

CARMEN VÁZQUEZ

(ED.)

**ESTÁNDARES DE PRUEBA
Y PRUEBA CIENTÍFICA**

Ensayos de epistemología jurídica

Marcial Pons

MADRID | BARCELONA | BUENOS AIRES | SÃO PAULO
2013

La colección *Filosofía y Derecho* publica aquellos trabajos que han superado una evaluación anónima realizada por especialistas en la materia, con arreglo a los estándares usuales en la comunidad académica internacional.

Los autores interesados en publicar en esta colección deberán enviar sus manuscritos en documento *Word* a la dirección de correo electrónico manuscritos@filosofiyderecho.es. Los datos personales del autor deben ser aportados en documento aparte y el manuscrito no debe contener ninguna referencia, directa o indirecta, que permita identificar al autor.

En caso de ser aceptada la publicación del original, el autor deberá adaptarlo a los criterios de la colección, los cuales se pueden encontrar, en formato PDF, en la página web www.filosofiyderecho.es.

Quedan rigurosamente prohibidas, sin la autorización escrita de los titulares del «Copyright», bajo las sanciones establecidas en las leyes, la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, y la distribución de ejemplares de ella mediante alquiler o préstamo públicos.

© Carmen Vázquez (ed.)

© Los autores

© MARCIAL PONS

EDICIONES JURÍDICAS Y SOCIALES, S. A.

San Sotero, 6 - 28037 MADRID

☎ (91) 304 33 03

www.marcialpons.es

ISBN: 978-84-15664-53-6

Depósito legal: M. 8.434-2013

Diseño de la cubierta: Manuel Estrada. Diseño Gráfico

Fotocomposición: JOSER TRATAMIENTO DE TEXTOS, S. L.

Impresión: ELECÉ INDUSTRIA GRÁFICA, S. L.

Polígono El Nogal - Río Tiétar, 24 - 28110 Algete (Madrid)
MADRID, 2013

ÍNDICE

	Pág.
A MODO DE PRESENTACIÓN, Carmen Vázquez	11
1. SOBRE LA EPISTEMOLOGÍA JURÍDICA	11
2. ESTÁNDARES DE PRUEBA	13
3. LA PRUEBA PERICIAL-CIENTÍFICA	16
4. ELEMENTOS Y ESTÁNDARES DE PRUEBA	19
I. LA PRUEBA ES LIBERTAD, PERO NO TANTO: UNA TEORÍA DE LA PRUEBA CUASIBENTHAMIANA, Jordi Ferrer Beltrán	21
1. INTRODUCCIÓN	21
2. LOS MOMENTOS DE LA ACTIVIDAD PROBATORIA EN EL PROCESO JUDICIAL	24
2.1. La conformación del conjunto de elementos de juicio o pruebas	24
2.2. La valoración de los elementos de juicio o pruebas	26
2.3. La adopción de la decisión sobre los hechos probados	27
3. LA CRÍTICA DE STEIN A LA TRADICIÓN RACIONALISTA	28
4. RECONSIDERANDO LAS TESIS DE PARTIDA	33
BIBLIOGRAFÍA	37
II. LOS ESTÁNDARES DE PRUEBA Y LOS LÍMITES DEL ANÁLISIS JURÍDICO, Ronald J. Allen	41
1. LA IMPORTANCIA DE LAS CARGAS DE PERSUASIÓN	43

	INDICE
	Pág.
2. LOS LÍMITES DEL ANÁLISIS JURÍDICO	48
BIBLIOGRAFÍA	63
III. EL PROBABILISMO JURÍDICO: UNA DISENSIÓN EPISTEMOLÓGICA, Susan Haack.....	65
1. LOS ESTÁNDARES DE PRUEBA DEBEN SER ENTENDIDOS COMO GRADOS DE AVAL.....	66
2. LOS GRADOS DE AVAL NO SON PROBABILIDADES MATEMÁTICAS.....	75
3. EL CASO <i>COMMONWEALTH V. SACCO AND VANZETTI</i> : EL FUNDAMENTO DE LA FUNCIÓN MEJOR QUE EL BAYESIANISMO SUBJETIVO.....	83
4. <i>PEOPLE V. COLLINS</i> : UNA VEZ MÁS, EL FUNDAMENTO DE LA FUNCIÓN MEJOR QUE EL BAYESIANISMO SUBJETIVO.....	89
BIBLIOGRAFÍA.....	95
IV. ESTÁNDARES DE PRUEBA Y TEORÍA DE LA PRUEBA, Michael S. Pardo.....	99
1. CRITERIOS TEÓRICOS Y PRESUPUESTOS CLARIFICADOS.....	100
1.1. El requisito de nivel micro.....	101
1.2. El requisito de nivel macro.....	102
1.3. El requisito de integración.....	104
2. TEORÍAS PROBABILÍSTICAS DE LA PRUEBA Y ESTÁNDARES DE PRUEBA PROBABILÍSTICOS.....	105
2.1. El nivel micro.....	105
2.2. El nivel macro.....	110
2.3. Integración.....	112
3. UNA TEORÍA EXPLICATIVA DE LA PRUEBA.....	113
3.1. El nivel micro.....	114
3.2. El nivel macro.....	115
3.3. Integración.....	116
BIBLIOGRAFÍA.....	117
V. LA ELEMENTAL ARITMÉTICA EPISTEMICA DEL DEBERECHO II: LOS INAPROPIADOS RECURSOS DE LA TEORÍA MORAL PARA ABORDAR EL DERECHO PENAL, Larry Laudan.....	119
1. PERSPECTIVA GENERAL.....	120

	INDICE
	Pág.
2. PARTE I: LOS FORCEJOS DE LA DEONTOLOGÍA CON EL ESTÁNDAR DE PRUEBA.....	121
3. PARTE II: ¿PUEDE LA TEORÍA MORAL DEONTOLOGICA EXPLICAR O JUSTIFICAR REGÍMENES PUNITIVOS ESPECÍFICOS?.....	126
4. CONCLUSIÓN.....	132
BIBLIOGRAFÍA.....	133
VI. ¿FUE EL BEBÉ SACUDIDO? PRUEBA, PERICIA Y EPISTEMOLOGÍA JURÍDICA EN EL PROCESO PENAL INGLÉS, Paul Roberts.....	135
1. INTRODUCCIÓN: LA PREOCUPACIÓN EPISTEMICA EN LA JURISDICCIÓN PENAL.....	135
2. HISTORIA DE LOS CASOS.....	138
3. LA JURISDICCIÓN DE APELACIONES EN INGLATERRA Y GALES: EL OBJETO Y LA RAZÓN DE SER DE LAS APELACIONES PENALES EN UN SISTEMA PROCESAL ADVERSARIAL.....	140
4. ERRORES DE DERECHO Y DETERMINACIÓN DE LOS HECHOS EN LAS APELACIONES PENALES DE CONDENAS DICTADAS POR JURADOS.....	148
5. PRECEDENTES JURÍDICOS Y HECHOS CIENTÍFICOS.....	159
6. INCERTIDUMBRE CIENTÍFICA Y DESACUERDO ENTRE PERITOS.....	164
7. CONFLICTOS SOBRE LA AUTORIDAD DEL PERITO Y SU CREDIBILIDAD.....	169
8. CONCLUSIÓN: HACIENDO JUSTICIA A LA PRUEBA PERICIAL Y A LA PRUEBA CIENTÍFICA.....	175
BIBLIOGRAFÍA.....	177
VII. PRUEBA CIENTÍFICA. UN MAPA DE RETOS, Martha Gascón Abellán.....	181
1. PRUEBA CIENTÍFICA. SOBREALVALORACIONES Y PARADIGMAS.....	181
1.1. El mito.....	181
1.2. Sobrealvaloración epistémica.....	183
1.3. Sobrealvaloración semántica.....	185
2. EL PARADIGMA DE LA VEROSIMILITUD.....	188
2.1. Lo que dicen los datos.....	188

	Pág.
2.2. Lo que debe creerse.....	188
2.3. Lo que debe hacerse.....	189
3. LOS RETOS.....	190
3.1. Sobre la admisibilidad de la prueba científica. El juez como <i>gatekeeper</i>	191
3.2. Sobre el modo de transmitir los resultados de la prueba científica.....	194
3.3. Sobre la valoración de la prueba.....	196
BIBLIOGRAFÍA.....	199
VIII. LA APLICACIÓN DE ESTÁNDARES CIENTÍFICOS A LAS CIENCIAS SOCIALES Y FORENSES, <i>Michele Taruffo</i>.....	203
BIBLIOGRAFÍA.....	212
IX. UNA APROXIMACIÓN PRAGMATISTA AL TESTIMONIO COMO EVIDENCIA, <i>Andrés Páez</i>.....	215
1. INTRODUCCIÓN.....	215
2. LOS PROBLEMAS DE LAS APROXIMACIONES REDUCTIVAS AL TESTIMONIO.....	218
3. APROXIMACIONES REGLA-REDUCTIVAS AL TESTIMONIO.....	219
4. LA INVESTIGACIÓN COMO UN PROBLEMA DE DECISIONES.....	222
5. LA BÚSQUEDA DE UNA JUSTIFICACIÓN PARA EL TESTIMONIO.....	224
6. LA ACEPTACIÓN DEL TESTIMONIO COMO EVIDENCIA.....	233
7. CONCLUSIÓN.....	236
BIBLIOGRAFÍA.....	237

A MODO DE PRESENTACIÓN

Carmen VÁZQUEZ
Universitat de Girona

1. SOBRE LA EPISTEMOLOGÍA JURÍDICA

Esta compilación de artículos es el resultado de un taller que, con el mismo título del libro, organizó el grupo de filosofía del derecho de la Universidad de Girona en mayo de 2011¹. Los temas abordados, el diseño de estándares de prueba y la valoración de la prueba científica, son sin lugar a dudas dos de los más problemáticos en el ámbito de la epistemología jurídica. Por ello, decidimos dedicarle esas jornadas de estudio. En el evento los autores presentaron una versión preliminar de los trabajos aquí compilados, que fueron debatidos con todos los asistentes durante dos días de intenso y muy fructífero trabajo. A los autores, al público asistente, así como a los diversos traductores les agradezco su valiosa contribución al resultado de esta obra².

Son dos las características que inicialmente quiero destacar en esta breve introducción, pues hacen poco común este escenario de análisis en el ámbito

¹ Para el cual se recibió financiamiento de la AGAUR (Generalitat de Catalunya) a través de la ayuda para la organización de congresos ARCS (2010_ARCS1_00141) y de la ahora Secretaría de Estado de Investigación, Desarrollo e Innovación mediante la Acción Complementaria 2010 (DER2010-11788-E —subprograma JURI—) al proyecto «Ciencia y proceso judicial: prueba y atribución de responsabilidad» (DER2010-21331-C02-02).

² A título personal quiero agradecer, en primer lugar, a Jordi FERRER la invaluable oportunidad de editar este libro y todo el apoyo recibido durante esta edición; así como a mis compañeros del área de filosofía del derecho de la UdG que colaboraron en la organización del evento. Y en segundo lugar, a Maximiliano ARAMBURU, Gonzalo SEUAS, Diego DEI VECCI, Carlos BERNAL y María José VIANA muchas gracias por toda su disponibilidad (y paciente apertura para discutir y revisar los textos) durante la traducción de los artículos.

jurídico y, además, permiten entender mejor la estructura de este libro: la confluencia de dos tradiciones (la del *common law* y la de derecho romano-germánico) y la aplicación de desarrollos teóricos de un área del conocimiento no-jurídica (la epistemología) a problemas eminentemente jurídicos, y quizá como corolario de esto, la participación de expertos de esta área.

En cuanto a la confluencia de tradiciones, comúnmente se destacan las diferencias generales entre sus modelos de proceso judicial. Así, por ejemplo, se suele decir que el *common law* tiene un proceso marcadamente adversarial en el que las partes (o, quizá mejor dicho, los abogados de éstas) juegan el rol primordial y, además, se caracterizaría por la potencial presencia como juzgador de los hechos de un jurado lego que no tiene obligación de motivar sus veredictos. Mientras, el proceso de tradición romano-germánica se suele plantear como un escenario donde el juez sigue teniendo en mayor o menor medida un rol más activo entre los actores y, por otro lado, el juzgador de los hechos sería en la gran mayoría de los casos un juez profesional con obligación constitucional de motivar sus sentencias. Con cierta independencia tanto de la corrección de estas amplísimas descripciones así como de sus matices y alcances, estas diferencias se acentúan en los temas concretos sobre los que tratan los trabajos de este libro, es decir, en la necesidad de establecer y aplicar estándares de prueba y en el tratamiento de la ahora en boga prueba científica.³

Así pues, cabe advertir al lector que cuando en los diversos textos se habla de «sistema» o «práctica judicial», muchas veces se está haciendo referencia exclusivamente al contexto jurídico de los autores, por lo que se debe ser prevenido y crítico, no trasplantando de un sistema a otro posibles problemas o lo que se podrían calificar como soluciones viables sin el análisis oportuno. Hay que considerar con especial cuidado que en el mundo anglosajón muchas de las reglas probatorias y el análisis teórico respectivo versan sobre qué pruebas serán presentadas ante el jurado y, por ello, los controles de admisibilidad de los elementos de juicio resultan comparativamente mucho más importantes en su derecho probatorio que en los sistemas de tradición romano-germánica.⁴

³ Temas en los que la bibliografía y la jurisprudencia anglosajonas son, desde hace muchos años, extensas y ricas. Parte de esta tradición probatoria se ha reflejado en la elaboración de regulaciones concretas al respecto; ejemplo de ello son las *Federal Rules of Evidence* estadounidenses (FRE), de las que se encuentran varias referencias en el texto, y la aparición de influyentes escuelas de pensamiento, como la llamada *Rationalist Tradition of Evidence* o la *New Evidence Scholarship*.

⁴ Una de las características del sistema anglosajón, predominantemente en el estadounidense, son las reglas que excluyen pruebas *relevantes* no sólo por cuestiones de licitud o de limitaciones temporales, sino para mantener cierto control sobre el jurado y sobre los abogados. Muchas veces se enfatiza que la exclusión de pruebas se hace con el objetivo de proteger al jurado (e indirectamente a las partes) de los efectos que puedan tener en su decisión cierto tipo de pruebas; con independencia de la corrección de este argumento (psicologista), vale la pena tomar en cuenta que son pocos los procesos que realmente lleguen a fase de juicio oral ante jurado, la mayoría termina en fases anteriores, lo que cuanto menos resta importancia a esta situación. Por lo que toca a la exclusión de pruebas como medio para controlar a las partes y a sus abogados, se fundamenta por algunos autores estadounidenses como un mecanismo para obligar a las partes a presentar mejores pruebas, en una suerte de prevención de errores.

Ahora bien, a pesar de las posibles especificidades de cada práctica jurídica-procesal concreta, se ha de enfatizar que los trabajos compilados tienen una pretensión teórica mucho más amplia, dados los problemas que se abordan y la forma en que se hace. Sin duda, como advierte el subtítulo de la obra, se trata de una compilación de teoría de la prueba jurídica o, mejor aún, de *epistemología jurídica*.⁵ En este escenario, es sin duda un privilegio contar con grandes epistemólogos con intereses teóricos en el derecho, como HAACK y LAUDAN, y con un reconocido filósofo de la ciencia, PAEZ, aportando una mirada refrescante a las muchas veces cerradas estructuras del pensamiento jurídico.

Las concepciones sostenidas por los autores son desde luego diversas. Sin embargo, podríamos decir que todos asumen, como común denominador, la averiguación de la verdad como objetivo fundamental del proceso. Por otro lado, al centrar la atención en las decisiones (y las reglas afines) que *con base en las pruebas*, y no meramente como resultado de cierto estado mental, toma el juez y/o el juzgador de los hechos⁶, consideran indispensable el análisis de cuestiones relacionadas con las creencias, la justificación, el conocimiento, la calidad de los elementos de juicio (*evidence*), los estándares probatorios, qué cuenta como determinación de los hechos (*proof*), etc. Aunque el libro se centra sólo en el tema de estándares y cierto tipo de elementos de juicio, están presupuestas discusiones más profundas sobre todas estas cuestiones epistemológicas.

2. ESTÁNDARES DE PRUEBA

El análisis, e incluso la conciencia de la necesidad, de los estándares de prueba, desafortunadamente, es muy poco común en los sistemas de tradición romano-germánica.⁷ Pareciera que bajo el llamado sistema de la libre valoración

⁵ «Epistemología jurídica» o «*legal epistemology*» han sido expresiones hasta ahora usadas habitualmente para referirse de forma exclusiva a un tipo de análisis de los problemas (o de algunos problemas) de la prueba judicial. Sin embargo, el adjetivo «jurídica», obviamente, va mucho más allá del contexto probatorio y/o procesal, por lo que podrá alegarse que los problemas de la llamada epistemología jurídica tratan más allá de este ámbito. A manera de ejemplo, podría aludirse a la pregunta: ¿es posible conocer el derecho (entendido como proposiciones normativas) de forma neutral? Sin embargo, reconociendo esto, en el marco de éste libro, por epistemología jurídica sólo se hace alusión al contexto procesal probatorio.

⁶ «Juez» y «juzgador de los hechos» no son sinónimos. Es importante, sobre todo para la comprensión de muchos de los problemas que abordan los estudiosos de sistemas del *common law*, distinguir las actividades jurisdiccionales que en materia probatoria se efectúan *previamente* a la etapa de valoración de elementos de juicio y toma de decisión, v. gr. la admisión y exclusión de pruebas. Puede darse el caso de que un mismo sujeto (el juez profesional) efectúe todas estas funciones, como suele ser en los sistemas romano-germánicos o en los llamados *bench trials* estadounidenses. Ahora bien, en los sistemas del *common law* esta distinción se refleja sobre todo en la (debatida) separación entre «cuestiones de derecho» y «cuestiones sobre los hechos»: las primeras recaen siempre sobre un juez profesional, mientras las segundas pueden recaer sobre un juzgador de los hechos lego i. e., el jurado lego.

⁷ Diría que, en castellano, FERREER BERRÁN es el primer estudio del derecho en profundizar sobre este problema, ofreciendo no sólo todo un andamiaje para el análisis de la institución probatoria

ción de la prueba todo lo que importa es si, habiendo oído todos los elementos de juicio admitidos y practicados, la mera creencia de los juzgadores de los hechos es firme y sólida. Pero, cabe preguntarse, ¿en qué situación queda el ciudadano que está siendo parte de un proceso si aunque se le asegure un derecho al debido proceso (*in itinere*) y a la prueba, finalmente termina siendo condenado o declarado responsable cuando el juzgador considera a bien que está *persuadido, firmemente convencido*, etc.?

En este sentido, es indispensable enfatizar la imperiosa necesidad de establecer un umbral que determine el grado de prueba mínimo para «tener por probado que *p*» o, lo que es lo mismo, para que el juzgador tome la decisión en uno u otro sentido. Para esto último, no es suficiente la comparación entre todas las pruebas que pretenden avalar las dos hipótesis fácticas en juego, y no sólo porque posiblemente estén establecidas cargas de prueba distintas, sino porque la hipótesis que se declarará probada debería necesariamente alcanzar cierto grado de prueba. Este grado, por otro lado, debe ser previamente conocido por las partes procesales, o más precisamente, debería estar establecido en la legislación correspondiente.

A efectos de situar mejor la discusión sobre los estándares de prueba es necesario resaltar dos cuestiones. Primera, no se trata de una especie de regreso a las reglas de prueba tasada, pues los estándares de prueba suelen plantearse principalmente en la valoración del conjunto de pruebas a efectos de la toma de la decisión final; no afectan a elementos de prueba concretos estableciendo su valor a priori, como sí lo hacen las reglas de prueba tasada y sustituyendo en gran medida la actividad evaluativa o los criterios del juzgador (actuando, en este sentido, como una especie de presunción *iuris et de iure*)⁸. Segunda, el estándar de prueba presupone una decisión de política pública sobre el beneficio de la duda que se pretende dar a cada una de las partes implicadas y, con ello, la *distribución* de errores entre las mismas que se busca conseguir en un proceso judicial. En este sentido, a la epistemología sólo le corresponde la *construcción* del estándar de prueba, mas no determinar el grado mínimo indispensable, *i. e.*, epistemológicamente se podría decir el cómo pero no el cuánto.

en general, sino un ejemplo de cómo formular un estándar adecuado epistemológicamente y exigente, si así se ha decidido previamente, en materia penal.

⁸ En los textos compilados, las discusiones sobre los estándares son fundamentalmente en este sentido, *i. e.*, para la toma de la decisión final. Sin embargo, en la literatura especializada también se discute la necesidad de establecer estándares para la toma de decisiones previas a la decisión final como, por ejemplo, decretar un arraigo o incluso dictar prisión preventiva o medidas de alejamiento, etc. Para las cuales, el conjunto de elementos de pruebas disponibles es, obviamente, diferente.

Por otro lado, se ha planteado también cuán exigentes deberían ser partiendo de ciertos criterios establecidos para la valoración de elementos de juicio concretos; pensemos por ejemplo en la fiabilidad, al ser un criterio gradual, la pregunta sería ¿cuán fiable debería ser la prueba para ser admisible? Sin embargo, en mi opinión, esta respuesta corresponde al criterio mismo en el caso concreto y no a un estándar de prueba en juego.

Pues bien, en la práctica anglosajona se tienen establecidos diferentes estándares más o menos exigentes, no carentes de otros serios problemas, para los diversos tipos de procesos. Así, en el ámbito penal el conocido *beyond a reasonable doubt* (más allá de toda duda razonable), mucho más exigente que *the more probable than not* (más probable que su negación), *preponderance of the evidence* (preponderancia de la prueba)⁹ o *clear and convincing evidence* (prueba clara y convincente) usados en el ámbito civil. Se asume, claramente, que los costes de los errores en uno u otro tipo de procesos son muy diversos y, por ello, requieren estándares de prueba también diferentes; ahora bien, una de las discusiones actuales en materia procesal penal es si *todos* los delitos, presumiendo que la gravedad de las conductas consideradas delictivas (y su pena) son distintas, deberían decidirse uniformemente por un mismo umbral tan exigente o si lo más adecuado sería diseñar estándares ajustados a esta situación¹⁰.

Este ya complejo panorama quizá nos da la pauta de la gran diversidad de cuestiones que el tema de los estándares de prueba exige abordar, así pues, haciendo las veces de hilo conductor del análisis, veremos en los siguientes artículos distintos «puntos de salida» con el objetivo de entender mejor lo que está en juego en la práctica probatoria. Veámos.

Jordi FERRER, inicia la discusión sobre estándares de prueba con su artículo «La prueba es libertad pero no tanto: una teoría de la prueba cuasi-benthamiana», dibujando un panorama sobre los diversos momentos probatorios y el papel que en cada uno de ellos juega (o debería jugar) el derecho y la epistemología. Los argumentos vertidos vienen a propósito de la conocida tesis abolicionista de BENTHAM que, como el nombre indica, pretende sostener la necesidad de abolir el derecho probatorio en pro de la averiguación de la verdad. El autor cuestiona el alcance de dicha tesis sosteniendo que los criterios epistemológicos, al no tener carácter práctico, son insuficientes para guiarnos en la toma de decisión.

Ron ALLEN, por su parte, en «Los estándares de prueba y los límites del análisis jurídico» analiza algunos problemas en la conceptualización de la llamada «carga de persuasión». Asumiendo ciertas descripciones empíricas del ámbito probatorio, critica algunos presupuestos usados por teóricos que

⁹ Los estándares civiles de la «preponderancia de la prueba» y el «más probable que su negación», se han usado tradicionalmente de forma intercambiable, como sinónimos. Sin embargo, el primero de ellos parecería establecer que para su satisfacción es suficiente con tener simplemente *más pruebas* que la contraparte, lo que constituye un error: *ambas* partes podrían tener un conjunto de pruebas débil o insuficiente para cumplir con lo exigido por el estándar, aunque una siga preponderando sobre la otra. Por ello algunos teóricos han manifestado su preferencia en usar «más probable que su negación» en lugar de la preponderancia de la prueba.

¹⁰ Piénsese, por ejemplo, en los costes que puede llegar a tener la pérdida de la patria potestad en materia civil frente a los costes de un delito contra la seguridad vial. Pareciera que los bienes jurídicos en juego ameritan una mayor exigencia en el primero y no tan alta en el segundo.

pretenden desarrollar un proyecto normativo del mismo; y concluye que sus insuficiencias sugieren que las herramientas analíticas empleadas más que esclarecer el objeto de investigación, lo conciben erróneamente, por lo que son necesarias otras formas alternativas de análisis que intentará apuntar.

Susan HAACK, por otro lado, en «El probabilismo jurídico: Una dimensión epistemológica», aborda la compleja cuestión sobre qué tipo de grados son exactamente los grados de prueba, ofreciéndonos su teoría fundarrentista (una aproximación tripartita de la prueba) como alternativa a las concepciones probabilísticas en este contexto. Mostrando, además, su aplicabilidad en dos famosos casos estadounidenses que le sirven como un buen ejemplo práctico.

Michael PARDO, en su trabajo «Estándares de prueba y teoría de la prueba», nos brinda un panorama general sobre los aspectos que toda propuesta teórica sobre el proceso probatorio deberá tomar en cuenta; para ello, identifica un nivel macro, un nivel micro y la integración entre éstos, planteando problemas de distinta índole en cada uno. Después de analizar los distintos problemas que las concepciones probabilísticas de la prueba tienen en dichos niveles, ofrece como alternativa su concepción de la prueba jurídica como inferencia a la mejor explicación; ésta, en su opinión, resulta compatible con la mejor descripción empírica con que se cuenta hasta ahora en el ámbito anglosajón (*Story Model*).

Y Larry LAUDAN en «La elemental aritmética epistémica del derecho II», intenta demostrar la arbitrariedad de la idea de que una falsa condena es mucho peor que una falsa absolución, subyacente en los estándares de prueba en materia penal. Para ello, nos ofrece una fuerte crítica a los proyectos deontológicos, dada su imposibilidad para graduar objetivamente los daños morales relevantes y poder especificar cuál es la pena merecida para cada delito. Según el autor, deberíamos saber cuán lesivo es un daño para determinar cuán exigente debería ser el estándar probatorio.

3. LA PRUEBA PERICIAL-CIENTÍFICA

Por otro lado, en los últimos años, debido en parte a la especialización (y a veces hiperespecialización) de nuestras sociedades y/o el significativo y creciente impacto que la ciencia y la tecnología tienen en nuestra vida cotidiana, la prueba pericial, en concreto la llamada prueba científica, ha cobrado suma importancia en la práctica jurisdiccional, siendo también cada vez más los estudios al respecto en el ámbito teórico-dogmático. En este sentido, resulta muy conveniente preguntarse no sólo *cómo* han afrontado este panorama los sistemas jurídicos sino precisamente *cuál* es el reto (o los retos) que se afrontan con respecto al conocimiento experto relevante para la toma de decisiones judiciales.

Si pretendemos responder al cómo, deberíamos decir que los sistemas de derecho romano-germánico han conrado tradicionalmente con los llamados peritos oficiales y/o se ha considerado que todo perito tiene una obligación primordial con el tribunal y/o la impartición de justicia (estableciendo incluso diversos tipos de sanciones en casos de incumplimiento). Por ello, uno de los criterios fundamentales (si no el criterio o metacriterio fundamental) para valorar la prueba pericial ha sido la *imparcialidad* del perito, entendida básicamente como ausencia de relación con las partes. Sin embargo, aun cuando la imparcialidad del perito es importante, podría *no* determinar la calidad del peritaje: un perito imparcial puede cometer errores, tener creencias falsas, usar técnicas con grados de fiabilidad cuestionables, ser incompetente, etc. Entonces, considerar única y exclusivamente este criterio oscurece otro tipo de parcialidades, digamos, más cognoscitivas que personales, reduciendo en un sentido importante lo sustantivo a la forma.

Por lo que toca a los sistemas del *common law*, por otra parte, la prueba pericial tradicionalmente se ha vehiculado a través de un «*expert witness*» (testigo experto), implicando que el perito es considerado un *testigo de parte* y, por definición, parcial. Así pues, sabiendo de esta parcialidad pericial de origen, se han preocupado fundamentalmente en ofrecer criterios que permitían al juez directamente valorar la *información* ofrecida por el perito. Sitúndonos en los Estados Unidos, por ejemplo, desde 1923 con la resolución del caso *Frye v. United States*, ha estado sobre la mesa de discusión el criterio de la «aceptación general de la comunidad científica relevante»; o, a partir de 1993, con la resolución del caso hito en el tratamiento judicial de la *expert evidence*, *Daubert v. Merrill Dow, Inc.*, la llamada «fiabilidad probatoria» y sus muy diversos indicios sugeridos tanto por los tribunales como por la doctrina en los últimos años, centrando el debate básicamente en esa cuestión.¹¹

En este amplísimo panorama, se inserta el trabajo de Paul ROBERTS, ¿Fue el bebé sacudido?, analizando la resolución en apelación de un trío de casos penales donde se discutían hechos concernientes a la muerte y/o las lesiones de bebés o infantes en los que, dadas las circunstancias, la prueba fundamental fue de carácter pericial. Se explicitan entonces los problemas enfrentados en

¹¹ En los últimos años el impacto del caso *Daubert* ha sido muy significativo en diversos países, no sólo del *common law* sino también de nuestra tradición; aunque, desafortunadamente, pocas veces media un análisis profundo sobre la viabilidad del trasplante. Así, por ejemplo, en mi opinión, deben distinguirse al menos tres cuestiones implícitas en *Daubert* cuyas problemáticas son diversas: el rol del juez como *gatekeeper* o vigilante, el meta-criterio de fiabilidad y los criterios que servirían de indicios de la fiabilidad. Respecto al primero, los controles de admisibilidad en Estados Unidos suponen la presencia de un juez de la causa (precisamente, que hace las veces de vigilante de la entrada) diferente al juzgador de los hechos (jurado), al sugerir la adopción de *Daubert* en este sentido, ¿se está sugiriendo también que nuestros sistemas prevean un juez de admisión diverso al juzgador que valora las pruebas? O, de lo contrario, ¿esa «prevaloración» de la calidad de las pruebas no se convierte más bien en una doble valoración innecesaria? Por lo que toca a la fiabilidad y sus posibles indicios, ¿qué exigencias tiene este criterio? ¿los indicios de fiabilidad propuestos son aplicables a toda prueba pericial y accesibles para los juzgadores?, etcétera.

la revisión de la valoración probatoria de controvertidas hipótesis médico-científicas hecha por los jurados y la importancia de considerar ciertas especificidades de una práctica procesal concreta.

Marina GASCÓN, en su trabajo «Prueba Científica. Un mapa de retos», refiriéndose fundamentalmente a los problemas que presentan las llamadas ciencias forenses en lo que identifica como el paradigma (jurídico) de la individualización, sugiere el uso del llamado Teorema de Bayes para la valoración particular de este tipo de pruebas. Esto, en su opinión, formaría parte de un indispensable cambio de paradigma (al de la verosimilitud) en el que, además de haber filtros más fuertes de admisibilidad, se haga una clara distinción entre la tarea del juez y la del perito.

Por su parte, en «La aplicación de estándares científicos a las ciencias sociales y forenses», Michele TARUFFO nos ofrece un panorama general de la problemática de las pruebas científicas haciendo alusión a la experiencia estadounidense más reciente, más allá de *Daubert*, poniendo sobre la mesa la necesidad de contar con mayor información empírica sobre las ciencias que se usan como pruebas.

Ahora bien, una cuestión fundamental a responder es precisamente *cuál* es el reto que se pretende afrontar en el tratamiento de la prueba pericial. Si bien es cierto que la respuesta quizá dependería en gran parte del sistema jurídico concreto en un momento determinado, las posibilidades epistemológicas deberían ser planteadas, básicamente: cierta deferencia (racional) a la autoridad epistémica experta o, por el contrario, un razonamiento inductivo del juez o juzgador a partir de cierta (meta)información considerada pertinente¹².

De inicio parecería asumirse que se requiere sólo un modelo normativo de prueba pericial, que los criterios de valoración sugeridos servirían o fracasarían para ese modelo, bien se trate, por ejemplo, de un perito de confianza del juez (oficial o no) o de un perito de parte, ambos deberían responder a ese modelo¹³. Sin embargo, creo que éste ha sido un grave error en el análisis de

¹² Evidentemente todo esto requiere una mayor explicación y fundamentación que no se puede desarrollar aquí. Sin embargo creo oportuno hacer dos observaciones concretas. Primera, la deferencia de la que se habla aquí es epistémica, no práctica; dice al juez lo que habría que creer y no lo que habría que hacer, por lo que no estoy decantándome por convertir al perito en el juzgador de los hechos (preocupación siempre latente del sistema). En todo caso, creo que esta preocupación (justificada) ha sesgado mucho las posibilidades de discusión racional y generado prácticas cuestionables (como que los tribunales acudan a expertos que consideran de su confianza a fin de abordar cuestiones periciales, sin explicitar esto en el caso correspondiente). Y, segunda, para exigir al juzgador que razone inductivamente sobre la prueba pericial en cuestión, se requeriría contar con mucha mayor información sobre la información subyacente a esas pruebas; lo que exigiría llevar a cabo investigaciones empíricas sólidas. Una tarea pendiente y muy apremiante a la fecha.

¹³ Al respecto, la doctrina procesalista española ha discutido durante varios años si la prueba pericial es una prueba en sí misma o se trata más bien de un «auxiliar del juez». Discusión que incluso ha llevado al pronunciamiento en la Ley de Enjuiciamiento Civil. En mi opinión, esta formidable intuición de los grandes dogmáticos procesalistas ha estado absolutamente carente de fundamentos epistemológi-

la prueba pericial y que podrían estar en juego al menos dos modelos muy diversos conviviendo pacíficamente en un sistema jurídico, pero accionándose con presupuestos diversos.

El fundamento para la afirmación anterior es netamente de carácter epistémico, dependiente de la justificación de las llamadas creencias o conocimiento testimonial. En este sentido, primero habría que reconocer, o explicitar, que epistemológicamente la prueba pericial es un tipo de testimonio, con absoluta independencia de las diferencias jurídico-procesales entre la llamada «prueba testimonial o testifical» y la «prueba pericial». En otras palabras, la información experta que entra en el proceso y, con ello, en conocimiento del juzgador es mediante un *tercero*, situado en mejor posición que él, que le proporciona dicha información¹⁴. ¿En qué circunstancias, pues, estaría el juez justificado para adquirir información testimonial experta de ese tercero? Los criterios para ello dependen, a su vez, de los criterios que justifican las creencias o conocimientos *testimonial* en general. Una relativamente reciente y muy interesante discusión en la epistemología entre dos posiciones, los reduccionistas y los no-reduccionistas, intentan responder a esta cuestión.

Y es, en esta discusión más general (no exclusivamente sobre el testimonio experto), donde cobra especial relevancia el artículo de Andrés PÁEZ, «Una aproximación pragmatista al testimonio como evidencia». Desde una perspectiva reduccionista, el autor considera que las reglas de inferencia y los mecanismos de adquisición de creencias son iguales para las creencias testimoniales como para las no-testimoniales, ofreciendo una justificación epistemológica de las creencias testimoniales y su aceptación deliberada, en contraposición a lo que llama el grado de confianza moral que la fuente inspira o la contradicción con las creencias personales.

