez el 19 de abril de 2023). Disponible en línea en: <https://www.parlamento.pt/Legislacao/Paginas/ConstituicaoRepublicaPortuguesa.aspx>

Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal. Consultado por última vez el 17 de marzo de 2023. Disponible en línea en: <https://www.boe.es/buscar/act.php?id=BOE-A-1995-25444>

Ley Orgánica 15/2003, de 25 de noviembre, por la que se modifica la Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal. Disponible en línea en:
<https://www.boe.es/buscar/doc.php?id=BOE-A-2003-21538>

Ley Orgánica 1/2004, de 28 de diciembre, de Medidas de Protección Integral contra la Violencia de Género. Disponible en línea en:
<https://www.boe.es/buscar/act.php?id=BOE-A-2004-21760>

Ley Orgánica 1/2015, de 30 de marzo, por la que se modifica la Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal. Disponible en línea en:
<https://www.boe.es/buscar/doc.php?id=BOE-A-2015-3439>

Ley Orgánica 14/2022, de 22 de diciembre, de transposición de directivas europeas y otras disposiciones para la adaptación de la legislación penal al ordenamiento de la Unión Europea, y reforma de los delitos contra la integridad moral, desórdenes públicos y contrabando de armas de doble uso. Disponible en línea en:
<https://www.boe.es/buscar/act.php?id=BOE-A-2022-21800>

Gobierno de España. Carta de Derechos Digitales. Consultado por última vez el 17 de marzo de 2023. Disponible en línea en:
https://www.lamoncloa.gob.es/presidente/actividades/Documents/2021/140721-Carta_Derechos_Digitales_RedEs.pdf

NOTICIAS EN PRENSA:

La Stampa. *“E’ vulnerabile geneticamente”*. 25 de octubre de 2009. <https://www.lastampa.it/cronaca/2009/10/25/news/e-vulnerabile-geneticamente-1.37057128/>

El País. *Un tribunal italiano acepta que la predisposición genética sea una eximente*. 2 de noviembre de 2009. https://elpais.com/sociedad/2009/11/02/actualidad/1257116411_850215.html

Wired. *UK police are using AI to inform custodial decisions – but it could be discriminating against the poor*. 1 de marzo de 2018. <https://www.wired.co.uk/article/police-ai-uk-durham-hart-checkpoint-algorithm-edit> (consultado por última vez el 1 de mayo de 2023)

La Vanguardia. *Un algoritmo impreciso condiciona la libertad de los presos*. 6 de diciembre de 2021. <https://www.lavanguardia.com/vida/20211206/7888727/algoritmo-sirve-denegar-permisos-presos-pese-fallos.html> (consultado por última vez el 1 de mayo de 2023)

20 Minutos. *Cómo funciona VioGén, la plataforma del Gobierno para evitar asesinatos relacionados con la violencia de género*. 13 de enero de 2023. <https://www.20minutos.es/tecnologia/actualidad/como-funciona-viogen-la-plataforma-del-gobierno-para-evitar-asesinatos-relacionados-con-la-violencia-de-genero-5091796/> (consultado por última vez el 1 de mayo de 2023)

The New York Times. *The Godfather of A.I.’ Leaves Google and Warns of Danger Ahead*. 1 de mayo de 2023. <https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html>

BBC. *AI 'godfather' Geoffrey Hinton warns of dangers as he quits Google*. 3 de mayo de 2023. <https://www.bbc.com/news/world-us-canada-65452940> (consultado por última vez el 3 de mayo de 2023)

OTROS RECURSOS EN LÍNEA:

Sobre COMPAS:

Equivant (2019). *Practitioner's Guide to COMPAS Core*. Disponible en línea en: <https://www.equivant.com/wp-content/uploads/Practitioners-Guide-to-COMPAS-Core-040419.pdf> (consultado por última vez el 1 de mayo de 2023)

Northpoint (2009). *Measurement & Treatment Implications for COMPAS Core Scale*. [https://www.michigan.gov/-/media/Project/Websites/corrections/progserv/Folder1/Timothy Brenne PhD Meaning and Treatment Implications of COMPAS Core Scales.pdf](https://www.michigan.gov/-/media/Project/Websites/corrections/progserv/Folder1/Timothy_Brenne_PhD_Meaning_and_Treatment_Implications_of_COMPAS_Core_Scales.pdf) (consultado por última vez el 1 de mayo de 2023)

Sobre Geolítica:

<https://blog.predpol.com/geolitica-a-new-name-a-new-focus> (consultado por última vez el 1 de mayo de 2023)

<https://geolitica.com/public-safety/> (consultado por última vez el 1 de mayo de 2023)

<https://geolitica.com/technology/> (consultado por última vez el 1 de mayo de 2023)

<https://www.predpol.com/about/> (consultado por última vez el 1 de mayo de 2023)

<https://www.predpol.com/technology/> (consultado por última vez el 1 de mayo de 2023)

Sobre VioGén:

Radio Televisión Española. ¿Qué es VioGén y cómo funciona? *Objetivo Igualdad*, Episodio 83: ¿Romper el techo de cristal? 19 de febrero de 2023. Fragmento disponible en: <https://www.rtve.es/noticias/20230220/objetivo-igualdad-funcionamiento-viogen/2424326.shtml>, Programa completo disponible en línea en: <https://www.rtve.es/play/videos/objetivo-igualdad/programa-83-romper-techo-cristal/6809486/> (consultado por última vez el 1 de mayo de 2023)

Ministerio del Interior. *Sistema de seguimiento integral en los casos de Violencia de Género (Sistema VioGén). Datos Estadísticos, Marzo 2023*.

<https://www.interior.gob.es/opencms/pdf/servicios-al-ciudadano/violencia-contra-la-mujer/estadisticas/2023/Estadistica-31-de-marzo-2023.pdf> (consultado por última vez el 1 de mayo de 2023)

<https://www.interior.gob.es/opencms/es/servicios-al-ciudadano/violencia-contra-la-mujer/sistema-viogen/> (consultado por última vez el 1 de mayo de 2023)

<https://www.interior.gob.es/opencms/es/servicios-al-ciudadano/violencia-contra-la-mujer/sistema-viogen/plan-de-seguridad-personalizado-ppsp/> (consultado por última vez el 1 de mayo de 2023)

<https://www.interior.gob.es/opencms/es/servicios-al-ciudadano/violencia-contra-la-mujer/supresion-de-datos-sistema-viogen/> (consultado por última vez el 1 de mayo de 2023)

Sobre RisCanvi

Departamento de Justicia (Generalitat de Catalunya). *Manual d'aplicació del protocol de valoració RisCanvi*.

https://justicia.gencat.cat/web/.content/home/ambits/reinsercio_i_serveis_peni/manual-aplicacio-protocol-avaluacio-riscanvi.pdf (consultado por última vez el 1 de mayo de 2023)