4. ELEMENTOS Y ESTÁNDARES DE PRUEBA

Una vez presentados ambos temas, y ya para finalizar, es necesario aludir a la coherencia entre los niveles probatorios, esto es, entre los criterios con los que se valoran cada uno de los elementos de juicio (v. g.: las pruebas periciales) y los criterios empleados para la toma de decisión (estándares de prueba), a la par que sus diferencias. Pareciera un tanto absurdo, o al menos una incom-

co, lo que les ha llevado no sólo a seguir sin dar una respuesta medianamente clara al problema, sino a seguir planteándose la unicidad de la prueba pericial, en lugar de replantearse, por ejemplo, un modelo diferente para el perito de confianza del juez.

¹⁴ Es aquí donde se hace más latente el gran desafío que representa la prueba pericial para el proceso judicial. Al menos un desafío para una concepción del proceso que presuponga que el juzgador de los hechos debe tener cierta *experiencia directa* con los hechos en materia de prueba; así, por ejemplo, en nuestra tradición representa un tipo de reto para la inmediatez judicial, mientras en el *common law* se contraponen en cierta medida con la idea de ser juzgados por iguales con una cultura más bien promedio (los miembros del jurado).

parbididad pragmática, buscar criterios impersonales o interpersonales para la valoración de los diversos elementos y luego terminar asumiendo estándares de prueba eminentemente subjetivos; una situación particularmente recurrente en el análisis de la prueba científica donde, por muy diversos motivos, se plantean fuertes exigencias de valoración para luego ceder ante estándares subjetivos.

Hasta aquí parece que no es osado concluir también que el análisis de las diversas cuestiones probatorias debe verse como un todo o, si se quiere, como una especie de redes temáticas que interactúan constantemente. Así pues, los *cambios* que se sugieran para alguno de los aspectos probatorios tendrán, sin duda, influencia en otras cuestiones de la actividad probatoria en su conjunto; y, a su vez, las *explicaciones* o descripciones de esta actividad se entenderán mejor o serán más exitosas en la medida en que se considere ese todo. Ambos tipos de proyectos (normativos y descriptivos) sobre la prueba judicial, en mi opinión, deberían ir acompañados de criterios epistémicos que consideren la disminución del error global en las decisiones judiciales, sin olvidarse, en todo caso, de cuestiones como las especificidades relevantes de prácticas jurídico-procesales concretas, el análisis adecuado de ciertas decisiones político-institucionales y, por supuesto, la necesidad de tener más información empírica sobre cómo están funcionando nuestros actuales sistemas de justicia.

Pues bien, en términos sumamente generales, es este contexto donde tienen lugar una serie de fructíferas discusiones teóricas sobre dos temas trascendentales en el ámbito probatorio: los estándares de prueba y la prueba científica. Las implicaciones que ambos tienen en la práctica jurisdiccional aconsejan su análisis cuidadoso, un análisis comprometido con el objetivo de la búsqueda de la verdad en el proceso judicial y, con ello, con la protección de los derechos de los ciudadanos. Este volumen tiene la pretensión no sólo de presentar posturas y argumentos sobre esos temas, sino de invitar al debate de los mismos.

Girona, diciembre de 2012

LA PRUEBA ES LIBERTAD, PERO NO TANTO: UNA TEORÍA DE LA PRUEBA CUASIBENTHAMIANA *

I

Jordi FERRER BELTRÁN
Universitat de Girona

1. INTRODUCCIÓN

En el ámbito de estudio de la prueba, Jeremy BENTHAM es considerado uno de los principales precursores de la denominada tradición racionalista (especialmente en el mundo anglosajón, claro). Es más, se ha considerado, con razón, que la teoría benthamiana «constituye la más ambiciosa y completa teoría de la prueba en la historia del pensamiento jurídico» (TWNING, 1994: 39). Por ello, resulta de gran importancia hacer las cuentas con el pensamiento de BENTHAM si se quiere abordar el estudio de los problemas probatorios en el derecho.

No es mi intención, sin embargo, entrar en disquisiciones acerca de la mejor reconstrucción que pueda darse de ésta o aquella tesis del autor; tampoco iniciar un debate con intérpretes más autorizados que yo de su pensamiento. La modesta pretensión de este trabajo es reevaluar un aspecto muy específico de la teoría de la prueba propia de la tradición racionalista para el que BENTHAM puede ser llamado en causa como precursor: la idea de que la prueba

* Agradezco los muy útiles comentarios de Susan HAACK, Larry LAUDAN, Giovanni B. RATTI y Carmen VÁZQUEZ.

debe ser fundamentalmente libre y que cualquier interferencia del derecho procesal en ella debía ser limitada o eliminada. BENTHAM elabora una teoría prescriptiva de la prueba o, aún mejor para lo que nos importa, una teoría prescriptiva del derecho de la prueba, que es una parte más de su teoría general del derecho (KEETON y MARSHALL, 1948: 83; TWINING, 1985: 19-25; MORESO, 1992: 354).

La premisa básica de la que parte es simple y clara: la finalidad de la prueba en el proceso judicial es la averiguación de la verdad sobre lo ocurrido (BENTHAM, 1827: 337; MORESO, 1992: 354; TWINING, 1985: 27, y 1994: 38). Aunque, por supuesto, esta asunción de partida puede ser y ha sido muy discutida, creo que debe ser admitida salvo que se esté dispuesto a pagar altos costes en la reconstrucción del derecho (FERRER, 2005: 55 y ss.). El objetivo de la averiguación de la verdad deriva, por otro lado, de la asunción por parte de BENTHAM de que la finalidad fundamental del derecho procesal es la de ser garantía de la correcta aplicación del derecho sustantivo (BENTHAM, 1827: 17-19). Finalidad que, a su vez, se divide en la necesidad de una doble garantía: asegurar que *todos* los infractores del derecho sean sancionados y que sólo ellos lo sean¹. Y esto supone, evidentemente, la necesidad de que lo que se declare probado en el proceso coincida con la verdad de lo ocurrido: esto es, que los enunciados declarados probados sean verdaderos y los enunciados falsos no se declaren probados.

Siendo éstos los objetivos, queda por determinar la metodología adecuada para alcanzarlos. Y en este aspecto, BENTHAM es extraordinariamente crítico con el derecho probatorio de su época (críticas en buena medida aplicables a nuestros sistemas procesales actuales). Para BENTHAM, la mejor forma de alcanzar ese objetivo es mediante una metodología o sistema «natural» de conocimiento (TWINING, 1985: 47-52, y 1994: 39-41; HART, 1982: 32; POSTEMA, 1986: 344 y ss.), propia del sentido común y de la epistemología general (BENTHAM, 1827: 344), y no a través del denominado derecho probatorio que, en la Inglaterra de su época, estaba formado básicamente por un conjunto de reglas de exclusión que limitaban las posibilidades de conocimiento. Es más, BENTHAM (1838-1843, vol. 4: 334) considera que la justicia es algo bastante simple, «es lo mismo para un hombre que para otro: sólo los legisladores, aconsejados por los abogados, la han complicado y hecho trizas». Resulta tan simple que el mejor método para administrar justicia es el que él denomina modelo doméstico, propio del «paterfamilias» en el trato de los hijos. Por ello, lo mejor que podría hacer el derecho en materia de prueba es callar, dejando reinar al principio general de la libertad probatoria (*free proof*)².

¹ POSTEMA, 1986: 343, quien cita, a su vez, los *Bentham Manuscripts in the University College London Library*.

² En este sentido, sostiene TWINING que el derecho probatorio angloamericano de los dos últimos siglos puede ser entendido como un serie de excepciones al principio general de *free proof*, de inspiración thayeriana. Véase TWINING, 1994: 196. Véase también DAMAŠKA, 1997: 149.

Ésta es una característica propia de la denominada tradición racionalista clásica en materia de prueba. Como afirma el propio BENTHAM en el *Treatise on Judicial Evidence* «[e]n toda la trama de esta obra reaparece constantemente esta conclusión práctica: *No hay que excluir ninguna prueba, ningún testimonio por el sólo temor de ser engañado*» (BENTHAM, 1823: 441, cursiva original). Por ello, es necesario «sustituir el principio de exclusión por el de *suspiciata*» ante las pruebas disponibles³. El «paterfamilias» que juzga la conducta de sus hijos tiende a aceptar todos los elementos de juicio puesto que la racionalidad indica que cuanto más abundante sea el conjunto de elementos de juicio disponible mayor será la probabilidad de determinar la verdad de lo ocurrido. Es éste un corolario del reino de la epistemología en el proceso, buscado por BENTHAM, del método «natural» de la prueba (TWINING, 1994: 211, n. 3).

Y tampoco el derecho debe regular la valoración de la prueba (*weight of evidence*), puesto que ésta debe llevarla a cabo el juzgador, atendiendo a las circunstancias propias del caso y en aplicación del método de razonamiento «natural». Por ello, en conclusión, BENTHAM es absolutamente contrario a la existencia de reglas jurídicas vinculantes sobre la prueba: ésta es la denominada «*anti-nomian thesis*» de BENTHAM, que ofrece una imagen clásica de la tradición racionalista de la prueba y que ha sido resumida por TWINING en los siguientes términos: el sistema jurídico no debe contener «ninguna norma que excluya testigos o pruebas; ninguna norma sobre el peso o el *quantum* de la prueba; ninguna norma vinculante sobre la forma de presentación de la prueba; ninguna restricción artificial sobre los interrogatorios o el razonamiento probatorio; ningún derecho de silencio ni privilegios de los testigos; ninguna restricción al razonamiento que no sean las propias del razonamiento práctico; ninguna exclusión de pruebas excepto si son irrelevantes o superfluas o si su presentación supone perjuicios, gastos o retrasos excesivos en las circunstancias del caso específico» (TWINING, 1994: 195)⁴.

La recepción de estas tesis de BENTHAM, tanto en el área de *common law* como de *civil law*, ha conllevado interpretaciones en términos claramente abolicionistas del derecho probatorio. Un caso muy claro es, en la cultura jurídica de lengua castellana, el de SENTÍS MELENDO, quien afirma que «*Derecho y probatorio*, ese sustantivo y ese adjetivo, son términos antagónicos. La prueba es libertad; y en el momento en que el derecho intenta someterla a normas rígidas, deja de ser prueba, para convertirse, si no en una caricatura, en algo que ya no es prueba» (SENTÍS MELENDO, 1979: 336). Y en términos parecidos en

³ BENTHAM, 1823: 442. En la misma línea puede verse de forma reciente, por todos, LAUDAN, 2006: 122, 124.

⁴ La tarea del legislador, pues, no es la de dictar reglas imperativas sino simplemente instrucciones orientativas acerca de la valoración de la prueba y otorgar al juez los poderes apropiados para tomar las decisiones necesarias (BENTHAM, J.: «The Rationale of Judicial Evidence», en *id.*, 1838-43, vol. 6: 151. Véase también TWINING, 1985: 66-75; POSTEMA, 1986: 349; PARDO: 2005: 325).

cuanto al análisis, que no en cuanto a la propuesta abolicionista, puede entenderse el muy conocido pasaje de CARNELUTTI (1947: 33-34): «Basta un *limite minimo* a la libertad de búsqueda del juez para que el proceso de búsqueda de la verdad degenera en un proceso formal de fijación; en otros términos, la medida del límite es indiferente para la naturaleza del proceso». A SENTIS MELENDO y CARNELUTTI les separa muy claramente la apuesta posterior al análisis: el primero propugna la abolición del derecho probatorio y la vigencia de las reglas generales de la epistemología, mientras que el segundo propone abandonar los esquemas epistemológicos para estudiar la prueba jurídica; pero ambos comparten un claro punto de partida, *i. e.*, el derecho no debe regular la prueba ni las decisiones sobre esta si no se quiere deslindar completamente la prueba jurídica de la búsqueda de la verdad y de sus patrones epistemológicos.

Pues bien, en lo que sigue pretendo precisamente discutir este preciso punto de encuentro. Sostendré: i) que el derecho necesariamente debe regular (algunos aspectos de) la prueba judicial, aún asumiendo los postulados básicos de la tradición racionalista, y ii) que la influencia de BENTHAM y la tradición racionalista en el derecho probatorio (especialmente anglosajón) ha tenido éxito especialmente en aquellos puntos más débiles de su pensamiento.

2. LOS MOMENTOS DE LA ACTIVIDAD PROBATORIA EN EL PROCESO JUDICIAL

A los efectos de aportar nueva luz sobre las exigencias epistemológicas que pueden hacer compatible el derecho probatorio con el objetivo de la búsqueda de la verdad, conviene distinguir claramente tres momentos fundamentales en ese proceso de toma de decisiones sobre los hechos en el proceso judicial. Antes de identificarlos, no obstante, debo advertir que se trata de tres momentos lógicamente distintos y sucesivos, aunque en los procesos de toma de decisiones reales pueden presentarse entrelazados. Se trata de: *a)* la conformación del conjunto de elementos de juicio sobre cuya base se adoptará la decisión; *b)* la valoración de esos elementos, y *c)* propiamente, la adopción de la decisión.

2.1. La conformación del conjunto de elementos de juicio o pruebas

En primer lugar, el desarrollo del proceso judicial a través de la proposición y práctica de las pruebas debe permitir conformar un conjunto de elementos de juicio que apoyen o refuten las distintas hipótesis sobre los hechos del caso. Y aquí se da ya una de las especificidades jurídicas de mayor calado, que puede resumirse en la máxima *quid non est in actis non est in mundo*. Es

decir, a los efectos de la decisión jurídica el conjunto de elementos de juicio que podrá y deberá ser tomado en consideración está formado únicamente por las pruebas aportadas y admitidas al proceso⁵, no pudiéndose tomar en cuenta, por parte del órgano decisor, aquellas informaciones o elementos de juicio de los que disponga «privadamente» (STEIN, F., 1893: 71 y ss.; 150 y ss.; DAMAŠKA, 1986: 30, 138, 170-171) o aquellos que, habiéndose aportado al proceso han sido excluidos, por ejemplo, por su carácter ilícito. Creo que puede decirse sin excesivo riesgo de error que en todos los demás ámbitos del conocimiento, el conjunto de elementos de juicio que puede y debe ser analizado para adoptar una decisión racional es igual al total de las informaciones disponibles y relevantes para el caso⁶. En la prueba jurídica, en cambio, el conjunto de elementos a valorar es un subconjunto del conjunto formado por la totalidad de los elementos disponibles: aquellos de ellos que han sido incorporados al expediente judicial.

De este modo, adquiere una relevancia destacada el estudio de los filtros para la admisión de pruebas en el proceso. Un primer filtro, de orden epistemológico, prescribe la admisión de toda prueba que aporte información relevante sobre los hechos que se juzgan (BENTHAM, 1823: 391; TARUFFO, 1992: 337 y ss.; GASCON, 1999: 115). Una prueba es relevante si aporta apoyo o refutación de alguna de las hipótesis fácticas del caso a la luz de los principios generales de la lógica y de la ciencia (TWINING, 1994: 179). Puede considerarse este filtro, en realidad, como un principio general de inclusión. Funcionaría, así, prescribiendo la admisión de toda prueba relevante que no deba excluirse por aplicación de alguno de los filtros adicionales impuestos por las reglas jurídicas⁷. La diferencia con otros ámbitos de la experiencia es que en esos otros ámbitos es común que ese sea el único filtro. La particularidad jurídica es que se añaden un buen número de reglas jurídicas de exclusión: no se admiten las pruebas obtenidas en violación de derechos fundamentales, no se admiten los testigos de referencia, no se admiten, con salvedades, las pruebas que no puedan practicarse en el curso del proceso con la debida aplicación del principio de contradicción, etc. Además, algunos sujetos cuentan con lo que la doctrina anglosajona denomina «*privileges*»: esto es, se les exime de aportar información relevante para el caso (abogados —respecto de la información obtenida

⁵ A las que habrá que añadir, en su caso, los hechos admitidos por todas las partes y los denominados «hechos notorios», de los que suele decirse que no necesitan prueba.

⁶ Por ello, como he justificado con mayor detalle en FERRER (2006) no puede explicarse el resultado de la valoración de la prueba en el derecho mediante la noción de convicción judicial o de creencia. La creencia, por decirlo de un modo simple, es siempre *all things considered*, mientras que el resultado de la valoración de la prueba es siempre contextual: su justificación es relativa a ese conjunto de elementos de juicio.

⁷ Este es el punto de vista de THAYER (1898: 266 y ss.), del que puede verse una buena exposición en McNAMARA, 1986. Contra esta tesis de la concepción racionalista de la prueba, puede verse STEIN, A., 2005: 108 y ss. En mi opinión, como se verá, el análisis de STEIN sufre la confusión de no distinguir entre los distintos momentos de la actividad probatoria y, correspondientemente, entre los efectos de la regulación jurídica en cada uno de ellos.

de sus clientes—, sacerdotes—respecto de la información obtenida en confesión—, cualquier sujeto respecto de la información que pueda incriminarle, familiares directos—respecto de informaciones que puedan perjudicar o beneficiar a esos familiares—, etc.). Finalmente, pero sin menor importancia, y sin ánimo de exhaustividad, conviene señalar que también los propios plazos procesales juegan un papel de regla de exclusión: se excluye toda información—aunque sea relevante— que se presente mediante pruebas aportadas fuera de los plazos expresamente previstos.

Como ya señaló BENTHAM (1823: 391, 395), buena parte de esas reglas que excluyen elementos de juicio lógicamente relevantes se justifican en la protección por parte del derecho de valores distintos a la averiguación de la verdad: la intimidad y otros derechos fundamentales, la autonomía individual, las relaciones familiares, etc. En otros casos, en cambio, se excluyen elementos de juicio por considerarlos de bajo valor epistemológico ante el peligro de que una mala valoración tienda a darles más valor del que tienen. El número y el alcance de las reglas jurídicas que regulan la conformación del conjunto de elementos de juicio varía en función de los distintos ordenamientos y, dentro de ellos, en función de los tipos de proceso y jurisdicción.

2.2. La valoración de los elementos de juicio o pruebas

A partir de ahí, llegados al momento en que se cierra la composición del conjunto de elementos de juicio que deberán valorarse (con una simplificación quizá excesiva, pero elocuente: una vez declarado el proceso visto para sentencia), se pasa al segundo momento. Debe ahora procederse a la valoración de los elementos de juicio disponibles en el proceso. Si el sistema jurídico establece un régimen de libre valoración de la prueba, entonces deberá valorarse el apoyo que cada elemento de juicio aporta a las hipótesis en conflicto, de forma individual y en conjunto. Con ello, deberá obtenerse un resultado que nos permita saber el grado de confirmación del que dispone cada una de esas hipótesis.

Deben hacerse aquí dos observaciones importantes: en primer lugar, el resultado de la valoración de la prueba es siempre contextual, esto es, referido a un determinado conjunto de elementos de juicio (MENDONÇA, 1997: 77-78). Si cambia el conjunto, por adición o sustracción de algún elemento, el resultado puede perfectamente ser otro (FERRER, 2005: 104 y ss.). En segundo lugar, la libre valoración de la prueba es *libre* sólo en el sentido de que no está sujeta a normas jurídicas que predeterminen el resultado de esa valoración. La operación consistente en juzgar el apoyo empírico que un conjunto de elementos de juicio aportan a una hipótesis está sujeta a los criterios generales de la lógica y de la racionalidad (entre otros muchos, RESCHER-JOINT, 1959: 565; DAMAŠKA, 1986: 55, y 1997: 21; TWINING, 1994: 194; TARUFFO, 1990: 429).

Puede advertirse que ese apoyo empírico ofrecerá un grado de corroboración que, como ya señalará POPPER, nunca será concluyente, pero, como él mismo advierte también: «Aunque no podamos justificar una teoría [...] podemos, a veces, justificar nuestra preferencia por una teoría sobre otra; por ejemplo, si su grado de corroboración es mayor» (POPPER, 1974: 140).

Si el primero de los momentos, la conformación del conjunto de elementos de juicio, acoge la gran mayoría de especificidades jurídicas, este segundo momento (siempre que estemos en el ámbito de la libre valoración de la prueba) es el momento de la racionalidad (THAYER, 1898: 275). Se trata aquí, insisto, como en cualquier otro ámbito del conocimiento, de evaluar el apoyo empírico que un conjunto de elementos de juicio aportan a una determinada hipótesis o a su contraria. Si tengo, pues, razón en esto, negar la posibilidad de adoptar decisiones racionales en este momento, a pesar de las especificidades jurídicas que puedan regir en el primero y el tercero de los momentos de la prueba jurídica, supone necesariamente acoger argumentos escépticos generales ante la posibilidad del conocimiento o, más estrictamente, de la decisión racional. En otras palabras: no se puede negar la posibilidad de adoptar decisiones racionales en el ámbito de la prueba jurídica por razones específicamente jurídicas cuando no hay reglas jurídicas que limiten la libre valoración de la prueba.

Así pues, las limitaciones jurídicas que excluyen pruebas relevantes en el momento de la conformación del conjunto de elementos de juicio hacen que el conjunto sea epistemológicamente más rico o más pobre. Las limitaciones temporales, las limitaciones de recursos humanos y materiales, la falta de neutralidad de la intervención de las partes al aportar elementos de juicio, etc., producen, en general, que el conjunto de elementos de juicio disponible sea más pobre. Pero una vez ese conjunto está ya delimitado, la operación de valorar lo que de él se puede racionalmente inferir no difiere en nada de la que se puede realizar en cualquier otro ámbito de la experiencia y está sometida a los controles de la racionalidad general.

2.3. La adopción de la decisión sobre los hechos probados

Por último, el tercer momento es el que corresponde a la toma de la decisión. La valoración de la prueba habrá permitido otorgar a cada una de las hipótesis en conflicto un determinado grado de confirmación que nunca será igual a la certeza absoluta. Habrá que decidir ahora si la hipótesis *h* puede o no declararse probada con el grado de confirmación de que dispone. Esto depende del estándar de prueba que se utilice. Así, por ejemplo, es muy usual sostener (especialmente en la cultura jurídica anglosajona) que en el ámbito civil opera el estándar de la prueba prevalectante, de modo que una hipótesis está probada si su grado de confirmación es superior al de la hipótesis contraria.

En cambio, en el ámbito penal, operaría el estándar que exige que la hipótesis esté confirmada más allá de toda duda razonable. Es claro que aquí, de nuevo, la elección de uno u otro estándar es propiamente jurídica y se realiza en atención a los valores en juego en cada tipo de proceso (STEIN, A., 2005: 121-122). Así, puede justificarse la mayor exigencia probatoria en los casos penales, por ejemplo, en una especial protección del derecho a la libertad (vehiculada a través del derecho a la presunción de inocencia).

Conviene insistir en que el resultado de la valoración de la prueba que se obtenga en el segundo momento no implica por sí solo nada respecto de la decisión a adoptar. Para ello, es necesaria la intermediación de algún estándar de prueba. Y ni siquiera puede darse por descontado que la hipótesis que haya resultado más confirmada es aquella que deberá darse por probada. Si se observa bien el funcionamiento del estándar que exige una confirmación de la hipótesis de la culpabilidad, en un proceso penal, más allá de toda duda razonable, éste supone que la hipótesis no se considerará probada aunque disponga de un apoyo empírico mayor que la hipótesis de la inocencia (salvo que ese apoyo ofrezca una corroboración muy alta a la primera), de forma que se presumirá la verdad de la hipótesis menos confirmada (i. e., la de la inocencia).

3. LA CRÍTICA DE STEIN A LA TRADICIÓN RACIONALISTA

En el apartado anterior he señalado la vigencia, con algunas excepciones, del principio general de que cualquier elemento de juicio relevante para la adopción de una decisión debe ser admitido como prueba en el proceso judicial. Este principio general se justifica epistemológicamente en la medida en que garantiza la mayor probabilidad de que los enunciados que se declaren probados coincidan con la verdad.

Ahora bien, la averiguación de la verdad es un fin en algún sentido prioritario del proceso en materia de prueba⁸, pero no es en absoluto el único. La celeridad en la toma de decisiones, la protección de derechos fundamentales, la protección de secretos de estado, el secreto de las relaciones abogado-cliente, etc., son también fines habitualmente reconocidos en la gran mayoría de ordenamientos jurídicos (LIEBMAN, 1955: 276-277; RESCHER y JOYNT, 1959: 568; WEINSTEIN, J. B., 1966: 241 y ss.; DAMAŠKA, 1986: 160 y ss., y 1997: 12 y ss.; POSNER, 1990: 206; y TARUFFO, 1992: 336-337; STEIN, A., 2005: 110; entre otros muchos). Por ello, para garantizar la obtención de esas otras finalidades, los ordenamientos establecen reglas jurídicas procesales (aunque no sólo procesales) que funcionan como un filtro de admisibilidad de la prueba,

⁸ Se trata seguramente del único fin funcionalmente necesario para que sea posible la aplicación del derecho y, por ello, para que el derecho como instrumento de control social pueda también funcionar.

adicional y posterior al juicio de relevancia. Resulta claro que la racionalidad instrumental de esas reglas jurídicas no puede ser evaluada teniendo en mente la finalidad de la averiguación de la verdad, sino la finalidad a la que en cada caso responden. Pero ello significa también que las reglas que persiguen la garantía de esos otros fines del proceso o del derecho pueden imponer algún sacrificio epistemológico, en el sentido de que pueden ser contraproducentes para el objetivo de la averiguación de la verdad. En general es así, en efecto, cada vez que la protección o maximización de esos otros fines del proceso supone la exclusión de pruebas relevantes para el caso que debe decidirse. Esto no implica que sea imposible alcanzar una determinación verdadera de los hechos ocurridos, sino, más modestamente, que las probabilidades de que ello ocurra, dado que el conjunto de elementos de juicio será más pobre, serán también más bajas. Y siendo consciente de ese coste, el legislador deberá decidir en cada caso de conflicto entre la maximización de los fines epistemológicos y de cualquier otro que se quiera proteger, en qué medida se sacrifican unos y otros (WRÓBLEWSKI, 1981: 183).

Para evaluar en cada caso si este tipo de reglas de exclusión están justificadas, habrá que juzgar su racionalidad teleológica atendiendo a su adecuación como medios para alcanzar los fines a los que responden. Y, además, evaluar, dado que entran en conflicto con la finalidad de la averiguación de la verdad, si están disponibles otros medios para alcanzar esos fines que no conlleven este conflicto (evitando así la regla de exclusión).

Ahora bien, resulta evidente que cualquier regla de exclusión de este tipo supone el rechazo de una prueba que ha pasado el filtro de la relevancia y que, por tanto, tiene a su favor el principio general de inclusión de toda prueba relevante. Es decir, esas reglas de exclusión suponen que existe un conflicto entre distintos fines considerados valiosos por el derecho. Así, por ejemplo, la averiguación de la verdad *versus* la inviolabilidad de la correspondencia. La justificación de la regla de exclusión no es, en esos casos, sólo una cuestión de racionalidad teleológica. Tiene también un aspecto valorativo de gran importancia. Se trata de la ordenación de los valores que permite decidir por uno o por otro en caso de conflicto. Y ésta ya no es una cuestión de racionalidad sino valorativa. El legislador deberá adoptar una decisión respecto de la ordenación de los valores en conflicto y esa decisión sólo podrá ser juzgada desde patrones valorativos.

Alex STEIN ha insistido mucho (2005: 118 y ss.) en que no se da aquí un conflicto entre la epistemología (entendida como actividad neutral desde el punto de vista valorativo) y ciertos valores sociales o morales que el derecho pretende proteger, sino, en último término, como he señalado, un conflicto entre distintos valores, que el legislador deberá ordenar si quiere resolver el conflicto. Esto es, la averiguación de la verdad es también un valor para el derecho. A partir de esta constatación, STEIN elabora un doble argumento contra el que denomina «movimiento abolicionista» del derecho probatorio, que,

liderado entre otros por BENTHAM y THAYER, sostiene que la valoración de la prueba es una cuestión puramente epistemológica respecto de la que el derecho haría bien en no inmiscuirse:

1) En primer lugar, la epistemología sólo nos sirve para determinar el grado de probabilidad de que una hipótesis sobre los hechos sea verdadera, pero nada nos dice sobre el punto en que esa probabilidad es suficiente para aceptar como verdadera la hipótesis. Es decir, la epistemología no puede determinar los estándares de prueba (STEIN, A., 2005: 121-122). Establecer el umbral a partir del cual el nivel de corroboración de una hipótesis es suficiente supone tomar una decisión sobre el reparto del riesgo de error. Así, en efecto, si el estándar de prueba en el proceso penal es más exigente se producirán menos condenas falsas y más absoluciones falsas, mientras que el efecto será exactamente el contrario si el estándar se sitúa en un punto menor de exigencia. Pues bien, esa es una decisión que queda absolutamente en el ámbito político-moral sobre el que la epistemología no tiene nada que decir. Para ello es necesario, pues, que se pronuncie el derecho.

2) En segundo lugar, como hemos visto, el denominado principio de inclusión impone la admisión de toda prueba relevante para la decisión sobre los hechos que debe tomarse en el proceso. O, dicho de otro modo, el principio implicaría, tomado estrictamente, la derogación de las reglas probatorias que excluyen pruebas relevantes. Esta es la clásica formulación de BENTHAM (1827, vol. 5: 255) y tiene su fundamento en que la probabilidad de alcanzar una decisión correcta respecto de la verdad sobre los hechos aumenta en la medida en que lo hace la información sobre lo ocurrido. La riqueza o peso del conjunto de elementos de juicio sobre el que se adopte la decisión estaría en directa relación con la probabilidad de que la decisión sea correcta. La objeción de STEIN, en cambio, ataca precisamente el fundamento del principio de inclusión:

Mientras la información de que disponga el decisor sobre los hechos siga siendo incompleta, la adquisición de información adicional con credenciales inciertas puede no mejorar su posición epistémica. La llegada de nueva información podría sólo sustituir el riesgo de error existente hasta el momento por un nuevo riesgo de error, vinculado a la credibilidad de la nueva información. Más aún, no hay garantía de que el nuevo riesgo de error sea menor que el anterior.

[...] mi tesis simplemente refuta la intuitiva (pero falaz) idea de que el aumento de información *necesariamente* produce mayor exactitud en la determinación de los hechos (STEIN, A., 2005: 123, cursiva original).

Pues bien, creo que el desafío de STEIN al movimiento abolicionista⁹ debe ser cuidadosamente analizado. Tiene razón el autor cuando plantea que la ave-

⁹ Como se habrá advertido, ese movimiento teórico coincide fundamentalmente con las tesis de la concepción racionalista de la prueba, mencionada en el primer epígrafe, y que otros autores han denominado movimiento *Free proof*.

riguación de la verdad es un valor más que el derecho persigue y que puede entrar en conflicto con otros valores asumidos por el mismo derecho¹⁰. En esos casos, evidentemente, estaremos ante un conflicto de valores que habrá que resolver. Y la decisión que se adopte no será en ningún caso informada por la epistemología sino por la política y la moral. Sin embargo, conviene tener en cuenta que, en mi opinión, el valor o la finalidad de la averiguación de la verdad no tiene estructuralmente la misma posición que los demás. Como ya he defendido en otro lugar (FERRER, 2007: 29-32), el objetivo institucional de la prueba en el proceso es la averiguación de la verdad. Y ello no puede ser de otra manera, puesto que ese objetivo es estructuralmente necesario para que funcione el propio derecho como mecanismo de motivación de la conducta¹¹. Sólo si las consecuencias jurídicas previstas por el derecho para acciones determinadas se aplican efectivamente a esas acciones (idealmente siempre a esas acciones y nunca a otras), los ciudadanos tendrán motivos (jurídicos) para actuar conforme a lo prescrito por el derecho y éste podrá cumplir su función de mecanismo de resolución de conflictos. Esta faceta estructural no es compartida por otros valores con los que la averiguación de la verdad puede entrar en conflicto. Un sistema jurídico puede funcionar perfectamente sin asumir el valor de la inviolabilidad del domicilio o de las comunicaciones, por ejemplo. Sería este un sistema jurídico indeseable por otras razones, pero no hay nada que impida que un sistema así sea perfectamente efectivo y eficiente en el cumplimiento de su función motivadora de la conducta. En cambio, el sistema no puede prescindir de la averiguación de la verdad como objetivo institucional del proceso (y, por tanto, de la aplicación del derecho), puesto que de otro modo el sistema colapsaría. Esto no le otorga un mayor valor moral a la averiguación de la verdad, pero muestra que en el conflicto entre ese objetivo del derecho y otros posibles objetivos, el primero tiene una preferencia estructural que hace que no pueda ceder siempre.

Dicho esto, creo que el primer argumento de STEIN es perfectamente correcto. Efectivamente, la decisión sobre el nivel de suficiencia probatoria no es en absoluto epistemológica. La epistemología nos puede ayudar a delinear un estándar de prueba que refleje correctamente el nivel de suficiencia probatoria que se haya decidido adoptar, pero no nos dice nada sobre el nivel mismo. Esa

¹⁰ Punto este que no niega en absoluto la tradición racionalista (que coincide básicamente con lo que STEIN denomina «movimiento abolicionista»). Véase ANDERSON, SCHUM y TWINING, 1991: 79-80.

¹¹ No pretendo entrar aquí en el debate de las funciones del derecho, que es muy complejo e innecesario para el argumento que estoy sosteniendo. Basta, por el momento, con aceptar que cualquiera que sea la función que se pretenda atribuir al derecho, el cumplimiento de esa función exige que el derecho tenga un alto grado de eficacia como mecanismo de guía de la conducta (HART, 1994: 248-249). Y para que esa eficacia sea posible, pretendo mostrar que es necesario que en la aplicación del derecho se imponga la consecuencia jurídica prevista por las normas generales a los casos en que se haya dado el supuesto de hecho previsto también por las normas (y no se imponga a los casos en que el supuesto de hecho no se dé). Si esto es así, se obtiene ya que el proceso judicial, por lo que atañe a la prueba, no puede tener otro objetivo fundamental que la averiguación de la verdad.

es una decisión política. Esto, de todos modos, no es específico del derecho. También la farmacología, la historia, la epidemiología o la astronomía necesitan de estándares de prueba claramente definidos. Y tampoco la epistemología sirve en esos casos de decisiones sociales (*i. e.*, de la comunidad científica) que toman en cuenta también los bienes en juego. No es extraño, pues, que el estándar de prueba farmacológico sea mucho más elevado que el de la historia. Y ello no por razones epistemológicas, sino por el coste del riesgo de error sobre los bienes en juego¹². La única diferencia en el caso del derecho es que tenemos mecanismos institucionales para adoptar las decisiones sobre los estándares de prueba.

En cambio, el segundo argumento de STEN, que afecta más centralmente al objeto de discusión de este epígrafe, me parece erróneo. Desde luego, el aumento de información relevante para la adopción de una decisión no necesariamente produce un aumento en la exactitud de la decisión. En ello tiene razón STEN. Pero no es esto lo que necesita el movimiento abolicionista o la tradición racionalista para sostener que tendencialmente toda prueba relevante debe ser admitida (y, por tanto, deben abolirse las reglas de exclusión correspondientes). El fundamento necesario para esa tesis es algo más débil que el rechazado por STEN: que el aumento de información relevante aumenta, *ceteris paribus*, la probabilidad de que se adopte una decisión adecuada materialmente, esto es, que se declaren probados enunciados verdaderos sobre los hechos. Afirmar únicamente que una decisión es adecuada o correcta es ambiguo, por cuanto podría significar que es la decisión fundamentada en el conjunto de elementos de juicio disponible (con independencia de su valor de verdad) o que es una decisión cuyo contenido se corresponde con la verdad de lo ocurrido (FERRER, 2005: 98-100). Se puede distinguir así entre la validez y la verdad de la decisión (LAUDAN, 2006: 12-17). Una decisión sobre los hechos será válida si está fundada en los elementos de juicio disponibles. Será, en cambio, verdadera si se corresponde con los hechos del mundo. Pues bien, el aumento de la riqueza del conjunto de elementos de juicio (o de la información, en palabras de STEN) aumenta la probabilidad de que la decisión válida sea también verdadera. Y ello es fundamento suficiente para justificar el principio de inclusión¹³.

¹² No somos indiferentes ante los falsos positivos o falsos negativos en investigación farmacológica, de modo que preferimos un falso negativo (un medicamento que no supere el estándar y que, en cambio, sea beneficioso para la salud) a un falso positivo (un medicamento que supere el estándar y que, en cambio, sea perjudicial para la salud). Por ello, se define un estándar de prueba muy exigente que distribuye desigualmente los riesgos de error. En cambio, si somos indiferentes ante la distribución de errores entre los falsos positivos y negativos en investigación histórica, por lo que el estándar de prueba se sitúa en la proporcionalidad de la prueba (es decir, en la exigencia de que la probabilidad de una hipótesis sea mayor que la de su negación).

¹³ La crítica de STEN tiene un argumento más que debe ser analizado. Sostiene el autor que el aumento de información puede ser contraproducente para la corrección de la decisión si hay dudas sobre

4. RECONSIDERANDO LAS TESIS DE PARTIDA

Una vez analizados los distintos momentos de la actividad probatoria en el proceso judicial, estamos en condiciones de revisar las tesis de BENTHAM presentadas al inicio del trabajo y evaluar la incidencia, real y/o deseable, que tienen en nuestros sistemas procesales.

En este sentido, TWING (1994: 40-41) sostiene que las tesis abolicionistas del derecho probatorio propugnadas por BENTHAM han tenido éxito (en el mundo anglosajón) en lo que hace a la inexistencia de reglas sobre la valoración de la prueba y, en cambio, no lo han tenido tanto en lo que hace a las reglas de exclusión de pruebas relevantes. Desde un punto de vista estrictamente epistemológico, creo que puede decirse que, de ser así, el resultado ha sido el inverso al deseable.

En efecto, en los sistemas jurídicos anglosajones el derecho probatorio es, fundamentalmente, un conjunto a veces muy prolífico y casuístico de reglas sobre la admisibilidad y la exclusión de la prueba. En cambio, prácticamente nada se dice sobre la valoración de la prueba y las reglas que regulan el momento de la adopción de la decisión sobre los hechos probados son extraordinariamente vagas, al punto de resultar poco menos que inservibles.

En el momento de la valoración de la prueba, el juez, jurado o tribunal deberá evaluar qué hipótesis de las diversas en conflicto en el proceso, tiene un más alto grado de contrasación. Esta es una labor comparativa entre las distintas hipótesis, a la luz de los test realizados y presentados como pruebas por las partes, peritos, etc., u ordenadas por el propio juez si el sistema se lo permite. Pero el resultado de esta valoración no podrá ser más que una ordenación del nivel de contrasación de las distintas hipótesis. Para llegar a la conclusión que atribuye a una de ellas la categoría de hecho probado es necesario contar con un criterio que nos indique a partir de qué nivel de contrasación vamos a considerar probada la hipótesis. Necesitamos un estándar de prueba¹⁴.

Como he señalado anteriormente, esto no es distinto de lo que ocurre en cualquier otro ámbito de la experiencia. En paleontología, medicina, historia, farmacología, química orgánica o física nuclear son necesarios también estándares de prueba que permitan decidir cuándo una hipótesis en esos ámbitos de

las credenciales de la información, esto es, sobre su fiabilidad. Desde luego, ello puede suceder. En todo caso, en línea de principio, la respuesta es que en ese supuesto lo que corresponde es un nuevo aumento de la información que nos permita evaluar la fiabilidad de la nueva prueba adquirida. Desde luego, en algún lugar hay que poner límites a este continuo para evitar el regreso al infinito.

¹⁴ Esto es así incluso en el caso en que se adopten las tesis de la inferencia a la mejor explicación, puesto que en este caso, sencillamente, el estándar de prueba indicaría que se acepte como probada la hipótesis relativamente mejor contrastada (aunque su nivel de contrasación sea bajo), por el hecho de que es la mejor de que disponemos.

investigación puede ser considerada probada. En todos ellos, son las respectivas comunidades científicas las que, de manera normalmente informal y no institucionalizada, adoptan un estándar de prueba atendiendo a la ratio entre errores positivos y negativos que se consideran aceptables en esos ámbitos. Y para ello la epistemología no impone más que metodologías de formulación del estándar, no el estándar mismo. No es raro, pues, que eso sea necesario también en el ámbito jurídico (como bien señala STEIN, 2005: 133-134). La especificidad, como he indicado ya con antelación, es el carácter institucionalizado de la toma de decisiones jurídicas. Y, por razones democráticas, tampoco es raro que esa decisión corresponda adoptarla a los legisladores (para cada tipo de proceso jurídico).

Si esto es así, pues, conviene matizar las exigencias de las tesis abolicionistas: el derecho no puede dejar de regular todos los aspectos de la prueba en el proceso para que rijan sólo criterios epistemológicos por la sencilla razón de que la epistemología no es suficiente para adoptar decisiones (ni pretenderlo). En cambio, si sería la epistemología una guía suficiente para el momento de la conformación del conjunto de elementos de juicio (si el valor de la búsqueda de la verdad fuera el único que nos guiara o prevaleciera siempre en caso de conflicto con otros, cosa que no ocurre en ningún ordenamiento). Y también, especialmente, en el momento de la valoración de la prueba, si estamos bajo el régimen de la libre valoración de la prueba.

La situación en el ámbito de los países de *civil law* es bastante distinta. DAMAŠKA (1997: 41-46) sostiene la tesis, muy atendible, de que la falta de una exigencia jurídica de motivación del juicio sobre los hechos en los ordenamientos de *common law* explica la especial proliferación en ese ámbito de reglas jurídicas que excluyen pruebas que se considera que aportan información poco fiable. El objetivo es evitar que en la valoración de las pruebas se les otorgue un peso excesivo¹⁵. En algún sentido, estas reglas pretenden garantizar *ex ante* una mayor racionalidad general de las decisiones sobre los hechos a costa de excluir elementos de juicio que, aun con un valor relativamente escaso, pudieran aportar informaciones relevantes. En los ordenamientos de *civil law*, en cambio, el control de la racionalidad de la decisión se realizaría *ex post* mediante el control de la motivación (IGARRUA, 2003: 135). Así, por ejemplo, ante una prueba o un tipo de prueba de muy baja fiabilidad, el control de racionalidad podrá funcionar a posteriori siempre que se exija al decisor sobre los hechos probados que justifique por qué ha declarado probados esos hechos y cuál ha sido el apoyo empírico en el que ha basado su decisión. Podrá

¹⁵ Pero no es éste el único motivo que explica la tradicional proliferación de reglas de exclusión de pruebas en los sistemas de *common law*. Hay que añadir también, por ejemplo, la presencia generalizada de la institución del jurado compuesto exclusivamente por legos (DAMAŠKA, 1997: 53-54) como órgano decisor sobre los hechos del caso y al que se pretende instruir detalladamente sobre las reglas que deben guiar su función.

comprobarse, entonces, si el decisor ha otorgado un peso excesivo a una prueba cuya fiabilidad resulte cuestionada¹⁶. En cambio, si no se exige motivación de la decisión sobre los hechos, el único modo de asegurar normativamente que no se tome la decisión sobre la base de pruebas poco o nada fiables es, directamente, excluirlas del conjunto de elementos de juicio disponibles¹⁷.

Pero el problema en los sistemas de *civil law* respecto de los estándares de prueba es el mismo: su práctica inexistencia. Tomemos como ejemplo el proceso penal: la tradicional apelación a la íntima convicción o la progresiva incorporación (por vía legislativa o jurisprudencial) del «más allá de toda duda razonable» propio de los sistemas anglosajones (McCORMICK, 1954, vol. 2: 445-449) a los sistemas continentales¹⁸ no satisface las necesidades en absoluto, puesto que deja la decisión en manos de la subjetividad del decisor y no permite ningún control intersubjetivo. Como ha mostrado muy convincentemente LAUDAN (2003 y 2006: 29-62), si no disponemos de un criterio que nos indique las condiciones en que la duda será razonable (que no podrá ser cuantitativo), no podremos decir en ningún caso que disponemos de un estándar de decisión objetivo (o al menos intersubjetivo) y cognoscible a priori. Y esto producirá serios problemas para la finalidad motivadora de la conducta propia del derecho: si no disponemos de un estándar de prueba cognoscible a priori, las decisiones judiciales no podrán ser previsibles, lo que impide la adaptación estratégica de la conducta a la decisión prevista¹⁹.

En otro lugar he propuesto una formulación del estándar de prueba penal (para la decisión final sobre los hechos probados de la hipótesis acusatoria)

¹⁶ Esto, por supuesto, está también vinculado a la existencia de un sistema de recursos para la revisión de la decisión adoptada (que sea, además, simétrico, *i. e.*, igual para todas las partes).

¹⁷ A pesar de ello, en ausencia de motivación detallada no hay forma de comprobar las razones que se han usado para adoptar la decisión. Podrá solo controlarse si, a la luz de los elementos de juicio disponibles en el proceso, esa decisión resulta suficientemente justificada o no, realizando para ello una nueva valoración de los mismos.

¹⁸ El uso de esta fórmula se ha expandido en los últimos años también en muchos países de tradición romano-germánica. Es el caso, por ejemplo, de Italia y España. Véanse, al respecto, STELLA (2001: 195-218) para el caso italiano y FERNÁNDEZ LÓPEZ (2005: 199-208) para el caso español. También el Tribunal Europeo de Derechos Humanos ha manifestado que el derecho a la presunción de inocencia establecido en el art. 6.2 del CEDH (Convenio Europeo para la Protección de los Derechos Humanos y Libertades Fundamentales) implica que sobre la acusación pesa la carga de la prueba más allá de toda duda razonable (STRDH de 6 de diciembre de 1988; caso *Barberá, Mesegué y Jabardo v. España*). En mi opinión, no obstante, la presunción de inocencia como regla de juicio establece únicamente que en caso de duda se presume la inocencia del acusado, pero no implica ningún estándar de prueba específico. O, dicho de otro modo, el derecho a la presunción de inocencia es compatible con estándares de prueba distintos, que expresan diversos niveles de exigencia probatoria.

¹⁹ Por esta razón, no puedo estar de acuerdo con autores como STOFFELMAYER y SEIDMAN DIAMOND (2000: 781) y LILLQUIST (2002: 146 y ss., 162 y ss.), que consideran justificable o incluso adecuado disponer de un estándar de prueba «flexible» para que pueda ser adaptado por el juzgador a las circunstancias del caso concreto e incluso del acusado. Por ello, sostienen estos autores, puede ser conveniente una formulación extremadamente vaga y no definida del estándar de prueba (como la de «más allá de toda duda razonable»), que permita esa flexibilidad. En mi opinión, ello conlleva necesariamente la imposibilidad del control sobre la aplicación del estándar de prueba y, por consecuencia, del cumplimiento del derecho a la presunción de inocencia.

que pretende salvar los inconvenientes señalados. No abordaré aquí su justificación, ni siquiera su explicación detallada, pero sí lo formularé a los efectos de facilitar la comprensión sobre las exigencias mencionadas.

Así, según este estándar de prueba, para considerar probada la hipótesis de la culpabilidad deberán darse conjuntamente las siguientes condiciones:

- 1) La hipótesis debe ser capaz de explicar los datos disponibles, intergrandolos de forma coherente, y las predicciones de nuevos datos que la hipótesis permita formular deben haber resultado confirmadas.
- 2) Deben haberse refutado²⁰ todas las demás hipótesis plausibles explicativas de los mismos datos que sean compatibles con la inocencia del acusado, excluidas las meras hipótesis ad hoc.²¹

La falta de estándares de prueba claros y objetivos no puede atribuirse, claro está, a los supuestos efectos perniciosos de las tesis de BENTHAM y de la tradición racionalista o abolicionista. Pero esa tradición, y el importantísimo influjo de BENTHAM en particular, han contribuido en gran medida a que no

²⁰ Conviene aclarar el sentido en que uso el término «refutación» en este trabajo para evitar una problemática confusión: entiendo la refutación como el rechazo de una hipótesis derivado del incumplimiento de las predicciones que pueden formularse a partir de ellas. Hay que advertir, no obstante, que el incumplimiento de una predicción formulable a partir de una hipótesis no conlleva necesariamente la refutación de la hipótesis. Al respecto puede verse FERRER (2007: 132) y LITTEK (1997: 807-808). Por ello, contra lo que pudiera parecer (y sostuvo Popper), no hay asimetría entre la corroboración de una hipótesis y su refutación: ambos son procesos graduales y probabilísticos.

²¹ Este es el único requisito planteado por ALLEN (1994: 604), aunque él no hace la advertencia de la exclusión de las hipótesis ad hoc. Siendo así, se presenta el problema de que siempre es posible construir una historia coherente que dé cuenta a posteriori de los datos disponibles. El acusado podrá sostener, por ejemplo, la hipótesis del cuento en su contra, que quizás resulte irrefutable. Pero es una hipótesis posible. A ALLEN le queda como salida hacer más fuerte el requisito de la plausibilidad, de modo que no excluya únicamente las hipótesis incoherentes o incompatibles con nuestro conocimiento del mundo, pero entonces este se convierte en un requisito tan extremadamente vago que quizás sea inservible. Para evitar este problema, ALLEN y PARDO (2007: 135-136) han propuesto recurrir a la obra de HAYMAN (1965) y ha tenido en los últimos años algunos defensores de su aplicación al razonamiento probatorio jurídico (por todos, AMAYA, 2009). Desde luego, el modelo requeriría de un análisis detallado que no puedo realizar aquí. Sin embargo, conviene señalar dos problemas que a mi entender lo hacen difícilmente utilizable como criterio de decisión para el proceso penal (i. e., como modo de formular su estándar de prueba): 1) en muchas ocasiones, la mejor explicación disponible de los hechos del caso es, sin embargo, una mala explicación. Que sea la mejor explicación depende sólo de una comparación con las otras explicaciones o hipótesis consideradas. Y, de ese modo, por ejemplo, puede suceder que la hipótesis de la acusación sea una mejor explicación de los hechos que la hipótesis de la defensa, pero que a pesar de ello no sea una buena explicación de los mismos. De este modo, usar el modelo de la inferencia a la mejor explicación como forma de fijar el estándar de prueba penal hasta de este un estándar muy débil. 2) El segundo problema se presenta cuando las hipótesis explicativas en conflicto, tanto de la acusación como de la defensa, son buenas. En ese caso, aunque la hipótesis de la acusación sea mejor, no parece adecuado concluir con la aceptación de la hipótesis acusatoria, salvo que de nuevo se debilita mucho el estándar. Sobre estos y otros problemas de la teoría de la inferencia a la mejor explicación aplicada al razonamiento probatorio, véase LAUDAN, 2007.

se advirtiera tan importante carencia durante los dos últimos siglos²². Por el contrario, en los últimos años se está produciendo una reivindicación de las tesis benthamianas, especialmente en países de *civil law*, por lo que añade al momento de la conformación del conjunto de elementos de juicio sobre cuya base se deberá tomar la decisión, así como la progresiva derogación de muchas reglas de prueba tasada, que inciden en el segundo de los momentos de la actividad probatoria (el de su valoración). Y esta sí parece una influencia fructífera.

BIBLIOGRAFÍA

- ALLEN, R. J., 1994: «Factual Ambiguity and a Theory of Evidence», en *Northwestern University Law Review*, 88.
- ALLEN, R. J., y PARDO, M. S., 2007: «The Problematic Value of Mathematical Models of Evidence», en *Journal of Legal Studies*, 36.
- AMAYA, A., 2009: «In reference to the Best Legal Explanation», en H. KAPTEIN, H. PRAKKEN y B. VERHEI (eds.), *Legal Evidence and Proof: Statistics Stories, Logic, Aldershot: Ashgate*, en prensa.
- ANDERSON, T.; SCHUM, D., y TWANING, W., 1991: *Analysis of Evidence*, 2.ª ed., New York: Cambridge University Press, 2005.
- BAYÓN, J. C., 2009: «Epistemología, moral y prueba de los hechos: hacia un enfoque no benthamiano», en *Análisis e derecho*, 2009 (en prensa). Citado por el manuscrito presentado en el XIV Congreso Italo-Español de Teoría del Derecho, Girona, 2008.
- BENTHAM, J., 1823: *Traité des preuves judiciaires*, translated by Dumont, Paris: Bosses frères. Citado por la traducción de M. OSSORO: *Traité de las pruebas judiciales*, Granada: Comares, 2001.
- 1827: *Rationale of Judicial Evidence*, 7 vols., editado por J. STUART MILL. Citado por la edición incluida en BENTHAM, 1838-1843, vols. VI y VII.
- 1838-1843: *The Works of Jeremy Bentham*, 11 vols., ed. por J. BOWRING, Bristol: Thoemmes Press, 1995.

²² BAYÓN (2009: 18) argumenta convincentemente que la preocupación de BENTHAM estaba centrada en la minimización de errores y no en su distribución. Si esto es así, según BAYÓN, bastarían reglas epistemológicas para tomar la decisión (que debería coincidir con la hipótesis mejor fundada —lo que para BAYÓN nos lleva a la inferencia a la mejor explicación— y no serían necesarias múltiples reglas socio-políticas que establezcan estándares de prueba). Otra cosa sucedería si estuvieramos interesados no sólo (o no principalmente) en la minimización de errores sino en su distribución. Por supuesto, para distribuir los riesgos de error entre las partes en conflicto es necesario introducir una regla que no podrá ser nunca epistemológica. La cuestión es que, en el derecho, no distribuir los riesgos de error de manera equitativa entre las partes es también tomar una decisión sobre la distribución de riesgos de error. Se sostiene habitualmente que así es, por ejemplo, en el proceso civil, donde imperaría la regla de la preponderancia de la prueba (que distribuye igualmente los riesgos de error entre las partes). Y por ello se considerara también en el ámbito civil que esa regla constituye un estándar de prueba.

La inferencia a la mejor explicación no opera tampoco, en mi opinión, en el ámbito de la jurisdicción civil puesto que, también allí, la mejor explicación puede no superar el estándar de la preponderancia de la prueba (si no se acepta la aplicación del principio del cálculo de probabilidades de la complementariedad). En todo caso, no me resulta posible aquí desarrollar y justificar con detalle estas tesis.

- CARNELUCCI, F., 1947: *La prova civile*, 2.ª ed., Roma: Edizioni dell'Ateneo.
- DAMAŠKA, M., 1986: *The Faces of Justice and State Authority: A Comparative Approach to the Legal Process*, New Haven-London: Yale University Press.
- 1997: *Evidence Law Adrift*, New Haven-London: Yale University Press.
- FERNÁNDEZ LÓPEZ, M., 2005: *Prueba y presunción de inocencia*, Madrid: Iustel.
- FERRER BELTRÁN, J., 2005: *Prueba y verdad en el derecho*, 2.ª ed., Madrid: Marcial Pons.
- 2006: «Legal Proof and Fact Finders' Beliefs», en *Legal Theory*, 12.
- 2007: *La valoración racional de la prueba*, Madrid-Barcelona-Buenos Aires: Marcial Pons.
- GASCÓN, M., 1999: *Los hechos en el derecho. Bases argumentales de la prueba*, Madrid: Marcial Pons.
- HARMAN, G. (1965): «The inference to the best explanation», en *Philosophical Review*, 74.
- HART, H. L. A., 1982: *Essays on Bentham*, Oxford: Oxford University Press.
- 1994: «Postscript», en H. L. A. HART, *The Concept of Law*, 2.ª ed., Oxford: Oxford University Press.
- IGARTUA, J., 2003: *La motivación de las sentencias, imperativo constitucional*, Madrid: Centro de Estudios Políticos y Constitucionales.
- KEETON, G. W., y MARSHALL, O. K., 1948: «Bentham's Influence on the Law of Evidence», en G. W. KEETON y G. SCHWARZENBERGER (eds.), *Jeremy Bentham and the Law*, London: Stevens.
- LAUDAN, L., 2003: «Is Reasonable Doubt Reasonable?», en *Legal Theory*, 9.
- 2006: *Truth, Error and Criminal Law. An Essay in Legal Epistemology*, Cambridge: Cambridge University Press.
- 2007: «Strange bedfellows: inference to the best explanation and the criminal standard of proof», en *International Journal of Evidence and Proof*, 11.
- LEITER, B., 1997: «The Epistemology of Admissibility: Why Even Good Philosophy of Science Would Not Make for Good Philosophy of Evidence», en *Brighton Young University Law Review*, núm. 4.
- LIEBMAN, E. T., 1955: *Manuale di diritto processuale civile*, Milano: Giuffrè. Citado por la traducción castellana de S. SENTÍS MELENDO, *Manual de derecho procesal civil*, Buenos Aires: EJEA, 1980.
- LULLQUIST, E., 2002: «Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability», en *U. C. Davis Law Review*, 36, núm. 1.
- MCCORMACK, Ch. T., 1954: *On Evidence*, 4.ª ed. por J. W. STRONG (ed.), St. Paul (Minn.): West Publishing Co., 1992.
- MCMAMARA, Ph., 1986: «The Canons of Evidence. Rules of Exclusion or Rules of Use?», en *Adelaid Law Review*, 10.
- MENDONÇA, D., 1997: *Interpretación y aplicación del derecho*, Almería: Servicio de Publicaciones de la Universidad de Almería.
- MORESO, J. J., 1992: *La teoría del derecho de Bentham*, Barcelona: PPU.
- PARDO, M., 2005: «The field of Evidence and the Field of Knowledge», en *Law and Philosophy*, núm. 24, vol. 5.
- POPPER, K., 1974: *Unended Quest. An Intellectual Autobiography*, The Library of Living Philosophers. Citado por la traducción castellana de C. GARCÍA TREVIANO, *Búsqueda sin término. Una autobiografía intelectual*, Madrid: Tecnos, 2002.

- POSNER, R. A., 1990: *The Problems of Jurisprudence*, Cambridge (Mass.): Harvard University Press.
- POSTEMA, G. J., 1986: *Bentham and the Common Law Tradition*, Oxford: Oxford University Press.
- RESCHER, N., y JOYNT, C. B., 1959: «Evidence in History and in the Law», en *The Journal of Philosophy*, vol. LVI, 13.
- SENTÍS MELENDO, S., 1979: *La prueba*, Buenos Aires: EJEA.
- STEIN, A., 2005: *Foundations of Evidence Law*, Oxford: Oxford University Press.
- STEIN, F., 1893: *Das private Wissen des Richters. Untersuchungen zum Beweisrecht beider Prozesse*. Citado por la traducción castellana de A. DE LA OLIVA, *El conocimiento privado del juez*, Madrid: Centro de Estudios Ramón Areces, 1990.
- STELLA, F., 2001: *Giustizia e modernità. La protezione dell'innocente e la tutela delle vittime*, 3.ª ed., Milano: Giuffrè, 2003.
- STOFFELMAYR, E., y SIEDMAN DIAMOND, S., 2000: «The Conflict between Precision and Flexibility in Explaining "Beyond a Reasonable Doubt"», en *Psychology, Public Policy and Law*, 6.
- TARUFFO, M., 1990, «Modelli di prova e di procedimento probatorio», en *Rivista di diritto processuale*, XLV, núm. 2.
- 1992: *La prova dei fatti giuridici*, Milano: Giuffrè.
- THAYER, J. B., 1898: *A Preliminary Treatise on Evidence at the Common Law*, reed., Boston: Elibron Classics, 2005.
- TWINING, W., 1985: *Theories of Evidence: Bentham & Wigmore*, Stanford, California: Stanford University Press.
- 1994: *Rethinking Evidence. Exploratory Essays*, 2.ª ed., Evanston, Illinois: Northwestern University Press.
- WEINSTEIN, J. B., 1966: «Some difficulties in devising rules for determining truth in judicial trials», en *Columbia Law Review*, vol. 66, 2. Reeditado en W. TWINING y A. STEIN (eds.), *Evidence and Proof*, Aldershot: Dartmouth, 1992.
- WRÓBLEWSKI, J., 1981: «La preuve juridique: axiologie, logique et argumentation», en PERELMAN, Ch. y FORIERS, P. (eds.): *La preuve en droit*, Bruxelles: E. BRUYLANT, citado por la traducción castellana de F. J. EZQUIAGA: «La prueba jurídica: axiología, lógica y argumentación», en J. WRÓBLEWSKI, *Sentido y hecho en el derecho*, San Sebastián: Servicio editorial de la Universidad del País Vasco, 1989.

LOS ESTÁNDARES DE PRUEBA Y LOS LÍMITES DEL ANÁLISIS JURÍDICO *

II

Ronald J. ALLEN **
Northwestern University

En términos generales, a las partes en un proceso se les pueden imponer tres tipos de cargas: la carga de alegar, la carga de producción de prueba y la carga de persuasión. Las cargas de alegar especifican las condiciones bajo las cuales las pretensiones fácticas/jurídicas serán consideradas para su inclusión en un litigio. Estas cargas pueden ser distribuidas según la preferencia del legislador, aunque generalmente son asignadas a la parte que está pidiendo un cambio en el statu quo. Hay excepciones a esta regla general diseñadas para responder a diversas elecciones políticas, como el requerimiento de que el acusado por un delito alegue específicamente una coartada aun cuando una «coartada» no sea más que una negación general de la acusación. Las cargas de alegar son significativas en tanto que constituyen el andamio del proceso; sin embargo, no plantean ningún reto conceptual interesante y, por tanto, pueden ser dejadas a un lado¹.

* Traducción de Diego Del Vecchi y Carmen Vázquez.

** John Henry Wigmore Professor of Law, Northwestern University; Fellow, China Political Science and Law University; President of the Board of Foreign Advisors, Evidence Law and Forensic Sciences Institute.

Estoy en deuda con mis buenos amigos, coautores y joviales opositores en algunos puntos de este trabajo, Larry Laudan y Michael Parke, por sus comentarios a un borrador de este artículo.

¹ La revolución generada por el caso *Twombly* en la práctica federal estadounidense sobre el llamado *pleading* (alegación) *simpliciter*, de hecho, las cargas de alegar bajo la tesis central de este artículo.

Las otras cargas —de producción y de persuasión— comprenden conjuntamente lo que normalmente se conoce como «estándares de prueba». Estos dos conceptos se denominan de distintas formas en diversos países, como carga de la prueba y estándar de prueba; pero con independencia de los nombres, los conceptos subyacentes están claros. Se puede exigir a las partes que produzcan pruebas o que de lo contrario asuman perder alguna pretensión o el caso en su totalidad. Una vez que se cuenta con una base probatoria aceptable, la decisión deberá ser tomada en un marco de incertidumbre y la carga de persuasión específica que regla de decisión ha de ser empleada. Éstas son exigencias analíticas universales implícitas en la naturaleza conceptual del litigio². Se debe producir prueba si un tercero desinteresado ha de decidir sobre los hechos y, además, la decisión es prácticamente siempre tomada bajo incertidumbre.

Aunque sigue siendo usual distinguir entre carga de producción y carga de persuasión, el profesor McNAUGHTON hace tiempo demostró que la carga de producción es una función de la carga de persuasión³. Que una carga de producción sea satisfactoria, después de la presentación de pruebas, depende de si tiene sentido continuar el litigio sobre los hechos en cuestión —en otras palabras, de si el procedimiento al respecto puede continuar—. Ello depende tanto de las pruebas aportadas como, fundamentalmente, de la carga de persuasión*. Si la parte sobre la que pesa la carga de producción presenta pruebas sobre una cuestión fáctica pero, dada la carga de persuasión, ninguna persona razonable fallaría a su favor, entonces no tiene sentido gastar más recursos en dicha cuestión y el proceso debe resolverse en consecuencia (como se hace en cualquier contexto procedimental idiosincrático). Así, las principales cuestiones conceptuales sobre los «estándares de prueba» se reducen al significado y la aplicación de la carga de persuasión.

En este artículo pretendo analizar dichas cuestiones conceptuales. En primer lugar, demostraré que tanto la importancia jurídica como la capacidad explicativa de la carga de persuasión van más allá de la simple determinación de las implicaciones de las cargas de producción. En esta primera parte del artículo, asumiré como correcta la concepción tradicional de la carga de persuasión como grado de probabilidad. A partir de esta asunción, mostraré que

al hacer depender la adecuación de los alegatos directamente de un análisis costo/beneficio sobre la procedencia del caso, donde parte del análisis es la probabilidad de éxito sobre el fondo. Cfr. ALLEN y CURY, 2010: 1.

² Algunos teóricos continentales han tratado de negar esto afirmando que el estándar de la «mínima convicción» evita los errores judiciales. Sin embargo, esto es literalmente imposible, a menos que en Europa la designación como juez conguere omisiencia, cosa que dudo. En este contexto, es bueno ver que observadores inteligentes, como Jordi FERRER BELTRÁN, advierten los límites del eslogan. Véase, por ejemplo, FERRER BELTRÁN, 2013: 35.

³ McNAUGHTON, 1955: 1382.

* *N. de los T.*: es importante advertir que en el sistema estadounidense, el juez debe decidir, con posterioridad a la presentación de las pruebas por las partes, si vale la pena continuar el procedimiento y llevarlo ante el jurado. Si la parte que tiene la carga de producción no aporta pruebas suficientes para sostener mínimamente su pretensión, entonces el proceso no continúa.

la carga de persuasión es una de las herramientas conceptuales cruciales para entender el devenir del proceso. En segundo lugar, abordaré el debate actual respecto al contenido sustantivo de la carga de persuasión y, vinculada con ésta, la disputa sobre la relación entre las cargas de persuasión y los errores. Este análisis comprenderá un examen crítico de la asunción de que la carga de persuasión es una medida probabilística. Además, intentaré demostrar que la concepción tradicional de las cargas de persuasión y los recientes esfuerzos por responder a sus deficiencias son inadecuados como descripciones empíricas del fenómeno en cuestión; asimismo, sostendré que esta insuficiencia sugiere que las herramientas analíticas empleadas en lugar de esclarecer el objeto de investigación, lo conciben erróneamente.

Una última cuestión introductoria: gran parte de la doctrina jurídica que analiza la carga de la persuasión, como en casi todos los demás temas, no tiene pretensiones descriptivas o cognitivas, sino normativas. Mi esfuerzo aquí va en sentido contrario. Voy a tratar de desvelar ciertos aspectos interesantes del problema para avanzar en su conocimiento y no voy a expresar mis propias preferencias sobre el tema. Sin lugar a dudas, una comprensión más clara sobre los hechos en cuestión puede dar lugar a un análisis normativo más convincente; sin embargo, mi objetivo en este trabajo sigue siendo comprender y no opinar sobre la materia objeto de investigación. Tal vez, irónicamente, una consecuencia de este estudio sea que si se quiere obtener una comprensión más profunda del fenómeno jurídico se deberían emplear herramientas distintas a las del análisis jurídico. Claramente, el desiderátum de buscar o no una comprensión más profunda es completamente subjetivo y, en este sentido, mi argumento será que si éste es el desiderátum entonces se siguen ciertas consecuencias.

1. LA IMPORTANCIA DE LAS CARGAS DE PERSUASIÓN

Como se ha visto, la carga de persuasión determina el significado de la carga de producción. Ésta es también la clave para comprender los distintos mecanismos de control y de limitación del jurado, tales como las defensas afirmativas, las presunciones, las inferencias, los comentarios que el juez tiene permitido hacerle al jurado a efectos de su deliberación (*judicial comment*) y los hechos notorios (*judicial notice*), así como otros mecanismos de control más macro como los veredictos dictados directamente por instrucción del juez en un juicio oral que ha sido seguido ante jurado (*directed verdicts*), los fallos sumarios y las resoluciones dictadas por el juez por considerarlas cuestiones de derecho. En cada caso, se impone la carga de persuasión directa o indirectamente a una parte o se modifica en lo que atañe a una parte. Por razones que pronto se tomarán más claras, es útil distinguir las manipulaciones directas de la carga de persuasión de las manipulaciones indirectas. Comenzaré con las primeras.

Existen numerosas asignaciones directas de la carga de persuasión. El derecho sustantivo determina quién tiene la carga de persuasión mediante la articulación de los elementos del delito y de las defensas*, cuya única diferencia es tribo en sobre quién recaer la carga de persuasión. En el ámbito angloamericano, esta asignación a menudo se hace de forma explícita en las leyes o en normas de *common law*. No hay nada conceptualmente complicado respecto de este proceso, aunque puede haber desacuerdos razonables como, por ejemplo, quién debería soportar qué carga. Por ejemplo, ¿debe la ausencia de culpa concurrente ser un requisito para ejercer una acción por negligencia o una defensa frente a una acción de este tipo?, ¿deben los acusados por un delito probar su propia demencia o debe la acusación probar más allá de toda duda razonable la ausencia de tal demencia? Puede haber consideraciones tanto de orden moral como de prácticas a favor de ambas opciones pero, una vez tomada la decisión, su aplicación es sencilla. Si el demandante o el Estado soportan la carga de persuasión respecto a una cuestión, entonces se ha calificado como un requisito para la acción; si lo hace el demandado, entonces se ha calificado como una defensa (en algunas ocasiones una «defensa afirmativa», para distinguir cosas como la negligencia concurrente de otras como la negación directa). Además, el derecho puede mezclar y combinar las cargas de producción y de persuasión. No es inherente exigir al demandado que soporte una carga de producción respecto a una cuestión y, a la vez, exigir al demandante que soporte la carga de persuasión. Por supuesto, aunque sea coherente no está claro lo que se logra con ello dado que, una vez más, la carga de producción es una función de las cargas de persuasión. Así, sólo los casos que están literalmente en equilibrio deberían tener resultados distintos, lo que es, sin duda, un conjunto excesivamente pequeño.⁴

La atribución de cargas de persuasión y, por ello mismo, la creación de defensas, se realiza también de una forma un tanto oscura, aunque directa, mediante el uso de presunciones. Esto se lleva a cabo a través del lenguaje jurídico, de modo que «de la prueba de X se extrae la presunción de Y, salvo prueba en contrario» de una parte o de la otra. El lenguaje de las «presunciones» está omnipresente pero es poco efectivo; todo lo que supone es la asignación directa de una carga de persuasión. Existen otros usos del lenguaje

* *N. de los T.*: En un sentido amplio, en el derecho anglosajón, se entiende por «defensa» cualquier alegación que, si tiene éxito, alcanzará para lograr la absolución del acusado. En sentido estricto, en cambio, sólo designa las alegaciones que suponen el reconocimiento de que el acusado ha cometido el delito pero que, en cualquier caso, debe ser absuelto. La coacción, la legítima defensa o la inmadurez son casos de este último tipo. Véase, al respecto, GARDNER, J., 2007: *Offences and Defences*. Oxford: Oxford University Press. Citado por la traducción castellana de M.^a L. MANRIQUE Y J. MITTON PERALTA, *Ofensas y defensas*. Madrid-Barcelona-Buenos Aires: Marcial Pons, 2012.

⁴ El modo en que las cosas funcionan en la práctica puede diferir de la teoría. Algunos estudios sugieren que la preponderancia de la prueba es entendida como considerablemente más alta que una probabilidad de 0,5. Desafortunadamente, estos estudios no muestran ninguna consecuencia operativa a partir de sus descripciones. Véanse, por ejemplo, SIMON Y MARSH (1971: 319); KAGEBRO Y STANTON (1985: 159), que hacen referencia a estudios empíricos acerca de cómo los decisores sobre los hechos interpretan las cargas.

de las «presunciones», pero en todos los casos es, en realidad, epifenomenal, oscureciendo relaciones más simples. En general, las presunciones sirven para cuatro funciones diferentes, tres de las cuales asignan directamente las cargas de persuasión o las afectan. Son usadas para formular reglas jurídicas —las así llamadas presunciones *iuris et de iure*—; como ya he señalado, para asignar directamente cargas de persuasión y para asignar cargas de producción, lo que como hemos visto implica cargas de persuasión.⁵ A continuación discutiré estos tres usos del lenguaje de las «presunciones» y luego volveré sobre el cuarto uso —utilizado para realizar inferencias y comentarios sobre las pruebas—. Este uso del término implica manipulaciones indirectas de la carga de persuasión y será discutido en tal contexto.

Por varias razones, el derecho establece soluciones directas para hechos que de lo contrario serían discutidos. Por ejemplo, la prueba de que una persona conducía por encima del límite de velocidad o con más de una cierta cantidad de alcohol en las venas o de que había violado una norma de seguridad, fueron en algún momento consideradas como presunciones *iuris et de iure* de negligencia. Lo que esto significa, sin embargo, es que la prueba de la condición subyacente —exceso de velocidad, conducción bajo los efectos del alcohol, violación de una norma de seguridad— satisface los requisitos de la acción y, en consecuencia, cumple el papel que de otro modo hubiese tenido la carga de persuasión. En otros casos, se asume esa misma prueba a los efectos de establecer una presunción *iuris tantum*, como la presunción de que una carta enviada por correo fue recibida por el destinatario. Esta conclusión podría ser derrotada demostrando el hecho contrario por preponderancia de la prueba, o demostrando que una persona razonable podría llegar a inferir por preponderancia de la prueba el hecho contrario —y ésta es la distinción crucial entre la asignación de una carga de persuasión y una carga de producción—. Sin embargo, la diferencia central entre los dos casos es el tipo de manipulación directa de la carga de persuasión.

De igual forma, como ha demostrado recientemente PARDO (2010: 1451), este mismo punto se aplica a los diversos mecanismos de control macro, como los veredictos pronunciados directamente por instrucción del juez, los fallos sumarios, los fallos dictados por el juez como cuestión de derecho (apartándose del veredicto de condena emitido por el jurado), y las órdenes de nuevo proceso dictadas por un tribunal de apelación. Estos mecanismos sólo deben desplegarse cuando no haya ninguna cuestión que deba ser resuelta por el jurado o cuando el jurado haya llegado a un veredicto que resulte indefendible, sin embargo, determinar cuándo es el caso depende de la carga de persuasión. Las órdenes de nuevo proceso dan al juez una mayor libertad para manejar problemas de prueba de segundo orden, ¿cuán seguro está el juez de que sólo un resultado es

⁵ ALLEN, 1981: 84-3.

defendible o de que se ha pronunciado un veredicto indefendible? Si está bastante seguro, hace uso de los llamados veredictos pronunciados directamente por instrucción del juez, de los fallos sumarios o de los fallos dictados como cuestión de derecho. Si sólo está razonablemente seguro, un nuevo proceso ordenado por el tribunal de apelación permite un control mediante un segundo jurado. El punto importante aquí, sin embargo, es que todos estos diversos tipos de análisis están guiados directamente por las cargas de persuasión.

También hay manipulaciones indirectas de la carga de persuasión. Muchos mecanismos de prueba, además de la admisión o exclusión de elementos probatorios, pueden afectar a la carga relativa de persuasión aun cuando la carga explícita de persuasión pueda no ser afectada⁶. Los comentarios sobre las pruebas que haga el juez, las instrucciones sobre las inferencias que puede hacer o no hacer el jurado (que son comentarios sobre la prueba), la exclusión de pruebas relevantes o la admisión de pruebas engañosas, cambian, todos ellos, la carga relativa de persuasión⁷. El cuarto tipo de presunciones señalado anteriormente también lo hace. Se utiliza para introducir comentarios/instrucciones para hacer inferencias mediante el lenguaje «presuncional». Una instrucción típica es que «probado X, surge la presunción de Y, que es prueba de Y». No hay significado posible para «una presunción es prueba de algún hecho» excepto incitar al juzgador a determinar un hecho mediante la determinación del hecho principal de la presunción. Esto, obviamente, es un comentario del juez sobre la prueba.

El efecto de estas diversas manipulaciones probatorias sobre la carga de persuasión puede verse al comparar dos casos cuyo único rasgo distintivo es que uno carece de dichas manipulaciones y el otro no. Supongamos dos casos con hechos esencialmente idénticos que involucran el delito de receptación (posesión de bienes robados a sabiendas), donde la única prueba de ese conocimiento es el hecho de la posesión de los bienes recientemente robados. Supongamos además que los jueces tienen la creencia generalmente correlacionada con el conocimiento del origen de los bienes; sin embargo, los jueces también creen que el común de la población que compone los jurados no conoce generalmente esa correlación. Por último, supongamos que en el primer caso el juez hace comentarios al jurado sobre las pruebas o formula instrucciones sobre las posibles inferencias respecto a la relación entre la posesión y el conocimiento, mientras que en el segundo el juez nada dice sobre este punto.

En el caso en el que no se hizo ningún comentario judicial, el jurado podría posiblemente condenar, pero la posibilidad de que absuelva es más alta que

⁶ Este fenómeno fue identificado por primera vez en ALLEN, 1980. También hace referencia a esto LAUDAN, 2011: 11.

⁷ Véase ALLEN, 1980: 332-339. Parte de lo que sigue proviene en gran parte de dicho trabajo.

en el caso en el que el juez explica las implicaciones de la posesión de bienes recientemente robados. El comentario del juez aumenta la probabilidad de una condena al incrementar el efecto de las pruebas de cargo. Para que un acusado tenga la misma posibilidad de ser absuelto, en el primer caso se tendrían que proponer pruebas de descargo más persuasivas que las producidas en el segundo caso. A pesar de que la relación formal entre el Estado y el acusado ha permanecido igual —el jurado ha aplicado el estándar de la duda razonable en ambos casos—, los comentarios judiciales han modificado las cargas relativas de persuasión mediante la alteración de la matriz fáctica con la cual el jurado toma su decisión. Esto demuestra que el efecto de los comentarios judiciales es cambiar las posiciones relativas de las partes mediante la modificación de la carga relativa de persuasión que cada parte soporta sobre una cuestión, al igual que las defensas afirmativas y la distribución de la carga de producción. La exclusión de pruebas relevantes y la admisión de pruebas irrelevantes y engañosas tienen exactamente las mismas consecuencias: el afectado de forma adversa por la decisión probatoria debe hacer más para satisfacer la carga de persuasión⁸.

Se podría plantear una objeción a este análisis que debe ser tomada en consideración y respondida. Esta objeción es que la carga de persuasión pesa sobre una de las partes y no sobre la otra, así que podría parecer que hablar de inversión de la carga relativa de persuasión no tiene sentido. Esta frecuente objeción es falsa. Las cargas de persuasión son recíprocas. Decir que el demandante soporta la carga de probar el caso por preponderancia de la prueba equivale a decir que el demandado tiene que demostrar que las elementos de juicio disponibles no prueban preponderantemente que el demandante merece ganar o, formulado en positivo, que al menos hay un 0,5 de probabilidades de que los argumentos del demandado sean verdaderos. Incluso en los procesos penales donde se dice, erróneamente, que los acusados no tienen carga de persuasión sobre los elementos del delito, en realidad sí la tienen. Una vez más, la afirmación de que la acusación debe probar los elementos del delito más allá de toda duda razonable y la afirmación de que el acusado debe mostrar que hay una duda razonable, son equivalentes.

En suma, la carga de la persuasión concebida como una medida de probabilidad convencional tiene un enorme poder explicativo. Permite descifrar el significado del término «presunción»⁹, así como hacer un análisis unificado de los mecanismos probatorios y procedimentales¹⁰ e incluso darle orden a las complejidades de la exigencia constitucional estadounidense de la prueba más allá de toda duda razonable¹¹. Sin embargo, es cada vez más claro cuán

⁸ Por tanto, también lo hacen la admisión de prueba relevante y la exclusión de prueba irrelevante, pero supuestamente estas actividades llevan al juzgador a la idealizada decisión correcta.

⁹ ALLEN, 1980.

¹⁰ PARDO, 2010.

¹¹ ALLEN, 1980.

limitadas son la corrección y la utilidad de representar la carga de persuasión como medida de probabilidad convencional. Ello, a su vez, hace que surjan dudas acerca de las herramientas de análisis jurídico que unívocamente utilizan los estudiosos del derecho o, alternativamente, acerca de lo adecuado que resulta la concepción del fenómeno que estudian. Paso ahora a estas cuestiones.

2. LOS LÍMITES DEL ANÁLISIS JURÍDICO

El atractivo de un análisis probabilístico de las cargas de persuasión es obvio. Dicho análisis es consistente con una conceptualización igualmente atractiva del proceso probatorio como actualización de probabilidades a la luz de nuevas pruebas. Ambos parecen explicar, de manera conjunta, tanto la práctica probatoria en el proceso judicial como la aplicación de la regla de decisión al final de éste. A pesar de su atractivo, la insuficiencia empírica y analítica de estas conceptualizaciones es clara. Por ejemplo, las cargas de persuasión sólo pueden desempeñar el rol que estas conceptualizaciones le asignan si involucran frecuencias relativas (de otro modo ¿cómo se puede decir algo acerca de los errores?); sin embargo, la única interpretación de la teoría de la probabilidad que tiene algún poder explicativo posible para la prueba judicial es la de los estados subjetivos de creencia. En el proceso prácticamente no se dispone de frecuencias relativas, sin duda con la interesante excepción de la prueba de ADN; y, las teorías de la probabilidad lógica y de la probabilidad en términos de propensiones no tienen aplicación alguna¹². La teoría subjetiva es obviamente falsa como descripción de lo que ocurre en los Estados Unidos: no se instruye a ningún juzgador de los hechos para que consulte sus estados mentales de creencia o sus utilidades personales. Son instruidos, se podría pensar sensatamente, para descubrir los hechos. Por otra parte, la estructura del proceso es radicalmente incompatible con las exigencias de la teoría de la probabilidad subjetiva. Los juzgadores de los hechos no son instruidos para actualizar creencias a la luz de nuevas pruebas, y en cualquier caso, la actualización sería computacionalmente intratable¹³. Cualquiera que sea el atractivo normativo de las teorías subjetivas de la probabilidad o de otros usos para los que puedan servir (como los heurísticos), son descripciones empíricamente inadecuadas.

Hay además otras dificultades. Larry LAUDAN ha realizado un ataque multifacético al significado de «prueba más allá de toda duda razonable», concebido como una probabilidad alta¹⁴. En parte de su ataque usa un trabajo que

¹² Para una discusión sobre las teorías de la probabilidad, véase GILLIES, 2000.

¹³ ALLEN, 1997: 254.

¹⁴ LAUDAN, 2006.

elaboramos conjuntamente, donde se establece que la comprensión habitual de la justificación de las cargas de persuasión en general (no sólo en los procesos penales) no resulta en absoluto convincente¹⁵. Esta justificación consiste en la distribución de errores de acuerdo con diversas decisiones políticas. En los procesos civiles, la política común es tratar igual a ambas partes y reducir el número global de errores. En los casos penales, la política es proteger a los inocentes haciendo más difícil condenar a alguien y esto supuestamente se hace prefiriendo los errores favorables a la absolución de culpables (el mantra es que resulta diez veces peor condenar a un inocente que absolver a un culpable).

A pesar de que estos lemas y formas de pensar han estado presentes tal vez por siglos, es evidente que padecen defectos irremediables. En primer lugar, hay cuatro tipos de decisiones que pueden tomarse en un juicio oral y todas tienen beneficios o costos sociales. Además de los errores, se pueden tomar decisiones correctas. Descuidar las decisiones correctas es perverso. Por ejemplo, en los casos civiles, la política de igualdad de errores se satisface cometiendo errores en cada uno de los casos individuales, siempre y cuando la proporción de casos que lleguen a juicio oral comprenda aproximadamente el mismo número de demandantes y demandados que genuinamente merezcan tener una decisión favorable. Del mismo modo, en los casos penales la proporción de diez absoluciones incorrectas frente a una condena errónea es satisfecha por 99 de cada 100 casos decididos erróneamente. Además, si desconocemos la proporción de partes que van a juicio oral y merecen genuinamente una decisión favorable, así como la relación entre la valoración del juzgador de los hechos y el verdadero estado de cosas¹⁶, lisa y llanamente no hay forma de predecir el efecto de una carga de persuasión en la toma de decisiones correctas o incorrectas. Si todo aquel que va a juicio oral es culpable, por ejemplo, no puede condenarse a inocentes, sin importar cuánto se disminuya la exigencia del estándar de prueba. El discurso tradicional sobre las cargas supone además un sistema estático, cuando de hecho es dinámico. Un aspecto de este dinamismo es que las partes deciden qué casos se llevan a juicio y pueden ajustar su decisión a la luz de los cambios en las reglas. Por tanto, la simple asunción de que un estándar de prueba menos exigente en el proceso penal causará más absoluciones falsas, o cualquier otra relación causa a efecto sugerida entre las cargas y las decisiones, obviamente no es

¹⁵ ALLEN y LAUDAN, 2008; LAUDAN y ALLEN, 2010; ALLEN y LAUDAN, 2011.

¹⁶ Por ejemplo, trabajando con una concepción probabilística, uno necesita saber cuál es la relación entre los hallazgos de la probabilidad de responsabilidad y la responsabilidad real. Se encuentra implícito en el discurso convencional que un hallazgo según el cual la probabilidad de responsabilidad es del 0,8 significa que en ocho de cada diez casos similares, la verdad de los hechos es consistente con la responsabilidad. Sin embargo, prácticamente podría darse cualquier relación entre el hallazgo de la probabilidad por parte de los decisores y el real estado de cosas. En el conjunto de todos los casos en que los decisores hallan una probabilidad de responsabilidad de 0,8, podría ser cierto que todos los casos del subgrupo fueran casos en los que la responsabilidad no deba ser declarada.

analíticamente verdadera; esto depende de cómo el sistema se comporta dinámicamente¹⁷.

Los problemas de estas formas de análisis tradicionales son incluso más profundos. Las decisiones judiciales son sólo uno de los resultados del sistema jurídico. Las partes negocian los resultados tanto en los procesos civiles como en los penales. Lo hacen a la sombra del proceso, entre otras cosas, pero los resultados de estos casos obviamente forman parte del total de los efectos del sistema jurídico en el bienestar social. Planteado más claramente, podría darse el caso que decisiones un tanto azarosas en el juicio o costos relativamente altos fueran socialmente óptimos porque promoverían el acuerdo entre las partes¹⁸.

La inadecuación empírica de las concepciones tradicionales de las cargas de persuasión ha generado tres recientes esfuerzos por ofrecer una teoría alternativa. Dos de estas teorías, formuladas respectivamente por Larry LAUDAN y Jordi FERRER BELTRÁN, son explícitamente normativas y no descriptivas, a pesar de lo cual son igualmente interesantes en sus implicaciones para las conceptualizaciones actuales de las cargas de la prueba. La tercera, de Michael PARDO y mía, es explícitamente descriptiva y, aun cuando desde mi punto de vista esto es una virtud, tiene otras interesantes limitaciones. Examinaré brevemente estas tres teorías y luego volveré a los límites del análisis jurídico tradicional.

En un reciente trabajo, Larry LAUDAN amplía sus diversas críticas sobre el concepto de prueba más allá de toda duda razonable¹⁹. Crítica, acertadamente, su notable ambigüedad y advierte que es un concepto «absurdo» debido a su subjetividad²⁰. Nos recuerda su propuesta anterior respecto a cómo entender lo que debería significar «prueba más allá de toda duda razonable».

LAUDAN se basa en la idea proveniente de la estadística y la filosofía de la ciencia según la cual un test poderoso de la verdad de cualquier hipótesis empírica, H, es su capacidad para dar cuenta de hechos o datos que resultarían absolutamente ininteligibles si H fuera falsa. Aplicada al contexto jurídico, la idea sería que la afirmación de la acusación sobre la culpabilidad del acusado adquiere una alta plausibilidad en la medida en que algunas pruebas pertinentes se hayan *prima facie* determinado como verdaderas, siendo fácilmente explicables si el acusado fuese culpable y sumamente sorprendentes si el acusado fuese inocente. Si algunos de los rasgos importantes del caso a

¹⁷ LAUDAN y FERRER afirman lo contrario, pero es evidente que las tasas de error no se pueden deducir únicamente a partir de la regla de decisión. Véase FERRER BELTRÁN (2013: 32), en particular n. 12, afirmando que la regla de decisión media la probabilidad de error tipo I y tipo II. Esto no es analíticamente cierto en el contexto jurídico. Véase también, LAUDAN, 2011: 9-11.

¹⁸ Para la discusión sobre este punto, véase ALLEN, 2011; ALLEN y LAUDAN, 2011.

¹⁹ LAUDAN, 2011. Véase también el capítulo 2 de LAUDAN, 2006.

²⁰ LAUDAN, 2011: 6.

los que la hipótesis de la culpabilidad puede dotar de sentido son igualmente compatibles con la hipótesis de la inocencia, entonces la acusación no es lo suficientemente fuerte como para justificar una condena. Del mismo modo, si hay hechos pertinentes que tienen sentido sólo para la hipótesis de la inocencia, entonces se debe absolver. En resumen, esta propuesta de estándar considera que una condena está justificada si, y sólo si: a) la hipótesis de la culpabilidad puede explicar la mayor parte de los hechos más importantes del caso, y b) la hipótesis de la inocencia no puede dar cuenta de ninguna prueba importante que sea inexplicable por la hipótesis de la culpabilidad. Así como decimos que una hipótesis científica que no hace predicciones sorprendentes y corroboradas —no previstas por sus rivales— no es muy creíble, una hipótesis de culpabilidad no respaldada por pruebas muy sorprendentes carece de apoyo empírico fuerte lo cual debe ser una condición sine qua non para una condena²¹.

LAUDAN reconoce que, junto al atractivo de la definición en abstracto, deben tomarse en cuenta las consecuencias distributivas de una propuesta de definición del más allá de toda duda razonable. Sin embargo, yo me centraré sólo en la definición propuesta pues, con independencia de su creatividad, creo que éste es un ejemplo más de los límites del análisis tradicional. Y en cualquier caso, como se señaló anteriormente, las consecuencias distributivas de cualquier regla o de toda modificación de una regla son una cuestión empírica y no una pregunta analítica.

El paso analítico fundamental que hace LAUDAN es buscar una regla de decisión en una disciplina afín para sustituir la insatisfactoria regla de decisión jurídica y, en este caso, la encuentra en las generalizaciones sobre el progreso científico. Hay dos dificultades. Primero, la relación entre las actividades típicas de la investigación científica y las del sistema jurídico es sólo contingente. Y segundo, LAUDAN elige una etapa equivocada de la ciencia, la de descubrimiento o de justificación, en lugar de usar la etapa más parecida, *i. e.*, la de aplicación. Discutiré estos dos puntos a continuación.

Consideremos las diferencias entre el proceso judicial estándar y las actividades de investigación de la mayoría de los campos de la ciencia²². El objetivo del proceso judicial casi siempre es determinar lo que ha sucedido. El objetivo no es descubrir leyes causales generales, aunque las leyes causales generales puedan ser, y a menudo son, invocadas como prueba de que un evento ha sucedido. En la mayoría de los trabajos científicos los datos no son ambiguos aun cuando sus causas son generalmente complejas. Lo que sucede, como el hecho de que las partículas tienen masa o que se atraen mutuamente, es el punto de partida para intentar explicar por qué algo pasó (como la inves-

²¹ *Ibid.*: 7-8.

²² Los siguientes párrafos están basados en ALLEN, 1994: 604.

tigación del bosón de Higgs) —en otras palabras, la investigación de una ley causal general—. Los experimentos se pueden llevar a cabo una y otra vez y, a menudo, son ajustados durante su práctica con el fin de descubrir una posible explicación. En el derecho la situación es exactamente la opuesta. La regla es tener datos primarios inconsistentes («la luz era roja», «no, era verde») y su reproducibilidad es generalmente imposible. La controversia siempre se asienta en qué sucedió. También podría estar en disputa por qué algo sucedió (lo que sea que haya ocurrido) y las partes pueden presentar generalizaciones, científicas o de otro tipo, diversas y conflictivas, para que sean usadas por el juzgador de los hechos, pero éstas serán formuladas en una estructura inferencial que lleva a una conclusión sobre un hecho o una serie de hechos y no a un universal, a una teoría o a un principio.

Hay otras dos diferencias importantes entre la ciencia y el derecho. El conocimiento científico está organizado de forma deductiva, de arriba hacia abajo. La investigación científica se centra en la posterior sistematización de la estructura deductiva de los conocimientos preexistentes a través de la eliminación de anomalías (empíricas o teóricas) o bien en la modificación o reemplazo de la estructura conceptual de un área científica. Las diferencias entre estas actividades causaron un gran revuelo en la filosofía de la ciencia de hace medio siglo²³; sin embargo, las similitudes entre los practicantes de cada uno de dichos objetivos son notables. Comparten un cuerpo bien organizado de conocimientos sustantivos y de enfoques metodológicos. Presentan discrepancias, pero sobre problemas bien definidos. Incluso aquellos que atacan las bases conceptuales de una determinada área suelen compartir los principios metodológicos y matemáticos con aquellos que están investigando en dicha área, comparten también qué cuenta como elemento de juicio y son capaces de expresar comprensiblemente el alcance de su discrepancia.

Esta descripción de la actividad de la investigación científica no se corresponde con la investigación de los hechos en el proceso judicial. La diferencia no radica en que la investigación científica es altamente complicada y los razonamientos legos sobre la vida ordinaria bastante sencillos, sino exactamente todo lo contrario. El progreso científico procede en buena medida mediante la simplificación de los problemas, particularmente a través del control de tantas variables como sea posible. El razonamiento sobre eventos ordinarios, por el contrario, prácticamente nunca es, ni puede ser, reformulado como si lo son los resultados de los experimentos controlados. Muchas variables están constante y necesariamente en juego. Y las decisiones fácticas que se toman en un juicio son aún más resistentes a la adaptación porque las complicadas características del proceso judicial están en el caldero burbujeante de la vida real.

²³ Los más importantes actores en este drama fueron Kuhn, 1962; Popper, 1934. Para un compendio de perspectivas sobre este problema, véase Suppe, 1977.

Examinemos un sencillo ejemplo que plantee algunos años atrás:

Supongamos que un sujeto testifica y luego el juzgador de los hechos debe decidir qué hacer con el testimonio. ¿Cuáles son las variables relevantes? Primero, todos los problemas normales de la credibilidad, pero considérese cuán complicadas son. La conducta no es sólo la conducta; sino un conjunto complejo de variables. ¿El testigo está sudando o temblando?, si es así, ¿debe a nervios inocentes, a la presión de la preparación, a un problema médico o es simplemente un hábito desagradable adquirido durante una infancia lamentable? ¿El lenguaje corporal sugiere veracidad o evasión? ¿Encorvarse es prueba de que se está mintiendo o de la tranquilidad de estar contando una historia concisa? ¿El testigo mira directamente a los ojos del interrogador? Si es así, ¿esto se debe a un carácter enconitable o a la confianza de un charlatan experto? ¿La inflexión de la voz indica la rectitud de alguien honesto o, por el contrario, sería la voz intensa? Y una voz intensa, ¿indica invención o preocupación por el resultado del caso?, etcétera.

La lista de variables relevantes va mucho más allá de las cuestiones de credibilidad, de las que la conducta sólo es una de ellas. Cuando un testigo formula una proposición, el juzgador de los hechos debe determinar qué es lo que la proposición pretende afirmar y qué es lo que él cree que afirma. Esa tarea involucra también un inmenso número de variables. Además, el juzgador de los hechos poseerá algún conocimiento basado en sus observaciones anteriores a la primera proposición formulada por un testigo, adquirido a través de los abogados, por ejemplo. Y hay muchos ejemplos más. Para que el derecho opere como una ciencia se requeriría que muchas de estas variables estuvieran en una estructura deductiva con una enunciación de sus condiciones necesarias y suficientes. Pero tal estructura no puede crearse, sería demasiado complejo²⁴.

La segunda diferencia es igualmente crucial. La investigación científica se centra en aquello que puede ser explorado fructíferamente con las técnicas de la ciencia moderna. Si una pregunta no puede ser abordada provechosamente con esas técnicas, como por ejemplo qué sucedió antes del *Big Bang*, es simplemente ignorada. Por otra parte, no hay ninguna necesidad de tomar una decisión sobre una cuestión experimental. Si los experimentos arrojan resultados ambiguos o si hay cierta preocupación sobre alguna variable no controlada, etc., se puede seguir trabajando hasta que los investigadores estén más o menos satisfechos con los resultados. El derecho no tiene el privilegio de seleccionar sus disputas. Sea cual sea el problema, debe lidiarse con él. Y tampoco tiene el privilegio de posponer la decisión indefinidamente, porque no decidir es decidir. El litigio conlleva argumentos sobre la modificación o preservación del statu quo. Por lo general, no decidir preserva el statu quo y, por tanto, una parte obtiene lo que desea.

En suma, no es en absoluto obvio por qué una regla de decisión proveniente de una empresa intelectual tan radicalmente diferente pueda ser provecho-

²⁴ ALLEN, 1994: 626.

samente importada al derecho. Incluso si uno se inclina a hacerlo, la regla de decisión de la empresa del descubrimiento, que es la que el profesor LAUDAN propone, no parece ser la correcta. Los tipos de variables que él identifica son aquellas que muchos científicos y filósofos de la ciencia consideran que son las que mejor asisten al descubrimiento de nuevos conocimientos científicos. Si se ofrece una nueva y supuestamente mejor explicación de algún fenómeno, ésta será persuasiva si hace predicciones sorprendentes que luego sean confirmadas, tal como Einstein predijo que la luz se curvaría en las proximidades del Sol²⁵. Sin embargo, la mayor parte de los procesos judiciales no tienen que ver con predicciones sorprendentes, sino con las más mundanas de las disputas, por ejemplo, quién está diciendo la verdad y quién miente²⁶. Hacer que la condena dependa de la producción de «pruebas muy sorprendentes» explicadas por hipótesis creativas, asombrosas, contraintuitivas o cualquier otro tipo de hipótesis fuera de lo común, supone condenar a la acusación a perder prácticamente todos los casos y, más fundamentalmente, lleva a una concepción errónea de la naturaleza del proceso judicial. Finalmente, no está claro por qué las alegaciones de los acusados no deberían estar sujetas al mismo estándar, en cuyo caso habría menos absoluciones siempre que se presente una defensa afirmativa. Pocas «condenas» y «absoluciones» funciona bien en el contexto de la ciencia, pero no en el contexto procesal (esto depende de si uno piensa que el statu quo es objetable).

También es demasiado exigente una carga que prevea que los argumentos de la acusación y de la defensa estén más o menos herméticamente sellados uno al otro, incluso si esto se limita a un conjunto cualquiera de «características importantes» a que puedan referirse (la frase es muy ambigua, lo que se considere como «importante» será muy idiosincrático y dependerá de las afirmaciones hechas por las partes). El litigio de un caso típico implica la superposición masiva de pruebas con unos pocos desacuerdos significativos²⁷. La tarea del juzgador de los hechos es decidir a quién o qué creer para resolver esos pocos desacuerdos y juzgar conforme a ello. Una vez más, si para absolver fuese suficiente con que algunas de las pruebas de cargo fuesen compatibles con la inocencia, no habría ninguna condena, pues la mayoría de las pruebas de cargo siempre serán compatibles con la inocencia. Y nada es «absolutamente sorprendente» en el proceso. Los conflictos probatorios pueden darse debido a que alguien está mintiendo para satisfacer sus propios intereses. No hay nada incomprensible en eso. A pesar de que soy un gran admirador del trabajo del profesor LAUDAN y de que, en particular, compele a los estudiosos del derecho a comprometernos seriamente con los problemas conceptuales fundamentales,

²⁵ POPPER sostuvo que el progreso científico implica hacer predicciones audaces para luego intentar falsarlas. Véase POPPER, 1959. Para uno de los test empíricos respecto a la predicción de la relatividad general, véase, DYSON, EDMINGTON Y DAVIDSON, 1920: 291.

²⁶ Véase transcripción en ALLEN *et al.*, 2006: 2-75.

²⁷ Nuevamente, *ibid.*

en este punto creo que no ha observado las diferencias básicas entre el derecho y la mayoría de las áreas científicas.

Jordi FERRER BELTRÁN, más implícitamente que LAUDAN bajo la influencia de la filosofía de la ciencia, propone un test para la prueba más allá de toda duda razonable un tanto análogo al de LAUDAN:

Según este estándar de prueba, para considerar probada la hipótesis de culpabilidad, se deben reunir las dos condiciones siguientes:

1) La hipótesis debe ser capaz de explicar los datos disponibles de una manera coherente y las predicciones sobre nuevos datos que la hipótesis permite formular deben haber sido confirmadas.

2) Todas las hipótesis plausibles restantes que ofrezcan una explicación sobre los mismos datos y que sean compatibles con la inocencia del acusado deben haber sido refutadas, a excepción de las hipótesis *ad hoc*²⁸.

Aunque los contornos de esta propuesta están menos claramente definidos que los de LAUDAN, se pueden ver cuestiones conceptuales análogas. Con respecto a la primera parte de la definición, el juicio oral (al menos en los Estados Unidos) no supone hacer hipótesis que «predigan» nuevos datos; supone decir lo que sucedió a la luz de toda la información. Este test parece implicar una inútil confusión de los contextos jurídico y científico; como se discutió anteriormente, en la empresa científica un criterio de éxito de una teoría son sus predicciones, pero ello literalmente no tiene aplicación en la decisión judicial de los hechos. En cuanto a la segunda parte de la definición, no está claro a qué se refiere en dos aspectos. Tómese nuevamente el típico caso de testigos con testimonios opuestos. Obviamente uno está diciendo la verdad y el otro no. Si se concluye que el testigo del acusado está mintiendo, entonces se tiene una muy buena hipótesis de culpabilidad; si se concluye lo contrario, entonces se tiene una buena hipótesis de inocencia. Cuán compatible sea la «información» con los diferentes testimonios depende de cómo valora tal información el juzgador de los hechos. El testimonio del testigo presentado por el acusado es «refutado» si el juzgador de los hechos concluye que él o ella está mintiendo.

La segunda dificultad de esta propuesta es que la vía de escape que excluye las hipótesis *ad hoc* parece ser ella misma totalmente *ad hoc*. FERRER BELTRÁN ofrece el ejemplo de la afirmación de que ha habido «un complot» en contra del acusado²⁹. Esto equivale simplemente a una alegación de inocencia y a una explicación para las pruebas presentadas por la acusación. Su plausibilidad dependerá del resto del caso. Esto es cierto para todas las alegaciones, ya sean del Estado o del acusado. Es imposible ver por qué esta alegación de inocencia

²⁸ FERRER BELTRÁN, 2013: 36.

²⁹ *Ibid.*: 36, nota 21.

es diferente a cualquier otra ¿Son las coartadas hipótesis ad hoc? ¿Y alegar la ausencia de un estado mental adecuado o alegar demencia? Esto equivale a un test, contenido en el primer criterio y en la primera parte del segundo, aplicable a algún subconjunto indefinido cuando éste no funcione, alegando que cualquier problema que esté interfiriendo en él es ad hoc. Si el acusado plantea una cuestión que por alguna razón no puede ser refutada, puede ser etiquetada como ad hoc; a pesar de que a menudo será equiparable a un alegato directo de inocencia. Tal vez los procesos penales europeos sean diferentes de los procesos penales estadounidenses, pero esta propuesta obviamente no es una descripción correcta de lo que ocurre en los Estados Unidos, ni tampoco parece normativamente atractiva.

Michael PARDO y yo, con base en uno de mis trabajos anteriores sobre «la teoría de la plausibilidad relativa»³⁰, hemos explorado si el mejor planteamiento de las cargas de persuasión es considerarla como un ejemplo de inferencia a la mejor explicación, haciendo referencia a los procesos cognitivos y no como solución a la pregunta sobre qué es el conocimiento. La idea central es que los casos en litigio suponen la comparación de historias alternativas y que la decisión del juzgador de los hechos versa sobre éstas o está motivada por ellas, siendo aceptada como verdadera a efectos del juicio la más plausible de las historias. En las causas penales, la acusación debe presentar una explicación plausible de la culpabilidad o, en caso contrario, perderá; y, si la acusación presenta una explicación plausible, el acusado debe presentar una explicación plausible de su inocencia o, en caso contrario, perderá³¹.

Este planteamiento no es una propuesta normativa sino descriptiva y hemos tratado de mostrar en el transcurso de los años cómo ésta es, de hecho ampliamente, la mejor forma de dar cuenta del proceso judicial actual hasta ahora identificada. Tiene la clara virtud de ser consistente con lo que sabemos sobre cómo razona el juzgador de los hechos, a diferencia de todas las aproximaciones probabilísticas existentes y las de LAUDAN y FERRER BELTRÁN. Esta también explica y ordena muchos aspectos del proceso judicial, desde varias reglas probatorias³² hasta prácticamente todas las instrucciones dadas al jurado durante el proceso probatorio («use el sentido común», «escuche a su compañero del jurado», «usted no tiene la obligación de creer en todas las pruebas, pero trate de darles un sentido»), salvo ciertamente la definición formal de las cargas de persuasión. Pero, como he mostrado antes, tal definición formal no tiene ningún sentido entendida como medida de probabilidad³³ (LAUDAN y FERRER BELTRÁN proponen una aproximación completamente nueva al respecto),

³⁰ Véase, e. g., ALLEN, 1994.

³¹ ALLEN y PARDO, 2008: 223.

³² ALLEN, 1991: 373.

³³ Hay una multitud de problemas adicionales —el efecto de la conjunción, la intratabilidad del cálculo, y así sucesivamente—.

y lo que no puede explicar la teoría de la plausibilidad relativa no lo puede hacer tampoco ninguna otra teoría de las cargas.

Sin duda, y para que quede claro, muchos de los expertos, incluidos LAUDAN y FERRER BELTRÁN, están comprometidos con propuestas normativas y no descriptivas; la nuestra es casi exclusivamente descriptiva. Tratamos simplemente de entender lo que observamos en un rango relativamente estrecho de observación. No tratamos de explicar, por ejemplo, cómo encaja el proceso judicial en procesos sociales más amplios; no tratamos siquiera de estimar qué resultados se obtienen del tipo de proceso judicial que observamos. Estas cuestiones, diríamos, son empíricas en su mayor parte, no analíticas, y sólo pueden ser respondidas empíricamente. En resumen, las amplias cuestiones planteadas por investigadores como LAUDAN y FERRER BELTRÁN, están más allá del ámbito de nuestros esfuerzos analíticos.

Permítaseme ahora recapitular el estado de la cuestión en el terreno analítico respecto a las cargas de persuasión. La concepción probabilística tradicional es obviamente falsa en todos los aspectos interesantes; las alternativas que han sido propuestas son completamente aspiracionales o muy limitadas en sus implicaciones y no abordan las amplias cuestiones en las que se han enfocado investigadores como Larry LAUDAN. ¿Qué podría explicar esto? Una posible explicación, que quiero analizar para concluir este trabajo, es que se trata de una consecuencia de un método de análisis endémico en la investigación jurídica³⁴. El método estándar de investigación jurídica consiste en descomponer los problemas en piezas manejables y analizarlos utilizando las herramientas estándar de la lógica moderna, en particular la deducción³⁵. Soy un gran entusiasta de la deducción lógica, pero su utilidad tiene sus límites o, quizá, ha sido empleada en dominios que son en sí mismos muy amplios para tan limitadas herramientas. En cualquier caso, la forma estándar de análisis jurídico es efectuar una crítica lógica y luego argumentar a favor de sustituir la norma criticada por una nueva y mejor, como han hecho LAUDAN y FERRER BELTRÁN. El trabajo de ALLEN y PARDO, sin embargo, aunque no va exactamente en la misma dirección, sólo trata una parte muy pequeña del universo relevante.

Gran parte de esos trabajos de investigación jurídica asumen implícitamente que se ocupan de un entorno estático y, en mi opinión, es aquí donde se comete un error fundamental. Una concepción más acertada del sistema jurídico, en comparación con aquella que lo ve como un sistema de reglas estáticas, es considerarlo como una pequeña parte de un problema de optimización mayor.

³⁴ Este es el tema central que he tratado en mi *Meador Lecture*, ALLEN, 2011.

³⁵ Buena parte del trabajo de LAUDAN no responde a esta descripción; él está tratando de superar los límites conceptuales, en particular en su majestuoso libro, LAUDAN, 2006. Esto no es cierto, sin embargo, de su artículo sobre la carga de persuasión (LAUDAN, 2011) en el cual se rechazan algunas de las enseñanzas de su otro trabajo.

Consideremos en primer lugar el derecho probatorio, que estaría compuesto de reglas que establecen condiciones necesarias y suficientes, debilitadas en ocasiones asignando discrecionalidad a los jueces. Estas reglas se aplicarían a las «pruebas del juicio», es decir, a los objetos materiales y a los testimonios ofrecidos que supuestamente proporcionan una base exhaustiva para deducir la responsabilidad. Pero esto es falso. Las pruebas son literalmente incomprensibles sin las reglas del lenguaje y de la lógica, sin embargo, simplemente se acepta que el juzgador de los hechos puede procesar la información y deliberar al respecto. Esto no puede hacerse sin un vasto acervo de conocimientos y la inteligencia para usarlos apropiadamente. Para que el derecho probatorio pueda ser realmente regulatorio, debe normar la interacción entre lo que se aduce en juicio y el consiguiente proceso inferencial que depende del conocimiento y de las creencias del juzgador de los hechos. Es precisamente esta laguna lo que generó mi ejemplo anterior sobre la declaración de los testigos.

Un aspecto derivado del análisis conceptual anterior es también pertinente para entender la propuesta de LAUDAN de una forma alternativa. El Dr. LAUDAN, en algunas gentiles y muy provechosas observaciones al borrador de este artículo afirmó que, a pesar de los términos exactos de su propuesta, él pretendía aludir al contexto de aplicación y no al descubrimiento y que lo que tenía en mente era algo así como el diagnóstico diferencial en el área médica. Esto acerca significativamente su propuesta a la teoría de la plausibilidad relativa de ALLEN y PARDO, además refuerza la idea de la inferencia a la mejor explicación que, juntas, son la mejor explicación de la prueba en el proceso judicial hasta la fecha. El diagnóstico diferencial, muy similar al análisis químico cualitativo, implica la eliminación y la inclusión de hipótesis sobre la base de diferenciar la información (por ejemplo, si el papel tornasol es azul, es un alcalino; si es rojo, es un ácido). Hasta cierto punto es como la formación de historias, pero persisten las lagunas conceptuales dada la naturaleza de la prueba y el desigual contexto de la empresa científica señalados anteriormente. En el diagnóstico diferencial o en el análisis químico cualitativo, al igual que en el contexto de descubrimiento (de hecho, incluso más) los datos son claros —el papel tornasol era efectivamente rojo o azul—. Con las variaciones críticas sucede lo contrario. Un testigo declara que el papel tornasol era rojo, mientras otro testigo afirma que era azul. O un testigo declara que era rojo y la otra parte ofrece pruebas o señala ciertas circunstancias (como la parcialidad o el interés) para indicar que el testigo estaba mintiendo. O el juzgador de los hechos simplemente llega por su cuenta a una conclusión mediante sus observaciones sobre la conducta de los testigos. Los procesos conllevan una intratable complejidad y ambigüedad, simplemente inexistente en la empresa científica estándar, ni en la etapa de descubrimiento ni en la de aplicación. En la empresa científica no se tiene que conseguir la aceptación de personas encargadas de tomar una decisión mediante la experiencia acumulada a lo largo de su vida; más bien, como se señaló antes y es igualmente aplicable aquí,

prácticamente todos comparten conocimiento y metodología. Sin embargo, en tanto que el diagnóstico diferencial implica la articulación creativa de explicaciones posibles, está en efecto relacionado con la teoría de la plausibilidad relativa y con la inferencia a la mejor explicación. Al igual que el trabajo de ALLEN y PARDO, no obstante, sólo podría explicar una pequeña parte del universo conceptual relevante.

El reto de una regulación jurídica efectiva es más profundo. Considérese la distinción entre comportamiento primario y comportamiento en el litigio. El comportamiento primario está constituido por las actividades de cada día, incluyendo el comportamiento económico y privado (y de cualquier otro tipo). El comportamiento en el litigio es aquél dedicado a defender los derechos jurídicos. Para la mayoría de los estudiosos del derecho, los recursos que se gastan en el proceso son una pérdida social de eficiencia que no agrega nada a la comunidad y, por tanto, se deben minimizar; pero esto obviamente podría ser falso. El comportamiento en el litigio afecta al comportamiento primario y crea incentivos para éste, y viceversa. Por ejemplo, quizá las dificultades del litigio crean incentivos para que las personas ajusten su comportamiento del fin de evitar ir a juicio y optimizar así la productividad social. Alternativamente, los costos de los litigios podrían crear incentivos perversos que sobresudaran el comportamiento primario productivo. El punto es que el comportamiento primario y el comportamiento en el litigio son partes de problemas de optimización más grandes que necesitan ser analizados precisamente como tales.³⁶ Esto requiere pensar las cosas en términos analítico-sistemáticos más que en los términos de la lógica de predicados de primer orden o en los términos silogísticos comunes en la investigación jurídica.

Recientemente he abordado este tema en mi participación en las *Meador Lecture Series* de la Facultad de Derecho de la Universidad de Alabama:

[...] la realidad del sistema jurídico no es la de contextos nítidos, ordenados, simples y estáticos, sino más bien la de los procesos evolutivos, calderos en ebullición confusos, complicados y orgánicos. La herramienta estándar que se utiliza para regular este caldero en ebullición son las reglas, la fricción entre esta herramienta y muchos de los usos que se le da es lo que explica en general por qué la determinación de los hechos y la regulación jurídica son a menudo consideradas problemáticas. Esta misma relación explica muchos problemas jurídicos como, por ejemplo, en orden ascendente de importancia, las extrañas implicaciones del análisis estándar sobre el error, el debate sobre las reglas *versus* los estándares y el significado de «derechos».

El simple concepto de «regla», como fijación de condiciones necesarias y suficientes a partir de las cuales pueden deducirse las consecuencias, es un ejemplo de la lógica monotónica, en la que la suma de postulados o asunciones

³⁶ Larry LAUDAN comienza este proceso en LAUDAN, 2006, y algo de esto motiva también nuestro trabajo conjunto.

simplemente incrementa lo que puede ser deducido de las asunciones previas. Las lógicas monotónicas son herramientas poderosas, tal como lo demuestran el surgimiento de las matemáticas modernas y el éxito de muchas áreas científicas. Estas funcionan mejor cuando sus asunciones operativas capturan acertadamente los dominios respectivos; en la famosa dicotomía de HAYEK, esto significa que funcionan bastante bien en sistemas artificiales como los juegos y no tan bien en los sistemas evolutivos u orgánicos que caracterizan gran parte de la condición humana³⁷. Gran parte del debate sobre las reglas y sus límites versa, a menudo implícitamente, sobre la complejidad del dominio relevante y la tolerancia a errores de diversa índole. Cuando aumenta el número de variables pertinentes o cuando algunas de ellas son continuas y no discontinuas, el problema deductivo rápidamente deviene computacionalmente intratable, incluso para los ordenadores y no se diga para los seres humanos. Y, por supuesto, si aparece una nueva variable que no había sido prevista, todas las deducciones son posibles, por así decirlo. En cualquier caso (intratabilidad computacional o falta de imaginación), las aproximaciones algorítmicas que dependen de las normas existentes generan las críticas estándar de la naturaleza indeterminada de las reglas. En realidad, no es que las reglas sean indeterminadas sino que se les ha puesto a realizar una tarea para la que no son óptimas³⁸.

En esta *Lecture* me atreví a sugerir que el problema central del sistema jurídico es similar al problema central de la racionalidad, que es la tratabilidad de la complejidad. En ambos casos, a simples herramientas deductivas se les dan usos cuyos resultados son subóptimos. Esto da lugar a una importante cuestión: ¿qué otros enfoques pudieran ser más fructíferos? Una posible respuesta que entonces planteé, inspirada en un brillante artículo de un investigador en inteligencia artificial, con el que me encontré hace muchos años, TIM VAN GELDER, fue que la pugna de la racionalidad por tomar tratable la complejidad debe parecerse menos a la computación digital y más a un regulador dinámico, como el *Watt Centrifugal Governor* (regulador centrífugo) que fue una parte fundamental de la revolución industrial³⁹. De forma similar, el análisis jurídico necesita evolucionar para ocuparse de las complejidades de los sistemas. Para mi objetivo en este trabajo, el ejemplo de VAN GELDER es más una metáfora que un argumento, pues proporciona sólo la sugerencia de otras posibilidades y no un programa de investigación definido; a pesar de ello, no deja de ser interesante.

El crecimiento de la industria textil en Inglaterra dependía de una fuente de energía consistente y con una variabilidad muy limitada. La máquina de vapor proveyó la energía, pero sus pistones producían estallidos de energía periódicos y no una corriente fluida y continua. Los volantes de inercia o volantes motor ayudaron, pero siguieron siendo inadecuados. Como VAN GELDER puso de manifiesto, una solución potencial a este problema es computacional tal como se muestra en el siguiente esquema:

Estática

1. Medir la velocidad del volante.
2. Comparar la velocidad actual con la velocidad deseada.
3. En caso de no haber discrepancia, volver al paso 1. En caso contrario:
 - a) Medir la presión de vapor actual.
 - b) Calcular la alteración deseada de la presión de vapor.
 - c) Calcular el ajuste de válvula necesario.
4. Efectuar el ajuste de válvula necesario.

Volver al paso 1.

TIM VAN GELDER, «What Might Cognition Be, If Not Computation?», *The Journal of Philosophy*, vol. 92, núm. 7 (julio, 1995), p. 348.

Desafortunadamente, esta solución computacional requiere de una persona para ejecutarla, lo que es costoso, y raramente producirá una fuente de energía lo suficientemente fluida. JAMES WATT resolvió este problema, colocó brazos móviles sobre un eje en el centro del volante de inercia cuyo movimiento se transmite instantáneamente a la válvula que regulaba el flujo del vapor. A medida que la rotación del volante se acelera, los brazos se extienden, esto se transmite a la válvula y ésta se cierra hasta alcanzar el equilibrio adecuado, y viceversa, como se muestra en la siguiente figura:

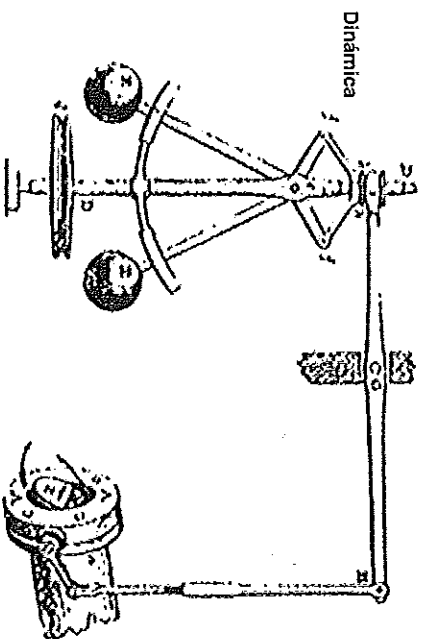


Figura 1 *

Dinámica

* Regulador centrífugo para controlar la velocidad del motor a vapor — de J. Farley, *A Treatise on the Steam Engine: Historical, Practical, and Descriptive* (London: Longman, Rees, Orme, Brown, and Green, 1827).
Tim van Gelder, «What Might Cognition Be, If Not Computation?», *The Journal of Philosophy*, vol. 92, núm. 7 (julio, 1995), p. 348.

³⁷ HAYEK, 1973.

³⁸ ALLEN, 2011.

³⁹ VAN GELDER, 1995: 348.

Con independencia de si el regulador centrífugo capta algo importante sobre la racionalidad, observar al sistema jurídico con esta metáfora en mente puede ser fructífero. El punto crucial es que algunos problemas pueden resolverse de forma diferente, no mediante argumentos deductivos⁴⁰. Ésa es la limitación crucial del trabajo que se discutió anteriormente y una constante limitación del estudio del derecho en general⁴¹. Sin lugar a dudas, es útil descomponer los problemas en partes más pequeñas, etc., sin embargo, al mismo tiempo tal proceso puede resultar contraproducente, distorsionando, en lugar de resaltar, la naturaleza de la entidad bajo examen. Los análisis que se han hecho sobre las cargas de persuasión, incluido el nuestro, ponen de manifiesto este punto. Tal vez la alternativa es pensar en el sistema jurídico como la dinámica de fluidos trata al flujo de los líquidos y gases para, en otras palabras, afrontar el desorden de la vida real en lugar de tratarlo de forma abstracta desde afuera.

Otra posibilidad, que actualmente estoy explorando, consiste en la adecuación empírica de concebir al sistema jurídico como un problema de optimización, suponiendo que es en esencia algo parecido a una red neuronal. La concepción tradicional del «sistema jurídico» sigue estando ampliamente basada en el concepto de regla. Esto es obviamente cierto para la mayoría de los ámbitos de la dogmática; de hecho, «dogmática» y, por tanto, «ámbitos de la dogmática» hacen referencia principalmente a reglas y conjuntos de reglas, aunque su ambigüedad y sus limitaciones son bien conocidas. Incluso el gran debate iusfilosófico de los últimos cincuenta años entre HART, DWORKIN y sus respectivos partidarios es en gran medida un debate sobre las reglas y sus límites, conceptualizándolas como determinadas en gran parte jerárquicamente suponiendo la articulación, la aplicación y la revisión por tribunales superiores. Mi intuición es que una concepción empírica más adecuada del sistema jurídico incluye un gran número de nodos decisionales o *loci*, como en una red neuronal, con comunicación casi instantánea con un incalculable número de otros nodos. Por ejemplo, cuando un tribunal toma una decisión, ésta es transmitida a muchas otras personas con problemáticas similares en todo el país, no sólo al tribunal de apelación jerárquicamente superior. Si la resolución es apelada, también el resultado de ésta es transmitido mediante una red de partes interesadas. Incluso cuando la resolución es definitiva, otras partes pueden refutarla directamente en otras jurisdicciones o reaccionar frente a ésta en una casi infinita variedad de formas. Cuanto más amplia sea la potencial ruptura de cualquier decisión concreta, mayor es el incentivo para otros «nodos» del sistema para reaccionar a ésta, para intentar modificarla, suscribirla o rechazarla. Por tanto, las reglas o las aplicaciones de las reglas son perturbaciones de un sistema que causan un efecto o tienen consecuencias que se propagan gradualmente en todo el sistema, dando lugar a diferentes reacciones hasta que finalmente se alcanza

⁴⁰ Para una discusión en este sentido, véase HAJACK, 2007: 20.

⁴¹ El aumento de la investigación empírica en las universidades de Estados Unidos es en cierta medida una reacción ante este problema.

algún nuevo equilibrio. Así, la cuestión que plantea una regla o su aplicación no es si ésta es «correcta» sino si es óptima, donde la optimización es juzgada en relación con sus efectos perturbadores en el equilibrio.

Creo que esta forma de concebir al sistema jurídico tiene un enorme poder explicativo. Explica, por ejemplo, muchas de las dificultades de las cargas de persuasión antes señaladas, que surgen por considerar a las cargas de persuasión como reglas deductivas y no como parte de un sistema complejo. La pregunta que debe plantearse no es «cuál es la carga de persuasión», sino qué efectos tienen las diferentes articulaciones de ésta.

Creo también que el poder explicativo de esta concepción es mucho más extenso. Por ejemplo, explica por qué la agenda de DWORKIN, tanto si se centra en el sistema jurídico o en la moral, es literalmente imposible⁴². Así como ningún algoritmo podría predecir cómo una molécula de agua realmente caerá por la montaña mediante un arroyo (aun cuando obviamente es un proceso completamente determinado), nadie podría saber cuál es la «mejor» reconstrucción ni del derecho ni de la moral. En cambio, uno puede participar en una negociación interminable sobre los diferentes efectos de las distintas articulaciones de las reglas (o lo que fuese). Hasta ahora, me parece que la mejor explicación del sistema jurídico es, precisamente, una negociación sin fin sobre tales efectos, llevada a cabo a través de un gran número de nodos interconectados, en donde las cargas de prueba estimulan las decisiones en una u otra dirección.

BIBLIOGRAFÍA

- ALLEN, R. J., 1980: «Structuring Jury Decisionmaking in Criminal Cases: A Unified Constitutional Approach to Evidentiary Devices», *Harvard Law Review*, 94 (2): 321-368.
- 1981: «Presumptions in Civil Actions Reconsidered», *Iowa Law Review*, 66: 843-867.
- 1991: «The Nature of Juridical Proof», *Cardozo Law Review*, 373: 413-420.
- 1994: «Factual ambiguity and theory of evidence», *Northwestern University Law Review*, 88: 604-640.
- 1997: «Rationality, algorithms and juridical proof: a preliminary inquiry», *International Journal of Evidence and Proof* (Special Issue): 254-275.
- 2011: «Rationality and the Taming of Complexity», *Northwestern Public Law Research Paper* (N. 11-51)
- ALLEN, R. J. et al., 2006: *Evidence: text, problems, and cases*, Aspen Publishers.
- ALLEN, R. J., y ALAN, E. G., 2010: «Conley as a Special Case of Twombly and Iqbal: Exploring the Intersection of Evidence and Procedure and the Nature of Rules», *Pennsylvania State Law Review*, 115: 1-47.

⁴² Véase, por ejemplo, DWORKIN, 1986 y 2011.

- ALLEN, R. J., y LAUDAN, L., 2008: «Deadly dilemmas», *Texas Tech Law Review*, 41: 65-92.
- 2011: «Deadly Dilemmas III: Some Kind Words for Preventive Detention», *Journal of Criminal Law and Criminology*, 101: 781-802.
- DWORKIN, R., 1986: *Law's empire*, Cambridge: Belknap Press of Harvard University Press.
- 2011: *Justice for hedgehogs*, Cambridge: Massachusetts: Belknap Press of Harvard University Press.
- DYSON, F. W.; EDDINGTON, A. S., y DAVISON, C., 1920: «A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations Made at the Total Eclipse of May 29, 1919», *Philosophical transactions of the Royal Society*, 220: 291-333.
- FERRER BELTRÁN, J., 2013: «La prueba es liberada, pero no tanto: una teoría de la prueba casilenthanniana», en este volumen.
- GILLES, D., 2000: *Philosophical theories of probability*, Philosophical Issues in Science, London-New York: Routledge
- HAACK, S., 2007: «Something but not All», *Ratio Juris*, 20: 1-31.
- HAYEK, F. A. VON, 1973: *Law legislation and liberty*, Chicago: University of Chicago.
- KAGEHIRO, D. K., y CLARK, W. S., 1985: «Legal vs. Quantified Definitions of Standards of Proof», *Law and Human Behavior* 9 (2): 159-178.
- KUHN, T. S., 1962: *The structure of scientific revolutions*, Phoenix Books in Science, Chicago: The University of Chicago Press.
- LAUDAN, L., 2006: *Truth, error and criminal law: an essay in legal epistemology*, Cambridge Studies in Philosophy and Law, Cambridge: Cambridge University Press.
- 2011: «Is it Finally Time to Put "Proof Beyond a Reasonable Doubt" Out to Pasture?», *University of Texas Law, Public Law Research Paper No. 194*.
- LAUDAN, L., y ALLEN, R. J., 2010: «Deadly Dilemmas II: Bail and Crime», *Chicago Kent Law Review*, 85: 23-42.
- MCKNAUGHTON, J. T., 1955: «Burden of Production of Evidence: A Function of Burden of Persuasion», *Harvard Law Review*, 68: 1382-1391.
- PARDO, M., 2010: «Pleadings, Proof, and Judgment: A Unified Theory of Civil Litigation», *Boston College Law Review*, 51 (5): 1451-1509.
- PARDO, M., y ALLEN, R. J., 2008: «Juridical Proof and the Best Explanation», *Law and Philosophy*, 27: 223-268.
- POPPER, K. R., 1934: *The logic of scientific discovery*, London: Hutchinson.
- SIMON, R. J., y MAHAN, L., 1971: «Quantifying Burdens of Proof: A View from the Bench, the Jury, and the Classroom», *Law & Society Review*, 5 (3): 319-330.
- SUPPE, F., 1977: *The structure of scientific theories*, citado por la 2.ª ed., Urbana: University of Illinois Press.
- VAN GELDER, T., 1995: «What Might Cognition Be, If Not Computation?», *Journal of Philosophy*, 92 (7): 345-381.

EL PROBABILISMO JURÍDICO: UNA DISENSIÓN EPISTEMOLÓGICA *

III

Susan HAACK **
University of Miami

[E]s claro que algunas cosas son más ciertas, mientras que otras son objeto de conjeturas azarosas. Para un hombre racional, hay una escala de [grados de dubitabilidad] ***, desde las proposiciones lógicas y aritméticas simples y los juicios de percepción, en un extremo, hasta cuestiones como la de la lengua que hablaban los micéneos o «qué canción cantaban las sirenas», en el otro [...] [E]l hombre racional, que asigna a cada proposición el grado correcto de credibilidad, se guiará por la teoría matemática de las probabilidades cuando ella es aplicable [...] Sin embargo, el concepto de «grado de credibilidad» es mucho más ampliamente aplicable que el de probabilidad matemática.¹

Russell está en lo correcto. El cálculo matemático de probabilidades está perfectamente bien *en su lugar*; pero ese lugar tiene límites. En particular, dicho cálculo matemático no puede ser la columna vertebral de una teoría epis-

* Traducción de María José VIANA y Carlos BERNAL. Revisión técnica de por Carmen Vázquez. © 2011 Susan HAACK. Derechos reservados. Este artículo retoma algunos aspectos discutidos en HAACK, 2012 (escrito en 2010) con el fin de precisar algunas ideas que previamente no había sido capaz de describir más que vagamente.

** *N. de los T.*: En la traducción al castellano de este texto de RUSSELL, la noción «*grades of doubtfulness*» ha sido traducida como «*duda*», sin embargo, la autora ha pretendido usar la traducción del concepto de *credibility* de DESCARTES, al entender que refleja mejor la idea a la que hace referencia RUSSELL. Por tanto, en este punto, la cita se aparta de la traducción castellana del texto comentado.

¹ Las cursivas son mías.

temológica sería, dado que no arroja luz sobre el concepto crucial que Russell llama «credibilidad racional» y yo «aval» (en inglés, «*warrant*»)*. Una de las consecuencias de esto, como defenderé aquí, es que para entender los grados de prueba y los estándares probatorios en el derecho, no debemos acudir a la teoría de la probabilidad, sino, en su lugar, a una rama de investigación más antigua y menos formal: la epistemología.

Así pues, en el apartado 1, argumentaré que los estándares de prueba en el derecho se entienden mejor en términos del grado en que las pruebas presentadas deben avalar la conclusión (de culpabilidad u obligación del acusado o de mandato) de un caso a resolver. En el apartado 2, esbozaré mi análisis sobre el concepto de aval y demostraré que los grados de aval no pueden identificarse con las probabilidades matemáticas. De esto inmediatamente se sigue que el «probabilismo jurídico» —es decir, la tesis según la cual los grados de prueba en el derecho deben ser identificados con las probabilidades que figuran en los cálculos matemáticos de probabilidad— es erróneo. Estos argumentos son completamente generales, es decir, aplicables a cualquier tipo de probabilismo jurídico. Sin embargo, el resto del trabajo se enfocará concretamente en una forma popular de probabilismo jurídico —el bayesianismo subjetivo; y, mostraré, a través del análisis de dos casos famosos, y famosamente complicados, los fallos de tales aproximaciones para ayudarnos en nuestra comprensión sobre la prueba jurídica—. Así, en el apartado 3 demostraré que el análisis bayesiano de KADANE y SCHUM sobre la prueba en el caso *Sacco y Vanzetti* es seriamente defectuoso y que una aproximación desde mi teoría puede ser significativamente mejor; y, en el apartado 4, sostendré que mi aproximación, a pesar de no ser probabilística, es perfectamente capaz de dar cuenta adecuadamente de la prueba estadística, para lo cual retomo el análisis bayesiano de FINKELSTEIN y FAIRLEY sobre la prueba en el caso *Collins*, demostrando que también es seriamente defectuoso —y, de nuevo, que mi aproximación a este problema resulta significativamente mejor—.

1. LOS ESTÁNDARES DE PRUEBA DEBEN SER ENTENDIDOS COMO GRADOS DE AVAL

En todo sistema jurídico es necesario determinar, de alguna forma, cuestiones de hecho; sin embargo, los distintos sistemas jurídicos tienen (o han tenido) diferentes maneras de hacerlo.

En la etapa anglosajona temprana, los tribunales confiaban en el juicio por juramento: el acusado debía jurar ante Dios que no era culpable y los «garantes del juramento» (también conocidos como «compurgadores» o, en inglés,

* *N. de los T.*: «Aval» debe entenderse como una noción gradual, no categórica; que no comunica ninguna implicación de «certeza»; y, finalmente, como un criterio objetivo y no subjetivo.

«*con-jurors*») ² debían ser llamados a jurar que dicho juramento no era falso. El libro de Lisi OLIVER sobre la historia del derecho inglés temprano nos dice que alrededor del 695 d. C., dependía del rango del acusado si era necesario que jurara y, de ser así, si se precisaba del apoyo de ayudantes del juramento y, en su caso, de cuántos ayudantes. Por ejemplo, la palabra del rey o de un obispo era suficiente sin juramento; en cambio, un sacerdote o un diácono tenían que jurar pero no necesitaban ayudantes del juramento como respaldo. El juramento de un hombre libre precisaba de tres ayudantes del juramento del mismo rango ³. El libro más popular de Sadakat KADRI sobre la historia de los procedimientos judiciales nos dice que en el año 899 la reina Uta de Alemania fue absuelta del cargo de adulterio sólo después de que ochenta y dos caballeros jurasen su inocencia. Cualquier defecto de forma en el procedimiento, según KADRI —¡incluso si el testigo mismo hubiese movido del relicario la reliquia sagrada sobre la cual juró!— era una causal de defensa en contra del perjurio ⁴. Frederick MATTLAND narra que la práctica de la «*Wager of Law*», es decir, el procedimiento de autodefensa mediante un juramento de inocencia, no fue formalmente abolida en Inglaterra sino hasta 1833 ⁵.

Según Robert BARTLETT, durante el reinado de Carlomagno surgieron una gran variedad de «ordalías»: juicios por inmersión en agua gélida o en caldera de agua hirviendo, juicio de la cruz, o mediante alambres de púas, etcétera ⁶. En un «juicio por hierro candente», por ejemplo, se le pedía al acusado que sostuviera una barra de hierro ardiendo y luego se revisaban sus heridas para ver si habían sanado adecuadamente o si se habían infectado —lo que se tomaba como señal de culpabilidad— ⁷. Sin embargo, continúa BARTLETT, el juicio por ordalías era usado como último recurso, sólo cuando no existiera otra forma de descubrir la verdad ⁸. El derecho de la ciudad en Enns (Austria), establecido en el año 1212, proveía que en procesos seguidos por violación, el acusado tenía la opción del juicio por ordalías si había sólo dos testigos, pero no si existían siete o más testigos ⁹. Para el año 1215, cuando el Cuarto Concilio Luterano prohibió a los sacerdotes formar parte de tales pruebas, esta práctica se había extinguido casi completamente ¹⁰; y para el año 1300, como lo señala Bartlett, era ya un «vestigio en todas las jurisdicciones» ¹¹.

² Asumo que este es el origen de la moderna palabra inglesa «*conjuror*», que significa «mago, prestidigitador».

³ OLIVER, 2002: 174 y ss.

⁴ KADRI, 2005: 20.

⁵ MATTLAND, 1919: 14.

⁶ BARTLETT, 1986: 9 y ss.

⁷ BARTLETT, 1986: 21, 33; Véanse también MATTLAND, 1909; *Lecture II*; KADRI, 2005: capítulo 1.

⁸ BARTLETT, 1986: 27-28.

⁹ *Ibid.*: 29.

¹⁰ MATTLAND, 1919; *Lecture II*, sugiere que la desaparición del juicio por ordalías fue el resultado de la decisión del Cuarto Consejo Luterano; pero el capítulo 1 del libro de FARUFFO (2009) muestra que la decisión de la Iglesia simplemente ratificó un cambio que ya ocurría en la práctica jurídica.

¹¹ BARTLETT, 1986: 34.

De acuerdo con George NELSON, el juicio por combate, en el cual las partes de un proceso debían literalmente pelear, se introdujo en Inglaterra bajo el gobierno de Guillermo el Conquistador. Un inglés que decidía evitar el duelo, señala NELSON, era sujeto a juicio por ordalía; sin embargo, si un normando escogía no batirse en duelo tenía la opción de defenderse a sí mismo mediante juramento.¹² Ya a inicios de la época de Enrique II, el juicio por combate se había confinado a una clase muy específica de casos; y su aplicación se restringió aún más con el paso del tiempo.¹³ Sin embargo, el derecho inglés no abolió oficialmente esta práctica hasta 1819, con la aprobación de la *Appeal of Murder Act* (ley de apelación de la pena de muerte), después de que Abraham THORNTON consiguiera escabullirse de una pena de muerte tras ofrecer defenderse «con su propio cuerpo»¹⁴.

En sus formas más agresivas la cultura adversarial del derecho estadounidense es algunas veces inquietantemente similar al juicio por combate; por ejemplo, alguna vez he escuchado a científicos describir su lamentable experiencia como peritos en interrogatorios cruzados como un «juicio por agua gélida»¹⁵; y, por supuesto, todavía exigimos a los testigos que juren «decir la verdad, solamente la verdad y nada más que la verdad» (aunque imagino que en nuestros días pocos creen seriamente que si mienten bajo juramento sufrirán un castigo divino). Por otra parte, incluso actualmente en países que adoptan el derecho islámico tradicional (*sharia*), se le da el doble de peso al testimonio de un testigo hombre que al de una mujer¹⁶; y hasta el año 2007 el derecho pakistaní que regula las conductas sexuales ilícitas (*zina*) exigía cuatro testigos oculares masculinos musulmanes para probar un delito de violación¹⁷.

Aunque me concentraré principalmente en el derecho actual de los Estados Unidos y utilizaré casos estadounidenses para ilustrar mi argumento, me ocuparé sobre todo de los grados de prueba que de las reglas sobre la prueba características del *common law*. De esta manera espero que al menos la mayor parte del trabajo sea también relevante para las jurisdicciones de derecho romano-germánico.

La necesidad jurídica de determinar las cuestiones de *hecho* naturalmente conduce hacia reglas sobre las cargas y los estándares de *prueba*. «La carga de la prueba» incluye principios respecto a qué parte está obligada a producir pruebas (también conocida como «carga de producción») así como principios que establecen qué parte está obligada a demostrar los elementos del caso al grado

¹² NELSON 1890: 31-32.

¹³ *Ibid.*: 33-36.

¹⁴ *Ibid.*: 328-331.

¹⁵ «*Pouring cold water on an idea*» es una locución inglesa que se usa para menospreciar algo.

¹⁶ JANNI y KARLMEYER, 2007: 32.

¹⁷ ANONIMO, 2006. Sobre el derecho pakistaní, en general, véase SIAL y JOBAL (2005) (si lo entendido correctamente, «*zina*» significa algo como «conducta sexual ilícita» e incluye, por ejemplo, el adulterio y la violación; lo cual posiblemente explica esta exigencia extraordinariamente alta de prueba).

exigido (también conocida como «carga de persuasión»)¹⁸. Los estándares de prueba especifican el grado o el nivel de prueba que debe satisfacerse en los diversos tipos de procesos¹⁹; en Estados Unidos, «más allá de toda duda razonable» en los procesos penales; «preponderancia de la prueba» o «más probable que su negación» en casos civiles ordinarios; y «prueba clara y convincente» en procesos especiales como la cesación de los derechos paterno-filiales, las cuestiones relacionadas con la ciudadanía, el contenido de un testamento, etc.²⁰ Y, finalmente, se encuentra la «sospecha razonable» o «causa probable» que se exige para un registro²¹; y también el requisito señalado en la regulación sobre la pena de muerte en Texas respecto a que el jurado deberá imponer dicha pena al acusado sólo si encuentra «más allá de toda duda razonable» que «hay una probabilidad» de que el sujeto será peligroso en el futuro²².

Alguna asignación de cargas de prueba y estándares probatorios es necesaria para asegurar que se ha alcanzado un resultado. Y la justificación que subyace en las cargas y los estándares de prueba *específicos* se funda también en consideraciones políticas: más claramente, la exigencia de que un cargo penal deba ser probado *por la acusación* y que deba ser probado *más allá de toda duda razonable* descansa en la idea de que es mucho peor condenar a alguien por un crimen que no cometió que absolver a alguien por un crimen que sí cometió. En el caso *Adington v. Texas* (1979), el magistrado presidente de la Corte Suprema estadounidense, magistrado Burger, escribió como criterio resolutivo de la Corte que en los casos típicos de derecho civil sobre controversias financieras, el estándar de la «preponderancia» asegura que «[l]os litigantes [...] comparten de manera más o menos igual el riesgo de error»; mientras que en un caso penal, dado que el acusado tiene en juego

¹⁸ GRAHAM, 2007: 577-579.

¹⁹ Al igual que el *common law* en general, también estos estándares han evolucionado gradualmente. Una estudiosa del área sugiere que el estándar más exigente y actualmente más conocido, «más allá de toda duda razonable», ha estado integrado en el derecho anglosajónico al menos durante setecientos años y posiblemente sea más de mil años» (DELOCCIO, 1986: 25). CHADBOURN (1981: 405). Sin embargo, data a principios de 1700 «la precisa distinción» de la exigencia de probar más allá de toda duda razonable en casos penales. El magistrado BENNAN, en 1970 en la sentencia del caso *In Re Whship*, aclara un poco el asunto: «[l]a exigencia de que la culpabilidad de una acusación penal sea establecida más allá de toda duda razonable data al menos de nuestros primeros años como nación»; y, como MCCORMACK sostiene, «la demanda de un grado más alto de persuasión en casos penales fue recurrentemente expresada desde tiempos antiguos, [aunque] su cristalización en la fórmula "más allá de toda duda razonable" parece haber ocurrido solo hasta 1878». *In re Whship*, 379 U. S. 358, 362, 90 S. Ct. 1068, 1071, citando a MCCORMACK, 1954: 681-682.

²⁰ BRONN, 2006: 488.

²¹ La terminología de «causa probable» deriva de la interpretación judicial de la Cuarta Enmienda de la Constitución estadounidense. Brevemente y muy aproximadamente, hay causa probable «cuando los hechos conocidos y las circunstancias son suficientes para garantizar a un hombre de prudencia razonable la creencia de que encontrará los bienes objeto de contrabando o las pruebas de un delito». *Minnis v. Gales*, 462 U. S. 213, 228 (1983); véase también *Ornelas v. United States*, 517 U.S. 690, 696 (1996).

²² Texas Code Crim. Proc. ANN. (Código de Procedimiento Penal de Texas), art. 37.071 (West Supp. 2009).

cuestiones sumamente importantes, exigimos probar «más allá de toda duda razonable» para asegurar que «nuestra sociedad carga sobre sí casi la totalidad del riesgo de error»²³. Tales consideraciones políticas ciertamente no quedan fuera de la reflexión filosófica²⁴; sin embargo, no son éstas las que me interesa abordar aquí —sino, más bien, comprender a qué equivalen los estándares de prueba—.

Es fácil ordenar los estándares desde el más exigente hasta el más débil: «más allá de toda duda razonable»; «preponderancia de la prueba»; «prueba clara y convincente»; «sospecha razonable»²⁵. Sin embargo, ninguno de ellos ha sido definido de manera precisa; ni tampoco es claro que lo ideal sería tener definiciones precisas, aun cuando éstas fuesen posibles (de hecho, algunos Circuitos federales en Estados Unidos aconsejan a los jueces no intentar definir qué significa «duda razonable», sino dejar que el jurado discierna su significado)²⁶. Sin embargo, es necesario responder a una pregunta más fundamental: ¿qué tipo de grados son exactamente los grados de prueba?

El magistrado HARLAN, en su voto particular concurrente en el caso *Winship* (1970), escribió que la función de los estándares de prueba es «instruir al juzgador de los hechos sobre el grado de confianza que la sociedad cree que él debería tener sobre la verdad de las conclusiones fácticas para un tipo particular de decisión»²⁷. Los autores de la sexta edición del reconocido libro *McCormick on Evidence* son aún más explícitos en su enfoque psicológico: la fórmula de la «duda razonable», afirman, «alude a algo en lo que estamos realmente interesados en el derecho, al estado de la mente del jurado»; mientras que «la preponderancia de la prueba» y «la prueba clara y convincente» han sido malentendidos porque «desvían la atención hacia la prueba». Pero esto, sostienen, es un paso que nos aleja de la cuestión esencial, es decir, el grado de creencia de los miembros del jurado; las pruebas sólo son «el instrumento mediante el cual la mente del jurado es influenciada»²⁸.

En mi opinión, las cosas son exactamente al revés. Es cierto que el lenguaje en que se expresan los estándares de prueba es parcialmente psicológico: al menos hablar de «carga de persuasión» suena subjetivo, sugiriendo que la

²³ *Addington v. Texas*, 441 U.S. 423, 423-4 (1979), 90 S. Ct. 1804, 1808 (1979).

²⁴ En este contexto pienso en el perspicaz comentario de Percy BRIDGMAN sobre la tendencia de los seres humanos a justificar (o pseudojustificar) instituciones sociales que han surgido de una manera más o menos arbitraria en el éxito o en el fracaso: «Un perro se conforma con dar tres veces una vuelta antes de echarse, pero un ser humano tendría que inventar una explicación de por qué es la mejor forma de echarse». BRIDGMAN, 1933: 368.

²⁵ Aunque puede valer la pena observar que el Tribunal de Apelaciones del caso *Winship* había sugerido que sólo hay una «diferencia tenue» entre los estándares de la duda razonable y la preponderancia —una sugerencia que, sin embargo, fue firmemente rechazada por la Corte Suprema, *In re Winship*, 368, 1074—.

²⁶ O'MALLEY *et al.* (eds.), 2000: § 12.10, p. 164.

²⁷ *In re Winship*, 370, 1073 (la cursiva es mía).

²⁸ BROWN *et al.* (eds.), 2006: vol. 2, p. 483.

tarea del abogado es simplemente persuadir, inducir cierto estado mental en los jurados; y «convencimiento» también suena como una cuestión psicológica. No obstante, el lenguaje de los estándares de prueba es también en parte epistemológico: «razonable» en «más allá de toda duda razonable», «clara» en «clara y convincente», parecen cuestiones objetivas dado que aparentemente se refieren a la calidad de las pruebas presentadas. Y el aspecto epistemológico, en mi opinión, es crucial.

Esto no quiere decir que el grado de confianza del juzgador de los hechos en la conclusión sea completamente irrelevante; después de todo, el grado de creencia de una persona razonable estará en proporción, al menos aproximadamente, a la fuerza de las pruebas —esto es, cuanto más avalen las pruebas a *p*, el juzgador de los hechos tendrá mayor confianza en que *p* es verdadero—. Sin embargo, esto es decir que el grado de creencia del juzgador de los hechos es una cuestión distinta y claramente secundaria; mientras *el peso de las pruebas* es lo principal.

Un muestreo de algunas instrucciones al jurado sobre los estándares de prueba confirma esta comprensión epistemológica. Comienzo con el estado de Florida:

— Las instrucciones al jurado por parte de la Corte Suprema de Florida respecto al estándar en los procesos penales contrastan «una convicción consistente de la culpabilidad» con una convicción «vacilante y tímbeante»; sin embargo, está muy claro que «es a las pruebas presentadas en un proceso, y sólo a éstas, a las que se debe acudir para tener como probado algo» y también que «[u]na duda razonable respecto a la culpabilidad del demandado puede surgir de las pruebas, del conflicto entre las pruebas o de la falta de pruebas»²⁹.

— Las instrucciones de la Corte Suprema de Florida sobre el estándar de la preponderancia se centran principalmente en el grado de persuasión: «Cuánto mayor sea “[e]l peso de las pruebas” significa que son más persuasivas y que tienen una mayor fuerza de convicción y un efecto mayor en el caso»³⁰.

— Sin embargo, las instrucciones de la Corte Suprema de Florida sobre «prueba clara y convincente» otra vez se centran principalmente en el aspecto epistemológico: «“prueba clara y convincente” son las pruebas que resultan precisas, explícitas, no confusas y de peso tal que producen una creencia firme o la convicción, sin duda, del caso en cuestión»³¹.

Ahora veamos algunos criterios federales estadounidenses —en éstos encontramos muchas ideas epistemológicas implícitas—:

²⁹ Fla. Std. Jury Inst. (Crim.) (Instrucciones estándares al jurado en casos penales en Florida), 3.7 (2009).

³⁰ Fla. Std. Jury Inst. (Civ.) (Instrucciones estándares al jurado en casos civiles en Florida) 40.1.3 (2010).

³¹ *Ibid.*

— A nivel federal, las instrucciones sobre los estándares de prueba penales comparan la «duda razonable» con la simple duda «posible»: la(s) «prueba(s) más allá de toda duda razonable [...] deben ser pruebas de carácter tan convincente que una persona razonable no dudaría en confiar y actuar conforme a ellas en el más importante de sus asuntos»³². El Primer Circuito ha aprobado la formulación propuesta por el juez KEERTON: «La duda razonable puede surgir no sólo de las pruebas presentadas sino también de la ausencia de prueba. [É]sta existe cuando, después de valorar y considerar todas las pruebas, usando la razón y el sentido común, los jurados no pueden decir que tienen una firme convicción sobre la verdad de los cargos»³³. Por su parte, el Tercer Circuito ha aprobado la siguiente formulación: «La duda debe ser razonable. No es una duda meramente posible o imaginaria [...] Una duda razonable es una duda imparcial, basada en la razón y en el sentido común [...] un acusado nunca debe ser condenado con fundamento en una simple suposición, conjetura o especulación»³⁴. El Sexto Circuito añade que una duda razonable «debe surgir de las pruebas, de la falta de pruebas o de la naturaleza de las pruebas»³⁵; el Octavo Circuito sostiene que la presunción de inocencia «sólo puede superarse si la acusación prueba más allá de toda duda razonable cada elemento esencial del delito imputado»³⁶; y el Noveno Circuito sostiene que el veredicto «debe surgir de una consideración cuidadosa e imparcial de todas las pruebas o de la falta de pruebas»³⁷.

— Las instrucciones al jurado sobre los estándares de prueba en los casos civiles ordinarios en el ámbito federal, citando las instrucciones modelo al jurado del Noveno Circuito, explican que para resolver un proceso por «preponderancia de la prueba» «se debe estar persuadido por las pruebas [...] de que [...] la verdad es más probable que la falsedad»³⁸.

— Y las instrucciones federales al jurado explican «la prueba clara y convincente» como el establecimiento de un estándar mayor que «la preponderancia de la prueba», que exige que el jurado «esté persuadido por las pruebas de que la alegación muy probablemente [...] es verdadera»³⁹.

Ninguno de estos estándares, por cierto, es completamente transparente; pero (con la posible excepción de la instrucción del estado de Florida sobre el estándar de «la preponderancia») todos dejan muy claro que los estándares de prueba no deben entenderse como una mera cuestión psicológica del grado de convicción del jurado, sino como una cuestión principalmente epistemológica, el grado de creencia *avalado por las pruebas*, o la falta de ello.

³² O'MALLEY *et al.*, 2000: § 12.10, p. 161.

³³ *Ibid.*: 165 (citando *U.S. v. Cleveland*, 106 F.3d 1056, 1062-63 (1st Cir. 1997)).

³⁴ *Ibid.*: 171 (citando *U.S. v. Isaac*, 134 F.3d 199, 202 (3rd Cir. 1998)).

³⁵ *Ibid.*: 174.

³⁶ *Ibid.*: 187.

³⁷ *Ibid.*: 189.

³⁸ O'MALLEY *et al.*, 2000: § 101.41, p. 53.

³⁹ *Ibid.*: § 104.03, p. 143.

Si es necesario tener mayor confirmación de lo anterior, puede hacerse analizando las circunstancias en las que un juez puede dictar o revocar un veredicto. El *Manual of Federal Practice* (Manual de la Práctica Federal) indica que un tribunal puede resolver un caso como cuestión de derecho (*Judgment as a Matter of Law*, en adelante, *JMOL*) cuando haya una absoluta ausencia de prueba respecto de una o más de las cuestiones materiales o bien cuando «no haya asuntos de hecho controvertidos sobre los cuales personas razonables pudiesen diferir»; esto es, para considerar la disputa como una cuestión de derecho, analizando la cuestión a la luz más favorable para la parte contra la que se presenta la moción, «sólo puede haber una única conclusión razonable». Sin embargo, si las pruebas están en conflicto o son insuficientes para conducir a una conclusión razonable, el *JMOL* es inapropiado⁴⁰. La regla federal 29 del procedimiento penal estadounidense señala que el juez debe ordenar al jurado que pronuncie un veredicto de absolución si «las pruebas son insuficientes para fundamentar una condena»⁴¹. Además, la Corte Suprema de Florida ha dispuesto que los «[tribunales no deberán conceder una petición de absolución a menos que las pruebas sean tales que ninguna opinión que el jurado pueda legalmente tomar favorable a la parte contraria pudiese ser sostenida en derecho»⁴².

Por otra parte, una comprensión de los grados de prueba fuertemente epistemológica no sólo es más fiel al lenguaje de las instrucciones al jurado, etc., que una puramente psicológica; es también, y como cuestión más importante, integral para aquello que exige el rol que juegan los estándares de prueba en los procesos judiciales. Sin embargo, articular cuál es ese rol exige otras sutilezas.

En algunos procesos (aunque no en todos) la cuestión fundamental es fáctica: ¿la Sra. Coppolino murió por causas naturales o fue envenenada?⁴³; ¿fue el demandado, Robert Downing, u otra persona quien se hizo pasar por el reverendo Claymore para obtener bienes de forma fraudulenta?⁴⁴; ¿fue Ethel Brownstone quien realmente firmó el documento dándole supuestamente los objetos de valor a su sobrina o dicha firma es falsa?⁴⁵; ¿fue la pandilla de Nicola Sacco y Bartolomeo Vanzetti o la de Morelli, la que cometió el robo de la nómina y el asesinato en la fábrica de calzado Slater and Morrill?⁴⁶; ¿fue Janet

⁴⁰ GIVENS y SHREVEY, 2010: § 7.51, pp. 790-791. El «JMOL» cubre los dos veredictos directos (cuando el juez quita el veredicto de las manos del jurado) y las sentencias *n.o.x.* «no obstante el veredicto» (donde un juez anula un veredicto que el jurado ya ha dictado).

⁴¹ WRIGHT y HENNING, 2009: § 467, p. 362 (la cita es tomada de la descripción de los autores, no del texto de la regla).

⁴² ALLEN, 2010: § 18.13, pp. 750-51 (citando *Fitzpatrick v. State*, 900 So.2d 495, 507 (Fla. 2005)).

⁴³ *Coppolino v. State*, 233 So.2d 68 (Fla.2d DCA 1968).

⁴⁴ *U.S. v. Downing*, 753 F.2d 1224 (3rd Cir. 1985).

⁴⁵ *U.S. v. Starzeczyal*, 880 F. Supp. 1027 (SDNY, 1995).

⁴⁶ Véase *infra* apdo. 3.

Collins quien empujó a la anciana en ese callejón y le robó su monedero y fue Malcolm Collins quien condujo el coche de la fuga o fueron otras dos personas completamente diferentes?⁴⁷; ¿cuál fue la causa de los agudos problemas respiratorios de Bob Moore, su exposición al escape de Tolueno o su historial como fumador y enfermo de asma?⁴⁸; ¿fue el consumo de Bendectin durante el embarazo lo que hizo que Jeffrey Blum naciera con pie equino varo o fue sólo mala suerte?⁴⁹, etcétera.

En 1996, la Corte Suprema estadounidense escribió en el caso *Tehan* que «el objetivo de un juicio es determinar la verdad»⁵⁰. Es cierto que un veredicto es sustantivamente justo sólo si la resolución es fácticamente correcta⁵¹; es decir, sólo si el acusado condenado por un delito es genuinamente el mismo o sólo si el demandado declarado responsable del daño es genuinamente el responsable de haberlo causado⁵². Sin embargo, un juicio *no* es como una investigación científica seria o la investigación académica sería que pueden tomarse el tiempo necesario para escudriñar todas las pruebas posibles; las determinaciones jurídicas de los hechos están sujetas a limitaciones de tiempo y a restricciones respecto a la forma de obtención y al tipo de pruebas que pueden legalmente ser presentadas⁵³. Entonces lo que se exige al juzgador de los hechos *no* es, estrictamente hablando, que determine si al acusado es culpable o es responsable, sino que —evitando las especulaciones, los caprichos, los prejuicios y, como lo expuso el Tercer Circuito, «las presuposiciones, las conjeturas, o las especulaciones»— determine si la culpabilidad del acusado o la responsabilidad del demandado *ha sido establecida por las pruebas presentadas al grado exigido*⁵⁴.

Esto obviamente exige que los estándares de prueba sean entendidos, no simplemente en términos del grado de confianza del juzgador de los hechos,

⁴⁷ Véase *infra* apdo. 4.

⁴⁸ *Moore v. Ashland Chemical, Inc.*, 151 F.3d 269 (5th Cir. 1998).

⁴⁹ *Blum v. Merrell Dow Pharmaceuticals, Inc.*, 1 Pa. D. & C.4th 634 (Ct. Com. Pleas Pa. 1988), *rev'd*, 560 A.2d 212 (Pa. Super. Ct. 1989), *aff'd*, 626 A.2d 537 (Pa. 1993), *remanded to* 33 Phil. Co. Rptr. 193 (Ct. Comm. Pleas. Pa. 1996), *rev'd*, 705 A.2d 1314 (Pa. Super. Ct. 1997), *aff'd*, *sub. nom. Blum et al. v. Blum* *Merrell Dow Pharmaceuticals, Inc.*, 747 A.2d 877 (2000).

⁵⁰ *Tehan v. United States*, 383 U.S. 406, 416 (1966).

⁵¹ Por supuesto, la verdad fáctica sólo es necesaria para la justicia sustantiva pero no es suficiente; esto exigirá tener también leyes justas.

⁵² Los tribunales estadounidenses, a veces deliberadamente, han hecho excepciones a este principio en procesos civiles; véase, por ejemplo, *Sindell v. Laboratorios de Abbott*, 26 Cal. 3os 588, 607 P.2d 924, 163 Cal. Rptr. 152 (1980); *Bichter v. Eli Lilly and Co.*, 79 D.2d 317, 436 N.Y.S.2d 623 (1981); *Hymowitz v. Eli Lilly and Co.*, 73 N.Y.2d 487, 539 N.E.2d 1069, 541 N.Y.S.2d 941 (1989).

⁵³ Cuando estaba terminando este trabajo leí a DESKOWITZ (2011) argumentando respecto al juicio de Casey Anthony (quien apenas había sido absuelta por el homicidio de su hija de dos años): «Los científicos buscan la verdad. Los filósofos buscan la moral. Un proceso penal busca sólo [...] la prueba más allá de toda duda razonable».

⁵⁴ Dado que estoy hablando principalmente del derecho estadounidense, esto, estrictamente hablando, debería leerse: «Por las pruebas admisibles presentadas». Pero no seguiré haciendo énfasis en ello; en adelante, deberá entenderse como implícito.

con independencia de si dicho grado de confianza es adecuado dadas las pruebas, sino en términos de lo que es *razonable* creer a la luz de las pruebas presentadas. En resumen, la tarea del juzgador de los hechos es paradigmáticamente epistemológica: dicho con una frase pintoresca pero exactamente adecuada de John Stuart MILL, «*to judge of evidence*» (que significa algo así como «valorar el mérito de las pruebas»)⁵⁵. El juzgador de los hechos debe determinar, como RUSSELL hubiese dicho, si las pruebas presentadas hacen creíble al grado exigido la(s) proposición(es) en cuestión; o, en mi vocabulario epistemológico, si las pruebas presentadas *avalan al grado requerido la(s) proposición(es) en cuestión*⁵⁶.

2. LOS GRADOS DE AVAL NO SON PROBABILIDADES MATEMÁTICAS

Como Ian HACKING observa, el concepto de probabilidad a lo largo de su historia «tiene dos aspectos»: está conectada con el «grado de creencia [avalado] por la evidencia» y con «la tendencia, exhibida por algunos dispositivos de azar, a producir frecuencias relativas estables»⁵⁷.

Cuando se habla de la credibilidad o del aval de una afirmación comúnmente se usa el lenguaje de las probabilidades (*probability* o *likelihood*)—que tan probable es, dadas ciertas pruebas, que la afirmación sea cierta, o, inconducionalmente, qué tan probable es dicha afirmación—. Yo misma hablo de esta manera, por ejemplo, cuando me pregunto qué tan probable es, dadas las pruebas que actualmente tenemos, que exista un nexo causal entre las vacunas y el autismo⁵⁸; o qué tan probable es que Egipto tenga un gobierno genuinamente democrático dentro de cinco años. De acuerdo con el *Oxford English Dictionary*, el sentido usual de «probable» (en inglés británico) es: «Dadas las pruebas existentes es posible esperar razonablemente que algo suceda o que sea el caso»⁵⁹. De modo similar, el *Merriam-Webster's Dictionary* da como definición principal de esta palabra (en inglés norteamericano), «algo suficientemente respaldado por las pruebas para establecer una presunción pero no para tenerlo como probado»⁶⁰.

⁵⁵ MILL, 1843: 5.

⁵⁶ Para una aclaración más completa de esto, véase HAACK, 2004a. Yo no iría tan lejos como Harry Hines, el abogado imaginario de Steve Martini, que cree que «la mayoría de los triunfos en los juzgados penales se deben a la preponderancia del perjurio. Usted tergiversa el suyo y ellos el de ellos y, al final, el lado que tenga más adeptos a la invención, gana»; y que «a lo largo de la historia la verdad se ha marchitado y ha muerto de soledad en la mayoría de los tribunales». MARTINI, 1994: 420. No obstante, no estoy argumentando que el sistema jurídico siempre llega mediante las pruebas presentadas a veredictos avalados al grado exigido.

⁵⁷ HACKING, 1975: 13.

⁵⁸ Una cuestión discutida brevemente en HAACK, 2011.

⁵⁹ *Oxford English Dictionary Online*, <http://www.oed.com/>.

⁶⁰ *Merriam-Webster Dictionary Online*, <http://www.merriam-webster.com/dictionary/probable>.

Estrechamente relacionado con este uso principal de la palabra está nuestro hábito de añadir, en el lenguaje habitual, «probablemente» cuando no queremos comprometernos categóricamente con algo. Cuando un colega me pregunta, «¿estará usted en el seminario mañana?» y yo espero estar allí pero no quiero comprometerme firmemente a ello, mi respuesta puede ser: «probablemente». Con mayor frecuencia usamos «probablemente» y otras expresiones similares cuando *pensamos* que algo es cierto pero no estamos totalmente seguros de ello: por ejemplo, cuando digo que las sobras de comida de la nevera, aunque ya no estén tan frescas como antes, «probablemente» están lo suficientemente bien para comerse. Cuando las pruebas de que *p* no son tan fuertes y reconociendo que *p* podría resultar falsa, usamos la expresión «probablemente» como un modo de limitar nuestro compromiso.

Y por supuesto, «probable», etc., aparecen también en el lenguaje de los estándares jurídicos de prueba: una formulación del estándar «de la preponderancia» es «más probable que su negación», una formulación «de clara y convincente» es «altamente probable», y una formulación «de la sospecha razonable» es una «causa probable». El modo natural y obvio de entender estos usos jurídicos está en el sentido epistemológico habitual: cuán razonable es una afirmación a la luz de las pruebas.

Pero la palabra «probable» y términos relacionados también forman parte del lenguaje característico de lótopatas, estadísticos y achartos. En estos usos, en vez de debilitar nuestro grado de compromiso respecto a una afirmación, «probable» forma parte del contenido de la afirmación misma: hay un 1 por 100 de probabilidades de que un hombre de cincuenta y cinco años muera en el próximo año, una posibilidad del 50 por 100 de que caiga cara en una moneda, un aumento del riesgo en más del doble de sufrir el síndrome Guillain-Barré entre aquellos que recientemente se han vacunado contra la gripe porcina, etc.⁶¹ Este es el lenguaje regulado por el cálculo de probabilidades estándar e interpretado en términos de frecuencias relativas o propensiones.

Muchos han buscado hacer calzar los grados de prueba en el derecho dentro de este molde más formal. LEBNIZ, por ejemplo, llama a la teoría de la probabilidad «*iurispudencia natural*»⁶²; y George BOOLE intentó aplicarla «a la valoración de la probabilidad de las decisiones judiciales»⁶³. Más recientemente, el probabilismo jurídico ha sido objeto de un largo debate entre los

⁶¹ Mi último ejemplo alude a la historia que lleva a la adopción del criterio «incremento del riesgo en más del doble» para la causalidad específica en casos de responsabilidad por daños tóxicos en algunos tribunales estadounidenses. Véase HAACK, próximamente.

⁶² JOHN LOCKE, sin embargo, ha negado que sean posibles reglas precisas para calcular los grados jurídicos de prueba. Mi referencia al respecto es HAACK, 1975: 167.

⁶³ BOOLE, 1854: capítulo XXI, p. 376. En la p. 382, sin embargo, BOOLE reconoce que: «[D]e los muchos expedientes de las decisiones de un tribunal... no es posible deducir ninguna conclusión definitiva respecto a la corrección de los juicios individuales de sus miembros».

teóricos que se autodescriben como *the New Evidence Scholarship*⁶⁴ —un debate que resurgió en el año 2010 como parte de un intercambio académico provocado por la reciente propuesta que sólo cambió la forma pero no el contenido de las *Federal Rules of Evidence* estadounidenses (FRE, Reglas Federales sobre la Prueba)⁶⁵—. Esta vez, al parecer, fue la definición de relevancia que aparece en la FRE 401 lo que desencadenó el debate: una prueba es relevante si *incrementa o disminuye* la probabilidad de algún hecho en cuestión. Sin embargo, es evidente que no se debe *asumir*, solamente porque aparezcan palabras como «probable» o «probabilidad» en contextos jurídicos, que se trata de probabilidades matemáticas y no de probabilidades epistemológicas.

Algunos críticos del probabilismo jurídico han expresado su escepticismo sobre la posibilidad de asignar valores numéricos a los grados de prueba. Thomas STARKIE escribió en 1842 «[q]ue [...] las probabilidades morales [...] puedan alguna vez representarse por medio de números [...] y así ser objeto de un análisis numérico sólo puede considerarse como ilusorio y quimérico»⁶⁶. Otros han sostenido que el cálculo preciso de probabilidades, incluso si fuese posible, es generalmente indeseable: Lawrence TRIBE escribe sobre «*the costs of precision*» (los costos de la precisión), como «el eclipse de las variables endebles»; e indica que en los procesos penales el enfoque bayesiano puede fomentar una «presunción de culpabilidad»⁶⁷. Estoy de acuerdo: no es factible asignar números precisos a los grados de prueba y tampoco sería necesariamente deseable hacerlo, incluso si esto fuese posible. No obstante, mi objeción al probabilismo jurídico es más fundamental: creo que identificar los grados de prueba en el derecho, es decir, grados de aval epistémico, con probabilidades matemáticas es un tipo de falacia de ambigüedad. En esto estoy de acuerdo con Richard VON MISES, quien hace mucho tiempo sostuvo que «la teoría de probabilidad no tiene nada que ver con preguntas tales como: ¿hay alguna probabilidad de que en un futuro Alemania esté en guerra con Libia?»⁶⁸.

Habría cierta superposición entre mis argumentos y los argumentos de otros críticos del probabilismo jurídico, como L. Jonathan COHEN⁶⁹ (desde el ámbito filosófico) y Leonard JAFFEE⁷⁰ y Ronald ALLEN⁷¹ (desde el ámbito jur-

⁶⁴ KAPLAN (1968) parece haber sido de gran influencia. Véase también, por ejemplo, LEWERT, 1983; PARKS y SAKS, 2005-2006.

⁶⁵ RISINGER, *et al.*, 2010 (aparentemente un prolongado intercambio de correos electrónicos entre los «autores»). Para la propuesta «reforma de forma pero no de fondo» de las FRE, véase ADVISORY COMMITTEE ON EVIDENCE RULES, 2009.

⁶⁶ STARKIE, 1842: vol. 1, p. 579.

⁶⁷ TRIBE, 1971: 1358 (los costos de la precisión). (Este trabajo fue en parte una respuesta a FINKELSTEIN y FARLEY, 1969-1970, discutido aquí en detalle en el apartado 4).

⁶⁸ VON MISES, 1928: 9.

⁶⁹ COHEN, 1977.

⁷⁰ JAFFEE, 1984-1985.

⁷¹ ALLEN, 1986.

dico)⁷². Sin embargo, yo no apelaré a ningún tipo de «lógica inductiva» ni cuasi-baconiana ni otra; dado que creo que la moraleja de la paradoja del «verdul» es que *no puede ser* una lógica inductiva formal⁷³. Y si bien hay cierta remota afinidad entre el concepto «integración explicativa» que desempeña un papel en mi explicación epistemológica y «las historias explicativas» que desempeñan un papel en la «reconceptualización de los procesos civiles» propuesta por ALLEN⁷⁴, estos no son conceptos exactamente iguales, tienen roles diferentes en las dos teorías —dos teorías que tienen también propósitos diferentes—.

Mis argumentos tendrán un carácter específicamente epistemológico: que los usos epistemológicos y estadísticos de «probables» son relevantemente diferentes; y además, que las probabilidades epistemológicas —es decir, los grados de credibilidad racional o de aval— no siguen los axiomas de KOLMOGOROV sobre el cálculo estándar de probabilidades. Así pues necesito empezar delineando los contornos principales de mi concepto de aval⁷⁵.

De forma breve y muy aproximadamente: cuán avalada está una afirmación depende de la calidad de las pruebas que hay respecto a esa afirmación. De forma un poco más precisa: las pruebas incluyen tanto las pruebas sensoriales como las razones⁷⁶; y éstas trabajan juntas como las pistas y las entradas ya completadas en un crucigrama. Crucialmente importante para mis propósitos aquí es que las razones se ramifican, como las entradas de un crucigrama; y que aquello que hace que las pruebas con respecto a una afirmación sean mejores o peores es análogo a aquello que hace que una entrada en un crucigrama sea más o menos razonable:

⁷² Quiero dejar constancia de que mi enfoque epistemológico fue desarrollado mucho antes de que leyera a COHEN o a ALLEN.

⁷³ Sobre la paradoja del «verdul», véase GOODMAN, 1954; sobre sus consecuencias para la posibilidad de una lógica inductiva sintácticamente caracterizable, véase HAACK, 2003: 84-86. Debería decirse, sin embargo, que —aunque COHEN no haga ninguna referencia a «verdul», y su única referencia a GOODMAN (COHEN, 1977: 184) no sea relevante para este punto— uno *podría* encontrar, leyendo entre líneas su discusión sobre las «variables relevantes» (pp. 133 y ss.) un reconocimiento implícito de que el grado de apoyo no es, después de todo, totalmente formal, sino que depende en parte de hechos materiales. Pero, lamentablemente, en lugar de buscar las consecuencias de este punto clave, COHEN ofrece «una sintaxis lógica de la probabilidad inductiva», presuponiendo una lista de variables pertinentes (capítulo 17).

⁷⁴ ALLEN, 1986: § III; LEMPERT (1988: 84 y ss.) resume la concepción de ALLEN. Como su título lo indica, la aproximación de ALLEN pretende ser aplicada solo en procesos civiles (ALLEN, 1986: 436) —lo que por sí mismo revela que su preocupación es más jurídica que, como la mía, epistemológica: como lo hace el hecho de que no da ninguna explicación sobre qué hace a una historia mejor o peor—. Además, dado que confía en unos juicios comparativos de las historias explicativas del demandante y del demandado, la concepción de ALLEN es —otra vez, como lo indica su título— una propuesta *revivista*, aparentemente imponiendo una carga sobre los demandados en los casos civiles mayor de la que actualmente tienen. [ALLEN reconoce la posible objeción de que su aproximación descansa en la diferencia entre «una simple denegación del caso del demandante y una defensa afirmativa del caso por parte del demandado» (p. 427); y responde (p. 428) que él la ve, no como una diferencia categórica, sino gradual].

⁷⁵ Por supuesto, éste *sólo* puede ser un esbozo. Para más detalles, véase HAACK, 1993; 2009; capítulo 4, y 2003: capítulo 3.

⁷⁶ El carácter exacto y el rol de las pruebas provenientes de la experiencia generan preguntas filosóficas desafiantes; pero en este contexto pueden dejarse de lado como sólo marginalmente relevantes.

i) Qué grado de apoyo ofrecen las pruebas; análogamente: cuán bien encaja una entrada de un crucigrama con la pista dada y las entradas ya completadas. Las pruebas pueden dar algún grado de apoyo (ser positivas, favorables) a determinada conclusión, pueden *socavarla* (ser negativas, desfavorables), o ser *neutrales* (irrelevantes).

ii) Cuán seguras son las razones con independencia de la afirmación en cuestión; análogamente: cuán razonables son las otras entradas entrecruzadas que ya se han completado con esta nueva entrada, independientemente de la palabra en cuestión. Cuanto mayor sea la seguridad independiente de las razones *positivas*, la conclusión estará *más* avalada; pero, cuanto mayor sea la seguridad independiente de las razones *negativas*, la conclusión estará *menos* avalada.

iii) Cuán *inclusivas* son las pruebas, es decir: cuántas de las pruebas relevantes tenemos; análogamente: qué proporción del crucigrama ha sido completado⁷⁷. Las pruebas con un mayor grado de inclusión dan más aval a una conclusión que las pruebas con menor grado de inclusión, si y sólo si, las pruebas adicionales son *al menos tan favorables* como las demás.

Dicho brevemente y de manera aproximada, el grado de apoyo otorgado por las pruebas depende de la aportación que éstas hagan para la integración explicativa de las pruebas-más-la-conclusión —o, aun más breve y aproximado— de cuán bien encajen las pruebas y la conclusión juntas en un relato explicativo. De acuerdo con esto, si un elemento de prueba concreto apoya a la conclusión y en qué grado lo hace dependerá de si éste contribuye a la integración explicativa como un todo y en qué grado lo hace⁷⁸. Pensemos en la entrada de un crucigrama que está bien apoyada por la pista y por las otras entradas, pero donde una de estas otras entradas específicamente —quizás aquella que tiene una «x» en medio— es especialmente significativa, porque sin ella el encaje total de las entradas sería mucho más flojo: lo análogo es un elemento de prueba concreto cuyo apoyo es crucial para cierta conclusión. Puede valer la pena añadir que, por esta razón, mi aproximación no es atomística pero tampoco totalmente holística; sino una especie del holismo articulado.

Esta teoría es *fundamentarista*: con ello quiero decir que, como el fundamentalismo y a diferencia del coherentismo, admite el rol de las pruebas sensoriales; y, al mismo tiempo, como el coherentismo, pero a diferencia del funda-

⁷⁷ La idea de la inclusividad parece estar implícita, por ejemplo, en la instrucción del Sexto Circuito respecto a que la duda razonable puede surgir no sólo de las pruebas, sino de la *falta* de pruebas.

⁷⁸ Mi concepto de integración explicativa difiere del concepto más familiar «de la inferencia a la mejor explicación» en dos maneras: no es, como «la inferencia a la mejor explicación», optimizador sino gradual; y no es, como «la inferencia a la mejor explicación», unidireccional sino multidireccional, concibiendo al *explorans* y al *explanzandum* como mutuamente explicativos. Y quizá, para prevenir cualquier malentendido, debería reiterar que mi concepto de integración explicativa juega su rol en la explicación del grado de apoyo —que es, para recordarlo, solo una dimensión del grado de aval—.

cionalismo, admite relaciones de apoyo mutuo omnipresentes entre nuestras creencias⁷⁹. Es también *material*: es decir, reconoce que lo que determina la calidad de las pruebas y, por tanto, del aval, no es totalmente formal sino también en parte material⁸⁰. Que una prueba sea relevante para una afirmación puede depender de hechos acerca del mundo: por ejemplo, si para determinar su culpabilidad es relevante que la herida del acusado haya sanado adecuadamente, como presuponia la práctica de los juicios mediante hierro candente, depende de si es verdad que Dios protegerá a aquel que ha sido acusado falsamente; si el efecto que cierta medicina tiene en las ratas es relevante para sus efectos en las personas, depende de si es verdad que las ratas son fisiológicamente similares a las personas en los aspectos importantes, etcétera.

Como esto pone de manifiesto, la explicación de relevancia de la prueba dada en las FRE (ya sea en su versión antigua o en la más nueva) —como una prueba que incrementa o disminuye la probabilidad de un hecho en cuestión— deja mucho que desear. Un elemento probatorio es relevante si y sólo si influye en el grado de soporte de las pruebas en global; es decir, si y sólo si su presencia contribuye o debilita la integración explicativa de las —pruebas—la-conclusión—lo cual, como hemos visto, dependerá de los hechos del mundo. Entonces la definición jurídica es potencialmente engañosa de dos maneras: abiertamente, sugiriendo que la relevancia es una cuestión que depende de si las pruebas cambian la probabilidad matemática de un hecho, y no de si éstas cambian su estatus epistémico; y encubiertamente, por su silencio sobre el carácter material de la relevancia.

Debería estar ya claro que si los conceptos de calidad de las pruebas y de grado de aval son tan complejos, tan sutiles, tan multidimensionales y tan mundanos como esta concepción sugiere, la teoría matemática de las probabilidades no puede de ninguna manera constituir, por sí misma, una teoría del aval. Sin embargo, esto aún no es suficiente para mostrar que si tenemos una adecuada teoría del aval, entonces el cálculo de probabilidades no podría servir para calcular los grados de aval. Pero no creo que el cálculo de probabilidades podría hacer esto, por al menos tres razones:

— Dado que la calidad de las pruebas tiene varias dimensiones diferentes (grado de apoyo, seguridad independiente e inclusión) y que no hay forma de ordenar el relativo éxito o fracaso a través de estos diferentes factores, no hay garantía incluso de un orden lineal de los grados de aval⁸¹.

— Mientras que la probabilidad de p y la probabilidad de $no-p$ debe ser igual a 1, cuando no tenemos prueba alguna o solamente tenemos pruebas

⁷⁹ El fundacionalismo y el coherentismo son tradicionalmente las dos teorías epistemológicas rivales en esta área; pero, como lo he demostrado, no son exhaustivas. Para más detalle, ver Haack, 1993, y 2009; capítulo 1, para un tratamiento más sumario, Haack, 1998.

⁸⁰ Todo esto se explica con mayor detalle en Haack, 2003, 76-77.

⁸¹ Este punto también ha sido resaltado en KEYNES, 1921: 27-28.

may débiles, en ambas direcciones, ni p ni $no-p$ pueden estar avaladas en ningún grado.

— Mientras que la probabilidad de p y q (para una p y una q independientes) es el producto de las dos y, por tanto, la probabilidad de su conjunción será menor a la probabilidad de cualquiera de ellas, a menos que las dos sean 1, el grado de aval de su conjunción puede ser *más alto* que el grado de aval de sus componentes⁸².

El segundo y el tercer punto tienen una clara relevancia para la prueba judicial.

El estándar de la «preponderancia de la prueba» no puede ser entendido de forma adecuada simplemente en términos de qué parte produce *más* pruebas; no se trata de contar el número de testigos presentados o de sopesar el peso de los documentos presentados por cada parte. Pero la forma alternativa de palabras «más probable que su negación» necesita un tratamiento cuidadoso. No es suficiente que una parte produzca mejores pruebas que la otra; sino que es necesario que la parte que tiene la carga de la prueba produzca pruebas *lo suficientemente buenas para avalar la conclusión al grado exigido*⁸³. De esta manera, exactamente como el segundo punto señalado nos llevaría a esperar, incluso el estándar de la preponderancia —que parece ser el más acorde con una interpretación probabilística, dada su formulación «más probable que su negación»— de hecho, se resiste.

Con respecto al tercer punto, vale la pena anotar, primeramente, que si los grados de prueba jurídica *fuessen* probabilidades matemáticas, las instrucciones dadas al jurado respecto a que cada elemento del caso debe establecerse al grado exigido no son suficientes para alcanzar el objetivo; como es obvio cuando uno se da cuenta de que —suponiendo por mor del argumento (aunque *sólo* por mor del argumento) que probar más allá de toda duda razonable exige una probabilidad matemática de 0,95⁸⁴— las pruebas que estableciesen a este grado cada uno de los tres elementos independientes sólo tendrían una

⁸² El punto negativo, que la aplicación de la teoría de la probabilidad convencional tropezaba aquí con problemas sobre la negación y la conjunción, fue planteado por COHEN. Pero, como espero resulte obvio, mi punto positivo respecto a que una adecuada aproximación al aval produce los resultados correctos en ambas cuestiones, no fue anticipado por COHEN.

⁸³ PABDO (2013:111-112) sugiere que si interpretamos probabilísticamente al estándar «de la preponderancia» como «más [matemáticamente] probable que no», esto exigirá un veredicto contra el demandante aun cuando, por ejemplo, la probabilidad de su teoría sobre lo que pasó sea 0,4, mientras la probabilidad de la teoría del demandado sólo es 0,2. Pero esto no es realista, como PABDO sugiere, un argumento en contra del probabilismo jurídico: es un argumento para reconstruir el estándar de la preponderancia (como ALLEN propone; véase la nota 75 *supra*) no para exigir que el demandante produzca pruebas para hacer su pretensión más probable, sino que sólo produzca pruebas para hacer que su pretensión parezca más probable que la historia del demandado. Si «probable», en este contexto, debe entenderse en términos de probabilidades matemáticas o en términos de grados de aval, es una pregunta completamente independiente.

⁸⁴ POSNER (2001: 367) reporta que las estimaciones de los jueces federales respecto a qué grado de probabilidad representa el «más allá de toda duda razonable», oscila entre el 75 y el 95 por 100.

probabilidad conjunta de 0,85, muy por debajo del umbral (y las pruebas que estableciesen cada uno de los tres elementos independientes en una probabilidad de 0,51, sólo tendrían una probabilidad en conjunto de 0,13).

De hecho, una agregación de pruebas puede avalar una conclusión en un grado *mayor* que el que pudiera avalar cualquiera de sus componentes por sí mismo. He sostenido este argumento en otro trabajo, en el contexto de los conflictos jurídicos en casos de daños por sustancias tóxicas, sobre la llamada «metodología del peso de las pruebas»⁸⁵. Dicho de forma breve y muy general: la combinación de las pruebas aumentará el grado de aval de una conclusión solamente si las pruebas en conjunto aumentan el grado de apoyo a la conclusión; y/o mejora la seguridad independiente de los elementos de prueba que la componen; y/o introducen nuevas pruebas que no sean menos favorables para la conclusión que una agregación más restringida de pruebas. Esto no sucederá invariablemente: la cuestión no es que la combinación de pruebas *siempre* es más fuerte que cualquiera de sus componentes sino que *a veces* lo es. Por ejemplo, una prueba epidemiológica sobre un incremento del riesgo mayor de D entre aquellos que han estado expuestos a S, puede o no entrelazarse con otros estudios. Si también sabemos que la misma sustancia, S, en dosis equivalentes, produce el mismo tipo de daño, D, en animales que se sabe son fisiológicamente similares en aspectos relevantes a los seres humanos y que la toxicología ha demostrado que los componentes químicos que son muy similares a S también están asociados a D, entonces la combinación de pruebas nos daría mayor aval para pensar que S causa D, a diferencia de lo que haría cualquiera de los elementos probatorios por sí mismos. Del mismo modo, la prueba de que el demandado tenía fuertes motivos para desear la muerte de la víctima puede entrelazarse con las pruebas que demuestran que poseía un arma del mismo calibre y con las pruebas de haber sido visto escapando del lugar del crimen, dando mayor aval a la conclusión de que él cometió el crimen, en comparación con la que cualquiera de los elementos de prueba por sí mismos podrían ofrecer.

Si, como he sostenido, los grados de prueba jurídica se explican mejor como grados de credibilidad racional o de aval y si, como también he sostenido, los grados de credibilidad racional o de aval no pueden identificarse con las probabilidades matemáticas, entonces el probabilismo jurídico es incorrecto. De todos modos, tal como decimos en inglés, «*the proof of the pudding is*

⁸⁵ Véase, por ejemplo, *Oxendine v. Merrill Dow Pharmaceuticals*, 506 2d 1100, 1108 (D. C. 1986) («[El doctor Done] reconoció su inhabilidad para concluir que el Bendectin era teratogénico sobre la base de cualquiera de los estudios individuales que discutí, pero también dejó muy claro que todos estos estudios deben analizarse en conjunto y que, vistos así, apoyan su conclusión»). *Gen. Elec. Co. v. Joiner*, 522 Estados Unidos 136, 153, 118 S.Ct. 512, 522 (1997) (El magistrado Stevens formuló un voto particular disintiendo) («No es intrínsecamente "no científico" para profesionales con experiencia llegar a una conclusión pesando todas las pruebas disponibles»). El argumento aquí expuesto es desarrollado en detalle en HAACK, 2008.

in the eating»⁸⁶. Mostrar simplemente que el probabilismo jurídico es erróneo no nos permite, en sí mismo, comprender mejor pruebas complejas. Entonces, el siguiente paso consistirá en poner en práctica mi enfoque epistemológico; y mostrar sus ventajas respecto a otros enfoques sobre la prueba manifestamente probabilísticos —bayesianos subjetivos— en dos casos famosos, o quizá, mejor dicho, tristemente celebres: *Commonwealth v. Sacco and Vanzetti*⁸⁷ y *People v. Collins*⁸⁸.

3. EL CASO COMMONWEALTH V. SACCO AND VANZETTI: EL FUNDAMENTALISMO FUNCIONA MEJOR QUE EL BAYESIANISMO SUBJETIVO

Los defensores de lo que quizá es la forma más desafiante de probabilismo jurídico, el bayesianismo subjetivo⁸⁹, podrían objetar que mi comparación entre un enfoque probabilístico y uno epistemológico es falso. Insistirán en que su enfoque probabilista es epistemológico. Así, se argumentará que dado que bajo este enfoque, las probabilidades se entienden subjetivamente como grados de creencia; los grados de creencia que una persona otorga a diversas proposiciones son racionales si son consistentes; y son consistentes si satisfacen los axiomas del cálculo de probabilidades. Entonces, el teorema de Bayes proporciona una forma de ajustar («actualizar») los grados de creencias o probabilidades previas que un sujeto hace al considerar nuevas pruebas. De este modo, el bayesianismo subjetivista concluye —dado que los estándares de prueba jurídicos exigen, precisamente, que el juzgador de los hechos tenga cierto grado de creencia racional en la conclusión y dado que el grado de creencia bajo la coherencia probabilista es un grado de creencia racional— que los grados de prueba en el derecho pueden identificarse con los grados de creencia racional, entendidos tal como éste propone.

Éste es el trasfondo teórico del análisis de Jay Kadane y David Schum sobre las pruebas presentadas en el juicio oral de Nicola Sacco y Bartolemeo

⁸⁶ Es decir, «sólo se sabe la calidad del pudín cuando se come». Este proverbio anglosajón revela que la palabra «*proof*» —que tiene la misma raíz que «*probable*»— originalmente significa «*test*» (como «*pruebas*»).

⁸⁷ Para ver las transcripciones del juicio y los procedimientos subsiguientes, véase ANONIMO, 1969.

⁸⁸ *People v. Collins*, 438 P.2d 33 (1968).

⁸⁹ La historia aquí es un tanto enredosa. El teorema de Bayes es un controvertido teorema del cálculo de probabilidades (debe su nombre al reverendo Thomas BAYES, quien lo demostró). Sin embargo, BAYES no publicó dicha prueba, que apareció en las *Philosophical Transactions of the Royal Academy* después de su muerte en 1763, gracias a su albacea Richard PRICE; y fue PRICE, y no BAYES, quien sugirió que el teorema podría tener implicaciones más amplias (SCHUM, 1994: 48-49). Lo que ahora es llamado «bayesianismo subjetivo» implica una interpretación específica de la probabilidad (como grados de creencia) y un uso específico del teorema de Bayes (como un medio para ajustar los grados de creencia de alguien a la luz de nuevas pruebas); y es, a decir lo menos, no incontrovertido.

Vanzetti⁹⁰, dos inmigrantes italianos acusados de un homicidio cometido durante el robo de una nómina en South Braintree, Massachusetts en 1920. La cuestión central del juicio era fáctica —si Sacco y Vanzetti eran realmente los delincuentes—. Las pruebas presentadas fueron, cuanto menos, complicadas; declararon ciento cincuenta y ocho testigos —la acusación presentó noventa y nueve y la defensa cincuenta y nueve—: algunos afirmando haber visto a Sacco y Vanzetti en la escena del crimen, otros sobre la fiabilidad de las coartadas dadas, o sobre la probabilidad de que la bala fatal hubiese salido del arma de Sacco, otros sobre una gorra que supuestamente pertenecía a Sacco y que fue encontrada en la escena del crimen, o sobre el coche en el que huyeron, o sobre las subsecuentes acciones de los acusados —que supuestamente revelaban cierta conciencia de ser culpables—, etc. El fallo del jurado fue que Sacco y Vanzetti eran culpables y fueron condenados a muerte. Después de numerosas objeciones y apelaciones —ninguna de las cuales prosperó, ni siquiera una fundada en la confesión de Celestino Maderios afirmando que él y otros miembros de la pandilla Morelli habían cometido el delito— Sacco y Vanzetti fueron ejecutados en agosto de 1927.⁹¹

Sacco y Vanzetti eran anarquistas; y en la época en que tuvo lugar el juicio, después de la revolución rusa, en Estados Unidos predominaba el «Temor Rojo». En 1918, cuando el país estaba en guerra, el Congreso había aprobado la *Sedition Act* (Ley de Sedición), tipificando como delito federal «el haber voluntariamente expresado de forma oral, haber estampado, haber escrito o publicado cualquier tipo de palabras desleales, profanas, groseras o impropias sobre la forma de gobierno de los Estados Unidos [...]»⁹²; y muchos pensaron que Sacco y Vanzetti habían sido señalados como culpables por su ideología política. En 1927, el año que fueron ejecutados, Felix FRANKFURTER publicó un breve libro señalando muchas deficiencias tanto en las pruebas que fundaron su condena como en las decisiones judiciales⁹³; el día de la ejecución, Upton SINCLAIR comenzó a escribir *Boston*, una apasionante novela basada en este caso⁹⁴; al año siguiente, Edna St. Vincent MILLAY escribió el poema «*Justice Denied in Massachusetts*» («Justicia denegada en Massachusetts») en memoria de los dos hombres⁹⁵; en el cincuenta aniversario de su ejecución, el entonces gobernador de Massachusetts, Michael Dukakis, emitió una declaración reconociendo que «la atmósfera de los juicios [seguidos contra Sacco y Vanzetti] y sus apelaciones estuvo impregnada

por prejuicios contra extranjeros y por la hostilidad hacia ideologías políticas no ortodoxas», y declaró el 23 de agosto de 1977 como «día conmemorativo de Nicola Sacco y Bartolomeo Vanzetti»⁹⁶. El caso sigue generando controversias hasta hoy. Por ello, si KADANE y SCHUM hubiesen aclarado exitosamente el análisis de las pruebas respectivas, esto tendría un interés más allá del ámbito técnico.

Sin embargo, el análisis de KADANE y SCHUM en realidad es más confuso que aclarador. Por un lado, su título, prometiendo un «análisis probabilístico» de las pruebas del caso, es engañoso. Es cierto que ofrecen un análisis desglosado de las pruebas en sus partes componentes (representado por medio de los diagramas de WIGMORE)⁹⁷; sin embargo, lo hacen completamente de forma manual —es decir, la teoría de las probabilidades no juega ningún papel en su análisis de las pruebas, solamente en la *síntesis* que hacen para llegar a ciertas conclusiones—⁹⁸. Por otro lado, aunque KADANE y SCHUM son conscientes de las objeciones que COHEN planteó contra la posibilidad de interpretar los grados de prueba como probabilidades matemáticas estándar, no ofrecen ninguna respuesta a éstas⁹⁹; por el contrario, luego de observar que las FRE definen como prueba relevante a aquellas pruebas que aumentan o disminuyen la probabilidad de algún hecho en cuestión¹⁰⁰, simplemente dan por supuesto que los grados de prueba en el derecho son probabilidades matemáticas. Además, no ofrecen conclusiones categóricas, sino únicamente diferentes probabilidades posteriores, de que Sacco o Vanzetti estuvieron implicados en el delito, asignando diferentes probabilidades previas a distintos elementos probatorios y varias asignaciones de probabilidades condicionales. Después de cientos de páginas de diagramas y cálculos esto es, como mínimo, decepcionante. Pero tienen problemas mucho peores.

Primero: el hecho de que KADANE y SCHUM ofrezcan toda una gama de conclusiones muy diferentes, *todas* ellas probabilísticamente consistentes, revela que la consistencia probabilística no es suficiente para garantizar grados racionales o razonables de creencias¹⁰¹ —lo que seguramente también

⁹⁰ Esta Proclamación aparece reimpresa en TOPP, 2005: 182-184.

⁹¹ WIGMORE, 1913; para una aproximación resumida, véase GODWIN, 2000.

⁹² En uno de sus libros anteriores, SCHUM había aludido a conceptos como la falta de pruebas, la relevancia y la credibilidad; pero, hasta donde yo puedo ver, no desarrolló una teoría del aval ni algo cercano a ésta. SCHUM, 1994: 15, 29 (sobre la falta de pruebas), 69 y ss. (sobre la relevancia) y 207 (sobre la credibilidad).

⁹³ KADANE y SCHUM, 1996: 239-240. Antes, SCHUM había discutido el relato de COHEN con cierto detalle (SCHUM, 1979 y 1994); pero tampoco encontré ahí respuestas a las objeciones de COHEN —sólo una sugerencia, en la parte de 1979— que los beneficios del bayesianismo sean de alguna manera más significativos que el problema «de las paradojas» de la negación, la conjunción, etcétera.

⁹⁴ *Ibid.*: 50. SCHUM (1994: 71-73) mantiene que los juicios de relevancia que hacen las personas dependen de su «punto de vista»; pero a veces parece desplazar esta afirmación hacia una muy diferente, que la relevancia es si misma es relativa al punto de vista.

¹⁰¹ *Qua fundamenteris*, por supuesto no creo que cualquier clase de consistencia o coherencia, por sí sola, sea suficiente como aval. Véase HAACK, 1993/2009: capítulo 3; y 2004b.

⁹⁰ KADANE y SCHUM, 1996.

⁹¹ Para una breve historia, véase TOPP, 2005: «Introduction», pp. 1-51, y «Chronology of Events Related to the Sacco and Vanzetti Case», pp. 185-188. Maderios fue ejecutado el mismo día pero por un asesinato diferente. *Ibid.*: 188.

⁹² 40 Stat. 553 1917-19, chap. 75 (Entramada a la *Espionage Act* de 1917). Dicha ley fue abrogada en 1921, 41. Stat. 1359 1914-1921, chap. 136.

⁹³ FRANKFURTER, 1927.

⁹⁴ SINCLAIR, 1928.

⁹⁵ MILLAY, 1928.

exigiría, si vamos a hablar en estos términos, grados razonables de creencias previas y grados razonables de creencias condicionales—.

Segundo: KADANE y SCHUM reconocen que las probabilidades que usan no pueden ser entendidas en términos de la teoría del azar y tampoco en términos estadísticos; más bien, sostienen, éstas son «personales, subjetivas, o [lo que supongo que quisieron decir, fue “es decir”] probabilidades epistémicas»¹⁰². Pero «personal» no es lo mismo que «subjetivo» y tampoco es lo mismo que «epistémico»; y la manipulación verbal aquí es sintomática de una cuestión más profunda. Supongamos por mor del argumento (pero sólo por mor del argumento) que las probabilidades que KADANE y SCHUM calculan son grados de creencia: ¿de quién son estos grados de creencia? Los suyos, responden KADANE y SCHUM, y los de otros estudiosos¹⁰³. Como esto revela, KADANE y SCHUM se apoyan en el juicio epistémico de quienes han estudiado el caso. La supuesta identificación de probabilidades con los grados de creencias (subjetivas) de estas personas es espuria; en realidad se trata de las *opiniones de los expertos sobre las probabilidades* [*likelihoods*] *epistémicas* (*objektivas*).

Y entonces, tercero: las muchas páginas de cálculos de KADANE y SCHUM pueden dar la impresión de que el cálculo de probabilidades hace un trabajo epistemológico muy serio; pero esto es una ilusión. La única epistemología aquí es el desglose totalmente informal que hacen KADANE y SCHUM de las pruebas en los elementos que la componen, así como su propia valoración de dichas pruebas y la de otros teóricos, también totalmente informal. Las matemáticas, cuando no son francamente engañosas, son en gran parte decorativas.

A fin de mostrar que mi concepción puede hacerlo mejor —en lugar de ofrecer un extenso análisis de todas las pruebas— seguiré el ejemplo de la discusión planteada por FRANKFURTER sobre los diferentes elementos de pruebas en el caso. Como mencioné anteriormente, «probabilidad» tiene usos tanto epistemológicos como matemático-estadísticos; y cuando FRANKFURTER escribe que «[t]oda probabilidad razonable apunta en dirección contraria a Sacco y Vanzetti [y] toda probabilidad razonable apunta hacia la pandilla de Morelli»¹⁰⁴, es muy claro que su intención es usar la expresión «probabilidad» en sentido epistemológico. Además, como veremos, sus observaciones epistemológicas sobre las debilidades de las pruebas tienen —si se me perdona el anacronismo¹⁰⁵— un patrón distintivamente fundaherentista.

¹⁰² KADANE y SCHUM, 1996: 24, 120, 159. Ellos incluso han escrito que: «[L]o que la no-independencia, significa que el conocimiento de un elemento de prueba *puede influenciar nuestro juicio* sobre la fuerza probatoria de los otros elementos». *Ibid.*: 129 (la cursiva es mía).

¹⁰³ *Ibid.*: 239-240.

¹⁰⁴ FRANKFURTER, 1927: 101.

¹⁰⁵ Debería añadir que sólo el término es anacrónico; porque mi fundaherentismo estuvo desde el inicio planteado en parte como una articulación de los estándares implícitos en nuestras evaluaciones diarias sobre qué pruebas son mejores y cuáles son peores y qué creencias están más o menos avaladas.

FRANKFURTER señala, por ejemplo, los numerosos defectos de los testimonios que ubicaban a Sacco en la escena del crimen. En el proceso, un año después de que ocurrieran el robo y el asesinato, Mary Splaine declaró con gran seguridad que había visto a Sacco —a quien describió con gran detalle (aunque no del todo de forma exacta)— en el coche en que huyeron los delincuentes. Ella sólo había visto a un hombre que podría identificar como Sacco —quien supuestamente estaba en un coche que viajaba a una velocidad entre 25 y 30 kilómetros por hora— a una distancia de 20 a 25 metros y en un lapso de tan sólo uno y medio a tres segundos. Enfatizó el hecho de que el hombre que había visto tenía «la mano izquierda de gran tamaño, una mano que denotaba fuerza»; pero las manos de Sacco eran más pequeñas que las manos promedio. Además, anteriormente había dicho a la policía que *no podría* identificar a la persona que vio e incluso en otra ocasión había identificado a otro hombre que luego resultó haber estado en la cárcel cuando el delito se cometió; y finalmente eligió a Sacco después de haberlo visto varias veces en la comisaría y en el juzgado¹⁰⁶. Otro testigo ocular, Frances Devlin, quien también declaró haber visto a Sacco en el coche, poco antes había dicho que no podría identificarlo positivamente¹⁰⁷. Louis Pelzer afirmó también haber visto a Sacco; pero sus compañeros del trabajo declararon que él había estado escondido debajo de una banca en el momento de los disparos¹⁰⁸. Otros dos testigos oculares, Ferguson y Pierce, que presenciaron los hechos delictivos desde un piso superior al que Splaine y Devlin estaban, señalaron que desde allí era imposible hacer cualquier tipo de identificación¹⁰⁹, etc. En resumen, yo diría que:

aun si fuera cierto que Splaine, Pelzer y Devlin vieron a Sacco en el coche de la huida, lo que supondría un fuerte grado de apoyo para la afirmación de que Sacco estuvo implicado, el testimonio de estos testigos oculares desafortunadamente carecería, por decir lo menos, de seguridad independiente. Hay una explicación mucho mejor del por qué Splaine identificó a Sacco, diferente a haberlo vio en la escena del crimen —que lo había visto en la comisaría y en el juzgado—; hay una explicación mucho mejor de por qué Pelzer identificó a Sacco, diferente a que él era el hombre que había visto en la escena del crimen —que Pelzer no quería admitir que se escondió después de escuchar los disparos—; y parece posible que Devlin, quien después afirmó tajantemente que *nunca había tenido duda alguna* de que Sacco era el hombre que había visto, con el tiempo había aumentado su seguridad quizás por «la intensidad del delito y todo lo demás»¹¹⁰.

Véase, especialmente HAACK, 1993: 2009; capítulo 1. Imaginar a FRANKFURTER leyendo mi libro, me hace recordar al Sr. Jourdain; personaje ficticio de MOLIÈRE, que descubre para su asombro que había «hablado prosa durante cuarenta años sin saber sobre éstas». MOLIÈRE, 1670: Act II, Scene VI.

¹⁰⁶ FRANKFURTER, 1927: 11-15; la cita es de la p. 11.

¹⁰⁷ *Ibid.*: 15.

¹⁰⁸ *Ibid.*: 17-18.

¹⁰⁹ *Ibid.*: 16.

¹¹⁰ *Ibid.*: 15.

FRANKFURTER también analiza cuidadosamente la prueba pericial sobre las armas de fuego. Como perito en el juicio, el capitán William H. Proctor testificó que «la bala 3» era «compatible con haber sido disparada con la pistola de Sacco»¹¹¹; y el juez Thayer interpretó que esto significaba que «la bala fue disparada con la pistola [de Sacco]». Sin embargo, no todo el testimonio de Proctor fue presentado en el proceso: en una declaración jurada subsiguiente afirmó que «en ningún momento [fue] capaz de encontrar prueba alguna que lograra convencer[lo] de que [esa bala] provenía de la pistola de Sacco»¹¹². «Por un acuerdo anterior», Frankfurt observó, «la acusación llevó ante el jurado un elemento de prueba aparentemente más perjudicial para los demandados, cuando de hecho la verdad completa sobre estas pruebas era muy favorable a ellos»¹¹³. En resumen, yo diría que:

la prueba pericial sobre las armas de fuego presentada en el juicio era muy deficientemente inclusiva, con la consecuencia de que la conclusión de que una de las balas mortales provenía del arma de Sacco estaba muy pobremente avalada.

La acusación presentó las mentiras que Sacco y Vanzetti dijeron en la comisión como un indicador de su «conciencia de culpabilidad»; pero dado que en ese tiempo hubo un gran número de detenciones y deportaciones de extranjeros sospechosos de ser comunistas simpatizantes, Sacco y Vanzetti pudieron haber mentido no porque fuesen responsables del delito de Brantree, sino porque pensaron que tendrían problemas con la policía por su radicalismo político¹¹⁴. Además, como FRANKFURTER señala, no se encontró en posesión de Sacco o Vanzetti ni uno sólo de los 16.000 dólares robados; y tampoco ninguno de los dos hombres se ocultó o huyó del país después del delito —ambos permanecieron en sus antiguos lugares de residencia y ejerciendo el mismo trabajo—¹¹⁵. Sin embargo, continúa, poco después del delito, Maderios misteriosamente entró en posesión de 2.800 dólares —el tipo de suma que uno esperaría que alguien recibiera como parte del botín—¹¹⁶. En resumen, yo diría que:

las mentiras que Sacco y Vanzetti dijeron a la policía ofrecen sólo un débil apoyo a la conclusión de que éstos trataban de ocultar su participación en el delito de Brantree; pues la integración explicativa de estas pruebas con dicha conclusión no es mejor que su integración con la conclusión de que ellos trataban de evitar ser castigados por sus convicciones políticas. Además, las pruebas adicionales sobre su comportamiento posterior dan un grado de apoyo mayor a esta última conclusión, al mismo tiempo que socavaban la primera.

¹¹¹ Otras cinco balas fueron encontradas en los cadáveres en la escena del crimen; las pruebas excluyeron la posibilidad de que cualquiera de estas otras balas hubiera sido disparada por Sacco o por Vanzetti. *Ibid.*: 76.

¹¹² *Ibid.*: 76-79.

¹¹³ *Ibid.*: 76.

¹¹⁴ *Ibid.*: 35 y ss.

¹¹⁵ *Ibid.*: 35-36.

¹¹⁶ *Ibid.*: 114.

Podría continuar; pero espero que esto sea suficiente para ilustrar las ventajas de una aproximación fundaherentista sobre la aproximación impresionantemente complicada pero finalmente incomprensible del «análisis probabilístico» de dichas pruebas ofrecido por KADANE Y SCHUM.

Bueno, sí, usted podría pensar: pero incluso si la aproximación epistemológica fundaherentista fuese, como aquí se dice, mejor *para este caso*, éste no es realmente el tipo de caso que ilustra mejor las virtudes del probabilismo jurídico. Después de todo, en el caso Sacco y Vanzetti no se presentó ninguna prueba probabilística o estadística; y seguramente es ahí donde el probabilismo jurídico realmente demuestra su utilidad. No estoy de acuerdo; sin embargo, para demostrar por qué es necesario presentar el segundo ejemplo.

4. PEOPLE V. COLLINS: UNA VEZ MÁS, EL FUNDAMENTISMO FUNCIONA MEJOR QUE EL BAYESIANISMO SUBJETIVO

No hay duda alguna de que las pruebas estadísticas —desde los análisis de ADN¹¹⁷ en casos penales hasta los estudios epidemiológicos en casos de daños por sustancias tóxicas y los cálculos actuariales en demandas de daños por muerte, etc.— juegan un papel muy significativo en el derecho; y de forma totalmente apropiada. Y tampoco hay duda de que el cálculo matemático de probabilidades sea aplicable a las pruebas estadísticas. Sin embargo, de esto no se sigue que el cálculo de probabilidades pueda esclarecer el papel que juegan las pruebas estadísticas en el contexto de un cuerpo probatorio más extenso en un caso; y, como pronto veremos, esto tampoco es así. Al contrario, de hecho: al incitarnos a confundir las probabilidades estadísticas con los grados de prueba, el probabilismo jurídico puede tentarnos a olvidar que las pruebas estadísticas de un caso deben tratarse como *una pieza probatoria más entre muchas otras*.

Repto deliberadamente las palabras de la Corte Suprema de Nueva Jersey en el caso *Landriagan* para rechazar la idea de que las pruebas epidemiológicas que demuestran un aumento del riesgo en más del doble son suficientes para establecer la causalidad específica en un caso de daños por sustancias tóxicas: «Un riesgo relativo de 2,0 no es determinante, sino solamente un elemento probatorio entre muchos otros para probar la causalidad»¹¹⁸. Esto es

¹¹⁷ Los cálculos de probabilidad de una coincidencia con el demandado son extrapolaciones matemáticas de los datos empíricos. No son datos matemáticos puros; pero tampoco son exactamente «estadísticos», pudieran ser cualquiera de los dos si esto se interpreta como «basados en estudios de patrones de [ADN] en la población». [Véase ALKEN, ROBERTS Y JACKSON (2010) para intentar delimitar las diferencias necesarias].

¹¹⁸ *Landriagan v. Celotex Corp.*, 127 N. J. 405, 419, 605 A.2d 1079, 1087 (1992). Compárese esto con un caso penal inglés de 1997: «el juez debería haber dejado al jurado evaluar, por una parte, la convincente prueba de ADN junto con otras pruebas que identifican al apellado como el potencial atacante

exactamente correcto en el punto epistemológico. Y también deliberadamente me posiciono en contra de la siguiente afirmación del artículo sobre epidemiología publicado en la primera edición del *Reference Manual on Scientific Evidence (Manual de Referencia sobre la Prueba Científica)* que data de un par de años atrás: «El riesgo relativo de un estudio epidemiológico puede ser adaptado al estándar [civil] del más del 50 por 100 para llegar a la conclusión de que hay una probabilidad o posibilidad de que un agente causara la enfermedad de un individuo».¹¹⁹ Éste comete exactamente la confusión contra la que estoy advirtiendo.

Aunque, como indica el párrafo anterior, mi explicación de los grados de prueba pretende una aplicación general, me enfocaré nuevamente en las pruebas de otro caso penal estadounidense. *People v. Collins* no fue, como el caso de *Sacco y Vanzetti*, una *cause célèbre*; por el contrario, fue un tipo de caso que podría apasionarle sólo a un estudioso de la prueba —francamente insignificante, excepto por un giro epistemológicamente curioso—. Fue este giro lo que condujo a FINKELSTEIN y FARLEY a empezar lo que sería un artículo clave en la literatura del probabilismo jurídico, ofreciendo «un acercamiento bayesiano a las pruebas de identificación», con un comentario sobre el caso *Collins*.¹²⁰ Sin embargo, como veremos, pese a que el trabajo de FINKELSTEIN y FARLEY sea filosóficamente un poco más sofisticado que el de KADANE y SCHUM, también es seriamente defectuoso.

Pero permítaseme comenzar por el inicio: con la anciana que fue atacada en un callejón para robarle su monedero. La víctima describió al ladrón como una mujer joven de un peso aproximado de 64 kilos, de cabello rubio peinada con cola de caballo; y otro testigo ocular dijo que había visto a una mujer rubia huyendo por el callejón y subiendo a un coche amarillo conducido por un hombre de raza negra con bigote y barba. Janet y Malcolm Collins fueron acusados del delito. Janet era rubia y Malcolm era negro; y conducían un antiguo coche amarillo.¹²¹ Sin embargo, la víctima no pudo identificar positivamente a Janet; y la identificación que el otro testigo ocular hizo de Malcolm era débil. Entonces, para reforzar una acusación que no era tan fuerte, la acusación presentó primero a un maestro de matemáticas para testificar —sobre la base de un argumento puramente probabilístico basado no en datos estadísticos sino en una

y, por otra parte, las pruebas del demandado y de las coartadas de sus testigos», R. v. *Doherty y Adams* (1997) 1 Cr. App. R. 369 (1997), 378.

¹¹⁹ BAILEY *et al.*, 1990: 168 (el artículo sobre epidemiología en la segunda edición del *Reference Manual* del 2000 hace una afirmación muy similar, pero ahora acumulando referencias a una serie de casos). GREEN *et al.*, 2000: 384. Para una crítica detallada del criterio del «aumento del riesgo en más del doble», véase HAACK, próximamente.

¹²⁰ FINKELSTEIN y FARLEY, 1969-1970.

¹²¹ El oficial Kinsey, mientras conducía a casa después del trabajo, vio a los demandados en su Lincoln amarillo, los vigió y los siguió a casa. La pareja, dicho sea de paso, no coincidió exactamente con la descripción de los testigos oculares: el pelo de Janet era rubio oscuro y Malcolm no tenía barba. *Collins* (nota 89 *supra*), 34.

mera asunción— que dada «la regla del producto», las probabilidades de que en las cercanías existiera otra pareja semejante (una mujer rubia, un hombre negro con vello facial conduciendo un coche amarillo) eran astronómicas: uno en 12 millones; y luego le dijo al jurado que estos números inventados eran estimaciones «bajas» y que eran la «prueba matemática» de que los Collins eran culpables.¹²² Como era de esperarse, el jurado condenó a la pareja.

Janet cumplió su condena; pero Malcolm apeló y en 1968 se le concedió la apertura de un nuevo juicio.¹²³ El argumento que el magistrado Sullivan planteó en la sentencia que concedió el nuevo juicio tenía dos flancos: primero, que las estadísticas que aparentemente influyeron en el jurado fueron inventadas, carecían por completo de cualquier base empírica; segundo, que aun si éstas hubiesen sido reales, «ninguna fórmula matemática podría jamás establecer [...] que los testigos oculares de la acusación habían observado correctamente y descrito con exactitud los rasgos distintivos [...] que relacionaban a los acusados con los hechos delictivos».¹²⁴ El primer punto es innegable; el segundo —aunque es también, en mi opinión, correcto— suscita preguntas sobre cuál es exactamente el papel que juegan las pruebas estadísticas en la identificación de un delincuente o delincuentes concretos.

Las pruebas estadísticas de identificación, según argumentaron FINKELSTEIN y FARLEY, normalmente no deberían ser suficientes por sí mismas, sino que necesitarían ir acompañadas de otras pruebas que forman la base de una «estimación "previa" de la identidad». El teorema de Bayes, continúan, justifica esta respuesta, pues nos permite «traducir» una probabilidad estadística en «un enunciado probabilístico que describe la fuerza probatoria de dicha estadística».¹²⁵ En resumen, el teorema de Bayes sirve para transformar de alguna manera el sentido estadístico de «probable» en un sentido epistémico. Claramente, un teorema del cálculo de probabilidades, que no es más que un teorema matemático puro, no puede hacer tal milagro «de traducción». Entonces no resulta muy sorprendente que de hecho FINKELSTEIN y FARLEY hayan establecido un puente entre las probabilidades estadísticas y las epistémicas, no aplicando el teorema de Bayes (como ya se ha advertido), sino jugando con la interpretación de «probable». Al igual que KADANE y SCHUM, FINKELSTEIN y FARLEY intentan explicar la probabilidad «subjensiva» como la probabilidad que un sujeto atribuye a la proposición en cuestión. Sin embargo, a diferencia de KADANE y SCHUM, FINKELSTEIN y FARLEY intentan explicar qué es esta última probabilidad: «la frecuencia relativa de la culpabilidad en los casos juzgados que son similares en el grado de creencia que generan».¹²⁶

¹²² *Collins* (nota 89 *supra*), 37 («estimaciones bajas»), 41 («prueba matemática»).

¹²³ Sin embargo, la acusación no pudo reunir a los testigos por segunda vez, por lo cual el nuevo juicio nunca se llevó a cabo. FISHER, 2006: 21-22.

¹²⁴ *Collins* (nota 89 *supra*), 40.

¹²⁵ FINKELSTEIN y FARLEY, 1969-1970: 498.

¹²⁶ *Ibid.*: 504.

Creo que esto significaría que la probabilidad «subjéctiva» [es decir, la probabilidad *objetiva* asignada por un sujeto *x*] de que el acusado es culpable sería la proporción de los casos [supuestamente, posibles pero no necesariamente reales] en los cuales los hechos son diferentes al caso en cuestión pero en los que *x* considera que su grado de creencia en la culpabilidad del acusado sería el mismo y en los cuales los acusados *serían* culpables. FINKELSTEIN y FARLEY admiten que su interpretación es «artificial»¹⁷, pero el hecho es que ésta es casi ininteligible. ¿Qué clase de posibles casos se supone que imagina un jurado?, ¿cómo deben individualizarse estos posibles casos?, ¿cómo se supone que un jurado estimaría la proporción de aquellos posibles casos en los cuales el acusado sería culpable? y ¿cómo se supone que todo esto funcionaría cuando del juicio que *x* hace sobre la probabilidad de la culpabilidad del acusado se extrapola a sus atribuciones de probabilidades previas o de probabilidades condicionales?

Para mostrar que mi propuesta puede hacerlo mejor, permítaseme comenzar con la observación del magistrado SULLIVAN —con la que FINKELSTEIN y FARLEY están de acuerdo— respecto a que las pruebas estadísticas versan sobre una población y por ello, sin otras pruebas, no pueden garantizar una conclusión sobre un individuo. Imaginemos un caso como *Collins* pero con pruebas estadísticas reales, no inventadas. Aunque pareciera demasiado obvio, quizá no lo sea tanto por lo que vale la pena decir que las estadísticas por sí mismas no tendrían ninguna relevancia sin: i) los testimonios de los testigos oculares, y ii) el hecho de que los acusados encajen en la descripción de los delinquentes dada por los testigos oculares. Entonces, supongamos que todas las pruebas que tenemos son:

— P1: testigos oculares que identifican a los delinquentes como una mujer rubia y un hombre de raza negra con bigote y barba que huyeron en un coche amarillo.

— P2: las parejas que encajan en esta descripción son muy raras (una en *n*) en la población del área donde ocurrió el robo.

— P3: los Collins encajan en esta descripción.

Estas pruebas dan *algún* grado de apoyo a la conclusión de que los Collins fueron los delinquentes: que los Collins cometieron el delito explicaría lo visto por los testigos oculares; y las pruebas estadísticas indican que hay otras relativamente pocas explicaciones posibles. Sin embargo, un cierto grado de apoyo no es suficiente; y el grado de aval que estas pruebas dan a la conclusión es obviamente muy bajo: el grado de apoyo no sólo es menos que abrumador, sino que también hay varias cuestiones relativas a la seguridad independiente (por ejemplo, la fiabilidad de los testigos oculares) y a la inclusión de las pruebas —hay muchas pruebas relevantes que faltan—.

Ahora imaginemos que tenemos algunas pruebas adicionales:

— P4: los testigos oculares no tienen problemas visuales, tienen buena visión; y no tienen razones para mentir, etcétera.

— P5: hay razones para creer que los delinquentes eran personas de la localidad.

— P6: los Collins no tienen ninguna coartada.

— P7: el comportamiento posterior de los Collins era evasivo, etcétera.

La adición de P4-P7 mejora considerablemente la integración explicativa de las pruebas a la conclusión: la conjunción de P1-P3 apoya en un grado bastante modesto la conclusión de que los Collins cometieron el delito, pero la conjunción de P1-P7 la apoya en un grado significativamente más alto. Además, han sido resueltas algunas cuestiones sobre la seguridad independiente; P1-P7 es considerablemente más inclusivo que P1-P3; y las pruebas adicionales no son menos favorables para la conclusión. Entonces P1-P7 avalan la conclusión en un grado considerablemente más alto de lo que hacen P1-P3.

Pero ahora imaginemos que, en cambio, tenemos estas otras pruebas adicionales:

— P4*: los testigos oculares tienen problemas de visión y/o no tenían una buena visión y/o tenían algún motivo para señalar a Janet y Malcolm Collins y/o los habían visto esposados en la comisaría antes de seleccionarlos en la rueda de identificación y/o etcétera.

— P5*: hay razones para creer que los autores del delito no eran personas de la localidad sino que provenían de otro estado.

— P6*: los Collins ofrecen una coartada que ha sido confirmada.

— P7*: los Collins presentaron documentos (verificados como legítimos) con los que demostraron que compraron su coche amarillo *después* de que el delito se hubiera cometido.

P4*, P5*, P6* y P7* *socavan* la conclusión de que los Collins eran los delinquentes; entonces, la conjunción P1-P3 con P4*-P7* está explicativamente mejor integrada, y por ende le da un mayor grado de apoyo, con la conclusión de que los Collins *no* fueron los autores, y no con la conclusión de que sí lo eran. Y la confirmación de las pruebas sobre la coartada y los documentos sobre la compra del coche aumenta el grado de aval de la conclusión de que ellos no eran los autores, resolviendo algunas cuestiones sobre la seguridad independiente de las pruebas que minan la conclusión de que sí lo eran.

Podría continuar, pero espero que esto sea suficiente para ilustrar las ventajas de una aproximación fundamente ingenuo pero finalmente confiable. Y si esto es así —dado que FINKELSTEIN y FARLEY no sólo están interesados en el caso *Collins* sino, de forma más general, en el papel de las pruebas

estadísticas de identificación— mi concepción debería ser más esclarecedora en este ámbito más amplio.

Cuando el artículo de FINKELSTEIN y FARLEY se publicó «la identificación de huellas genéticas» todavía no había entrado en el sistema jurídico¹²⁸. Sin embargo, los casos que involucran esta clase de pruebas cada vez más comunes pueden ilustrar particularmente bien el punto epistemológico general. El Sr. Raymond Easton fue arrestado por un robo con fundamento en la coincidencia positiva de su ADN con los registros en una base de datos policial; estadísticamente, era muy baja la probabilidad de que el ADN del Sr. Easton, archivado después de una detención por violencia doméstica, y el ADN encontrado en la escena de delito coincidieran por azar. No obstante, el Sr. Easton, que sufría de parkinson, era demasiado débil para vestirse él mismo o dar un paseo más allá de unos pocos metros —por no mencionar que lo era para haber conducido hasta el escenario del delito o haber cometido el delito—¹²⁹.

Éste es un caso estructuralmente muy parecido a mi segunda extrapolación imaginaria en el caso *Collins*: las pruebas probabilísticas parecen apoyar la conclusión de que el Sr. Easton era el autor del delito (pues el hecho de que el Sr. Easton hubiese estado en la escena del crimen explicaría que su ADN se encontrara allí); sin embargo, hay otras pruebas que socaban esta conclusión (dado que el Sr. Easton estaba físicamente muy limitado, no tenemos ninguna explicación de cómo llegó a estar en la escena del crimen o cómo podría haber cometido el robo). En resumen, la conjunción de las pruebas estadísticas con las demás pruebas da cierto grado de apoyo a la conclusión de que el Sr. Easton no cometió el delito y socaba la conclusión de que sí lo hizo. Y, entonces —asumiendo una seguridad independiente razonable (por ejemplo, que un prestigioso doctor confirmó que el Sr. Easton realmente padecía de parkinson) y asumiendo que no falta ninguna prueba significativa (por ejemplo, que el Sr. Easton pagó a un cómplice para que lo transportara al escenario de delito, etc.)—, la conclusión de que él no cometió el delito está bien avalada.

* * *

Por supuesto mi argumento no se restringe a los casos penales. Después de todo, las pruebas de ADN pueden desempeñar un papel importante en casos, por ejemplo, de paternidad o de sucesión; las pruebas epidemiológicas juegan un papel muy importante en los casos de daños por sustancias tóxicas. En todo caso —¿realmente tengo que decir esto otra vez?— mi «disensión epistemológica» hacia el probabilismo jurídico no es aplicable únicamente a los juicios que implican pruebas estadísticas, sino que es totalmente general. La cuestión

¹²⁸ Las pruebas de identificación por ADN fueron usadas por primera vez en un caso penal en Estados Unidos en 1987, en el proceso de Tommy Lee Andrews en Florida. Véase *Andrews v. State*, 533 So 2d 841 (Fla. App. 1988).

¹²⁹ GENEWATCH UK, Jan. 2005: 23.

es que en lugar de mejorar nuestra comprensión sobre qué son los grados de prueba en el derecho, el probabilismo la impide —y que una epistemología sólida puede ayudar ahí donde al probabilismo jurídico obstruye—¹³⁰.

BIBLIOGRAFÍA

- ADVISORY COMMITTEE ON EVIDENCE RULES, 2009: *Report. Committee on Rules of Practice and Procedure of the Judicial Conference of the United States*, Washington, D. C.
- AIKEN, C.; ROBERTS, P., y JACKSON, G., 2010: *Fundamentals of Probability and Statistical Evidence in Criminal Proceedings: Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses*, London: Royal Statistical Society.
- ALLEN, M. E., 2010: *Florida Criminal Procedure*, Eagan, MN: Thomson Reuters.
- ALLEN, R. J., 1986: «A Reconceptualization of Civil Trials», *Boston University Law Review*, 66: 401-437. Reimpreso en TILLERS y GREEN, 1988: 21-60.
- ANÓNIMO, 1969: *The Sacco and Vanzetti Case: Transcript of the Record of the Trial of Nicola Sacco and Bartolemeo Vanzetti in the Courts of Massachusetts and Subsequent Proceedings, 1920-27*, Mamaroneck, NY: Paul P. Appel (6 vols.). — 2006: «Islam and Rape», *Wall Street Journal*, 6 de agosto, A6.
- BAILEY, L.; GORDIS, L., y GREEN, M., 1990: «Reference Guide on Epidemiology», *Reference Manual on Scientific Evidence*, Washington, D. C.: Federal Judicial Center, 123-180.
- BARTLETT, R., 1986: *Trial by Fire and Water: The Medieval Judicial Ordeal*, Oxford: Clarendon Press.
- BOOLE, G., 1854: *The Laws of Thought*. Reimpreso, New York: Dover (sin fecha).
- BRIDGMAN, P., 1933: «The Struggle for Intellectual Integrity». Reimpreso en BRIDGMAN, *Reflections of a Physicist*, New York: Philosophical Library, 1955, 361-379.
- BROUN, K. (ed.), 2006: *McCormick on Evidence*, 6.ª ed., St. Paul, MN: Thomson/West.
- CHADBOURN, J. H. (ed.), 1981: *Wigmore on Evidence*, Boston: Little, Brown and Company.
- COHEN, L. J., 1977: *The Probable and the Probable*, Oxford: Clarendon Press.
- DELOGGIO, L., 1986: «Beyond a Reasonable Doubt: A Historic Analysis», *New York State Bar Journal*, abril, 19-25.
- DEKSHOWITZ, A. M., 2011: «Casey Anthony: The System Worked», *Wall Street Journal*, 7 de julio, A15.
- FINKELSTEIN, M. O., y FARLEY, W. B., 1969-70: «A Bayesian Approach to Identification Evidence», *Harvard Law Review*, 83.3: 489-517.
- FISHER, G., 2006: «Green Felt Jungle: The Story of *People v. Collins*», en LEMPERT, R. (ed.), *Evidence Stories*, New York: Foundation Press, 7-27.
- FRANKFURTER, F., 1927: *The Case of Sacco and Vanzetti: A Critical Analysis for Lawyers and Laymen*, Boston: Little, Brown and Company.

¹³⁰ Mi agradecimiento a Mark MCCORTI por sus útiles comentarios al borrador de este texto, a Paul ROBERTS por el provechoso intercambio de correspondencia y a Pamela LUCKEN por su hábil ayuda en la búsqueda de materiales relevantes.

- GENEWATCH, U. K., 2005: *The Police National DNA Database: Balancing Crime Detection, Human Rights and Privacy*, Derbyshire, UK. <http://www.genewatch.org/uploads/03c6d66a9b354557384831c349e4/NationalDNADatabase.pdf>.
- GIWENS, R. A., actualizado por SHREY, K., 2010: *Manual of Federal Practice, 2010 Cumulative Supplement*, New Providence, N. J.: LexisNexis.
- GODWIN, J., 2000: «Wigmore's Chart Methods», *Informal Logic*, 20.3: 223-243.
- GOODMAN, N., 1954: «The New Riddle of Induction», en GOODMAN, *Fact, Fiction and Forecasts*, 2.ª ed., Indianapolis, IN: Bobbs-Merrill, 59-83.
- GRAHAM, M., 2007: *Evidence*, 2.ª ed., St. Paul, MN: Thomson/West.
- GREEN, M.; FREEDMAN, D. M., y GORDIS, L., 2000: «Reference Guide on Epidemiology», *Reference Manual on Scientific Evidence*, 2.ª ed., Washington, D. C.: Federal Judicial Center, 333-400.
- HAACK, S., 1993/2009: *Evidence and Inquiry: A Pragmatist Reconstruction of Epistemology*, 2.ª ed., Amherst, NY: Prometheus Books (1.ª ed., 1993).
- 1998: «A Foundherentist Theory of Empirical Justification», en POJMAN, L. J. (ed.), *The Theory of Knowledge: Classic and Contemporary Sources*, Belmont, CA: Wadsworth, 2.ª ed., 283-293. Reimpreso en Sosa, E., y Kim, J. (eds.), *Epistemology: An Anthology*, Oxford: Blackwell, 2000, 226-236. Traducción al castellano de FALGUERA, J. I., *Agora*, 18.1, 1999: 35-53.
- 2003: *Defending Science Within Reason: Between Scientism and Cynicism*, Amherst, NY: Prometheus Books.
- 2004a: «Epistemology Legalized: Or, Truth, Justice, and the American Way», *American Journal of Jurisprudence*, 49: 43-61. Reimpreso en la segunda edición de HAACK, 1993/2009: 361-381.
- 2004b: «Coherence, Consistency, CoGENCY, Congruity, Cohesiveness, &c.: Remain Calm! Don't Go Overboard!», *New Literary History*, 35: 167-183. Reimpreso en HAACK, S., 2008: *Putting Philosophy to Work: Inquiry and Its Place in Culture*, Amherst, NY: Prometheus Books, pp. 61-78.
- 2008: «Proving Causation: The Holism of Warrant and the Atomism of Daubert», *Journal of Health and Biomedical Law*, IV.2: 253-289.
- 2009: «The Growth of Meaning and the Limits of Formalism, in Science and Law», *Analisis Filosófico*, XXXIX.1:5-29.
- 2011, en prensa: «Epistemology: Who Needs It?», manuscrito. Publicación al dadas en la traducción de STRINSELT, F., *Kritik*, 200 (2011): 26-35 (la versión en inglés está disponible bajo solicitud a la autora).
- 2012: «The Embedded Epistemologist: Dispatches from the Legal Front», *Ratio Juris*, 25.2: 206-235.
- próximamente: «Risky Business: Statistical Proof of Specific Causation», manuscrito, citado por la versión castellana de MURFARO, N.: «Asuntos arriesgados: sobre la prueba estadística de la causalidad específica», en PAPAYANNIS, D. (ed.), *Causalidad y atribución de responsabilidad*, Madrid: Marcial Pons, en prensa.
- HACKING, I., 1975: *The Emergence of Probability*, Madrid: Marcial Pons, en prensa.
- Press. Citado por la traducción al castellano de ALVAREZ, J. A., 1995: *El surgimiento de la Probabilidad. Un estudio filosófico de las ideas tempranas acerca de la probabilidad, la inducción y la inferencia estadística*, Barcelona: Gedisa.
- JAFFE, L., 1984-1985: «Of Probability and Probability», *Statistics, Scientific Evidence, and the Calculus of Chances at Trials*, *University of Pittsburgh Law Review*, 46: 925-1083.

- JANIN, H., y KAHLMEYER, A., 2007: *Islamic Law: The Shari'a from Muhammad's Time to the Present*, Jefferson, NC: McFarland and Company.
- KADANE, J., y SCHUM, D., 1996: *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*, New York: John Wiley and Sons.
- KADRI, S., 2005: *The Trial: A History, from Socrates to O. J. Simpson*, New York: Random House.
- KAPLAN, J., 1968: «Decision Theory and the Factfinding Process», *Stanford Law Review*, 20: 1065-1092.
- KENNES, J. M., 1921: *A Treatise on Probability*, London: MacMillan.
- LEMPERT, R., 1988: «The New Evidence Scholarship: Analyzing the Process of Proof», en TILERS y GREEN, 1988: 62-102.
- MATLAND, F., 1909: *The Forms of Action at Common Law*, Cambridge: Cambridge University Press. Hay traducción al castellano de CREMADDES UGARTE, I., 2005: *Las «formas de acción» en el Common Law*, Madrid: Marcial Pons.
- MARTIN, S., 1994: *Utah's Influence*, New York: G. P. Putnam. Las páginas citadas son de la edición en tapa blanda. New York: Jove Books, 1996.
- MCCORMACK, C. T., 1954: *Handbook of the Law of Evidence*, St. Paul, MN: West Publishing Co.
- MILL, J. S., 1843: *A System of Logic, Ratiocinative and Inductive, Being a Connected View of the Principles and Methods of Scientific Investigation*, 8.ª ed., London: Longman, 1970.
- MILLAY, E. ST. VINCENT, 1928: «Justice Denied in Massachusetts», en *The Buck in the Snow and Other Poems*, New York: Harper and Brothers Publishers.
- MOJERE, J. B. P., 1670: *Le Bourgeois Gentilhomme*.
- NEILSON, G., 1890: *Trial by Combat*, London: Williams and Norgate.
- OLIVER, L., 2002: *The Beginnings of English Law*, Toronto: University of Toronto Press.
- O'MALLEY, K. F.; GREING, J. E., y LEE, W. C., 2000: *Federal Jury Practice and Instructions: Civil*, vol. 3, 5.ª ed., Eagan, MN: West Group (y suplemento del 2010).
- 2008: *Federal Jury Practice and Instructions: Criminal*, vol. 1A, 6.ª ed., Eagan, MN: Thomson/West (y suplemento de 2010).
- PARDO, M., 2013: «Standards of Proof, and Evidence Theory». Citado por la traducción al castellano de SEVRS, G.: «Estándares de prueba y teoría de la prueba», publicado en este mismo volumen.
- PARK, R., y SAKS, M., 2005: «Evidence Scholarship Reconsidered: Results of the Interdisciplinary Turn», *Boston College Law Review*, 47: 949-1031.
- POSNER, R., 2001: *Frontiers in Legal Theory*, Cambridge, MA: Harvard University Press.
- RISINER, D. M. et al., 2010: «Bayes Wars Revisited: An Exchange», *International Commentary on Evidence*, 8.1, <http://www.beppress.com/cgi/viewcontent.cgi?article=1115&context=ice>.
- RUSSELL, B., 1948: *Human Knowledge, Its Scope and Limits*, New York: Simon and Schuster. Citado por la traducción al castellano de MIGUEZ, N., 2000: *El conocimiento humano*, Barcelona: Ediciones Folio.
- SAL, O., e IQBAL, S., 2005: *A Legal Research Guide to Pakistan*, <http://www.nyulaw-global.org/globalalex/pakistan.htm>.
- SCHUM, D., 1979: «A Review of a Case against Blaise Pascal and his Heirs», *Michigan Law Review*, 77: 446-463 (una recensión de COHEN, 1977).

- 1994: *Evidential Foundations of Probabilistic Reasoning*, New York: John Wiley and Sons.
- SINCLAIR, U., 1928: *Boston*, New York: A. C. Boni.
- STARKE, T., 1842: *A Practical Treatise of the Law of Evidence*, Philadelphia: T. & J. W. Johnson.
- TARUFFO, M., 2009: *La semplice verità. Il giudice e la costruzione de fatti*, Roma: Editori Laterza. Citado por la traducción al castellano de ACCATINO, D., 2010: *Simplemente la verdad: El juez y la construcción de los hechos*, Madrid: Marcial Pons.
- TILLERS, P., y GREEN, E. D. (eds.), 1988: *Probability and the Law of Evidence: Uses and Limits of Bayesianism*, Dordrecht: Kluwer.
- TOFF, M., 2005: *The Sacco and Vanzetti Case: A Brief History with Documents*, New York: Palgrave/MacMillan.
- TRIBE, L., 1971: «Trial by Mathematics: Precision and Ritual in the Legal Process», *Harvard Law Review*, 84: 1329-1393.
- VON MISES, R., 1928: *Probability, Statistics and Truth*, London: Allen and Unwin, 2.ª ed., 1928.
- WIGMORE, J. H., 1913: *The Principles of Judicial Proof as Given by Logic, Psychology and General Experience as Illustrated in Judicial Trials*, 5.ª ed. estadounidense, Littleton, CO: Fred B. Rothman & Co., 1981.
- WRIGHT, C. A., y HENNING, P. J., 2009: *Federal Practice and Procedure*, vol. 2A, 4.ª ed., Eagan, NM: Thomson Reuters (y suplemento 2010).

IV

ESTÁNDARES DE PRUEBA Y TEORÍA DE LA PRUEBA *

Michael S. PARDO **
University of Alabama

En las últimas décadas se ha realizado una gran cantidad de trabajo teórico acerca de los estándares de prueba en el derecho. Gran parte de este trabajo ha incluido la aplicación de la estadística, la teoría de la probabilidad o la teoría de la decisión a los estándares de prueba y, en términos más generales, al proceso probatorio¹. En este ensayo analizo este proyecto teórico acerca de los estándares de prueba en el contexto de las teorías generales del proceso probatorio e identifico criterios según los cuales deberíamos evaluar estas propuestas teóricas y el papel que en ellas juegan los estándares de prueba.

Mi conclusión es que la teoría dominante, que entiende los estándares de prueba en términos de probabilidades, fracasa a la luz de los criterios teóricos que he identificado. Según esta concepción probabilística², los enunciados

* Traducción de Gonzalo SEMAS.

** Profesor de derecho Henry Upson Sims, School of Law, University of Alabama. Una versión anterior de este trabajo fue presentada en la conferencia «Estándares de prueba y prueba científica», Universitat de Girona, mayo de 2011. Agradezco a Jordi FERRER BELTRÁN por su amable invitación a participar en la conferencia y a los demás participantes y asistentes por sus útiles comentarios.

¹ Para un breve resumen de este trabajo, véase RUSINGER, 2010. Para una taxonomía útil de las distintas cuestiones teóricas y prácticas para las que los diferentes modelos formales de prueba pueden ser de utilidad, véase TILLERS, 2011.

² Por *probabilística* entiendo a la probabilidad pascaliana (o matemática), como opuesta a la probabilidad baconiana (o inductiva). Sobre las discusiones acerca de esta última en el contexto de la prueba jurídica, véase COHEN, 1977; STEIN, 2005.

fácticos en disputa se pueden evaluar de acuerdo a su probabilidad, dadas las pruebas, conforme a los axiomas o reglas de la teoría de la probabilidad: por ejemplo:

— Se encuentran en una escala del 1 (que equivale a certeza respecto de la verdad) al 0 (que equivale a la certeza respecto de la falsedad).

— La probabilidad de un enunciado fáctico y de su negación es igual a 1.

— La probabilidad de dos proposiciones independientes se adecúa a la regla del producto.

— La probabilidad de un enunciado fáctico a la luz de un nuevo elemento de juicio se adecúa al teorema de Bayes.³

Según esta concepción, los estándares de prueba establecen un umbral mínimo de probabilidad que debe ser alcanzado. Propongo una concepción alternativa «explicativa» de la prueba que satisface mejor los criterios generales identificados. Según esta concepción alternativa, los enunciados fácticos en disputa son evaluados de acuerdo a cómo explicarían, si fueran verdaderos, los elementos de prueba y los eventos cuando son comparados con enunciados fácticos rivales, contrastantes. De acuerdo con esta concepción, los estándares de prueba establecen los umbrales explicativos que deben ser alcanzados.

Aunque se trata de un tema teórico, las consecuencias son prácticas y significativas. En los Estados Unidos, por ejemplo⁴, están en juego no sólo cuestiones de prueba en el proceso, sino también importantes cuestiones acerca de quién va a juicio oral, qué veredictos son irrecurribles y qué condenas deberían ser revocadas⁵. En los Estados Unidos, un contexto de reducción del número de casos que van a juicio en el ámbito civil y una preocupación (justificada) por las condenas erróneas en el ámbito penal, hacen que estas cuestiones doctrinarias prácticas adquieran una importancia creciente y que requieran clarificación.

1. CRITERIOS TEÓRICOS Y PRESUPUESTOS CLARIFICADORES

Las propuestas teóricas acerca de la prueba ofrecen concepciones acerca de la estructura y la naturaleza del proceso probatorio y de cómo éste se relaciona con los fines subyacentes de corrección fáctica y distribución de los riesgos de error entre las partes. Obviamente, esto tiene mucha relevancia teórica y

³ Véase KOLMOGOROV, 1956. Según esta concepción general, las probabilidades individuales pueden ser articuladas de varias maneras: por ejemplo, como relaciones lógicas, como frecuencias, como propensiones o como creencias subjetivas. Véase HAYEK, 2009.

⁴ Mi presentación se centrará únicamente en el derecho probatorio estadounidense; sin embargo, las cuestiones teóricas y epistemológicas discutidas pueden ser aplicadas de manera más general.

⁵ *Federal Rules of Civil Procedure*, 50 y 56; *Federal Rules of Criminal Procedure*, 29; *Jackson v. Virginia*, 443 U. S. 307 (1979).

práctica para el ejercicio probatorio⁶, e incluye aspectos morales, políticos y económicos, pero estas cuestiones epistemológicas centrales relacionadas con la corrección y los errores son necesarias —aunque no suficientes— para poder ofrecer una propuesta teórica satisfactoria acerca de la prueba⁷. Una propuesta teórica satisfactoria de los estándares de prueba debe ser parte de, o al menos consistente con, una propuesta teórica satisfactoria acerca del proceso probatorio como un todo en lo concerniente a estas cuestiones epistemológicas.

Una propuesta teórica satisfactoria del proceso probatorio debe cumplir con tres requisitos. Primero, debe ofrecer una propuesta plausible acerca de la relevancia y del valor probatorio de cada elemento de juicio concreto. Me referiré a esto como el requisito de «nivel micro». Segundo, debe dar cuenta de forma plausible de los diferentes estándares de decisión y de la suficiencia de la prueba. Me referiré a esto como el requisito de «nivel macro». Tercero, debe ofrecer una visión plausible acerca de la relación que hay entre las propuestas de estos dos niveles. Me referiré a esto como el requisito de «integración». Explicaré a continuación la naturaleza e importancia de cada requisito.

1.1. El requisito de nivel micro

Una propuesta teórica acerca de la prueba debería ser capaz de explicar la relevancia (o irrelevancia) y el valor probatorio de las pruebas⁸. Los dos principios fundamentales del derecho probatorio moderno son: 1) que la prueba irrelevante debe ser excluida, y 2) la prueba relevante debe ser admitida, salvo que hubiera una buena razón para excluirla⁹. Una teoría satisfactoria acerca de la prueba debería ser capaz de explicar si la prueba es relevante o no, y *por qué*. Al proveer explicaciones plausibles de estas cuestiones, una teoría satisfactoria está cumpliendo no solamente una función teórica explicativa, sino que también podría servir a una variedad de roles normativos. Estos roles normativos incluyen la provisión de criterios para justificar o criticar decisiones particulares al ofrecer consideraciones que guían y restringen las decisiones particulares.

Consideraciones teóricas y metodológicas similares se aplican al valor probatorio. El valor probatorio de la prueba se refiere a la fuerza que tiene una prueba relevante para confirmar un enunciado fáctico en disputa. Las decisio-

⁶ Un resumen de los recientes proyectos teóricos y empíricos relacionados con el derecho probatorio puede encontrarse en PARK y SAYS, 2006.

⁷ Este núcleo epistemológico conforma la base de lo que WILLIAM TWINING denomina «la tradición racionalista» del estudio de la prueba en el contexto anglosajón. Véase TWINING, 1994: 75.

⁸ *Federal Rules of Evidence*, 401, 402.

⁹ TRAYNER, 1898: 264-266. Las razones para excluir pruebas relevantes incluyen consideraciones que caen en una de las siguientes categorías: 1) económicas, 2) la promoción de fines no-epistémicos de las políticas, 3) el control sobre las partes y 4) el control sobre el jurado.

nes acerca del valor probatorio son en gran medida contextuales y dependen no solamente de la relación empírica o lógica que existe entre las pruebas y los enunciados fácticos en disputa, sino también de cuestiones como la importancia de las pruebas, lo necesario que es ese elemento de juicio para las partes, la disponibilidad de otras pruebas similares y la medida en que el asunto sea controvertido y la otra parte presente pruebas. En el nivel micro, como en el caso de la relevancia, el valor probatorio ofrece un criterio fundamental para la admisibilidad de pruebas relevantes; los jueces pueden excluir prácticamente cualquier elemento de prueba si su valor probatorio es superado por los peligros potenciales de producir un sesgo injusto, por confundir las cuestiones a tratar, por ser engañosas para el jurado o por razones de eficiencia (dilaciones indebidas, pérdida de tiempo, redundancia)¹⁰.

1.2. El requisito de nivel macro

Una teoría de la prueba debe también ser capaz de explicar la naturaleza y la estructura de la prueba en el nivel macro. En lugar de enfocarse en los elementos individuales de prueba, las explicaciones en este nivel se enfocan en si la prueba como un todo es suficiente para satisfacer ciertos estándares de prueba particulares —en otras palabras, si un hecho ha sido o no efectivamente «probado» a los fines jurídicos—.

En el nivel macro, en cualquier caso civil o penal, una parte tiene la carga de la prueba de cada cuestión litigiosa que constituye un elemento de una demanda, de un delito o de una defensa afirmativa. La carga de la prueba está formada por dos componentes: una carga de producción y una carga de persuasión. La carga de producción es una función de la carga de persuasión —en el sentido en que las partes habrán satisfecho una carga de producción cuando hubieran presentado pruebas sobre las cuales un juzgador de los hechos razonable podría concluir que la carga de persuasión ha sido satisfecha—.¹¹ La carga de persuasión está determinada por los distintos estándares de decisión: por ejemplo, «por preponderancia de la prueba», «prueba clara y convincente» o «más allá de toda duda razonable»¹².

Los estándares de decisión especifican e indican a los juzgadores de los hechos cuándo concluir que la prueba como un todo confirma un hecho par-

¹⁰ *Federal Rules of Evidence*, 403; pero véase *Federal Rules of Evidence*, 609(a)(2). El valor probatorio es también importante en el nivel macro porque las determinaciones acerca de si la prueba es suficiente para satisfacer un estándar de decisión son, esencialmente, determinaciones acerca de la conclusión a la que podría llegar un jurado razonable sobre el valor probatorio de la prueba como un todo.
¹¹ Véase McNAUGHTON, 1955.

¹² Los estándares se aplican a categorías generales de casos. Véase *Santosky v. Kramer*, 455 U. S. 745, 757 (1982) (rechazando la aplicación de reglas de decisión particulares caso por caso). Sobre la sensatez de tener tres estándares de decisión, a la luz de las limitaciones cognoscitivas generales de los jueces y jurados, véase CLERMONT, 1987.

ticular en disputa. Son dos los fines generales que subyacen y animan a los estándares: la corrección y la distribución del riesgo de error. Cada uno de estos fines ofrece un criterio independiente para evaluar la aceptabilidad de los estándares. Por ejemplo, el estándar de «la preponderancia de la prueba», en igualdad de circunstancias, funciona (o aparenta funcionar) para minimizar errores y maximizar la corrección. Esto se sigue del hecho de que las decisiones tomadas bajo este estándar son las que resultan más probables dadas las pruebas (en comparación con sus alternativas)¹³. Asumiendo que la prueba es un buen indicador de la verdad, las decisiones que parecen más probables dadas las pruebas serán más acertadas a largo plazo que las decisiones que parecen menos probables dadas las pruebas. Incluso si la regla de preponderancia maximiza la corrección, lo hace, sin embargo, distribuyendo el riesgo de error de forma más o menos equitativa entre las partes¹⁴ (cualquiera de las partes puede presentar la hipótesis menos probable y la decisión puede ser errónea). Aunque este riesgo de error resulta aceptable en la mayoría de los casos civiles (donde la igualdad entre las partes es una consideración poco importante), esto es generalmente entendido como un riesgo de error inaceptable para los acusados en un caso penal respecto de los elementos de un delito¹⁵.

De la misma manera, centrarse solamente en la distribución de los riesgos de error deja de lado el componente clave de la corrección. Por ejemplo, una interpretación del estándar «más allá de toda duda razonable» consistente con la advertencia blackstoniana de que una falsa condena es tan problemática como diez falsas absoluciones, implica consecuencias totalmente inaceptables. Implica que sobre 100 casos una distribución aceptable de resultados sería la siguiente: 90 falsas absoluciones, nueve falsas condenas y una decisión acertada¹⁶.

Una teoría de la prueba satisfactoria debería ser capaz de explicar cómo funcionan los estándares de decisión a la luz de estos fines. También debe ser capaz de ofrecer algún criterio o guía para la aplicación de las reglas por parte de los decisores, a la luz de los fines subyacentes a la propia regla en lo que hace a la corrección y al riesgo de error¹⁷. Los estándares de decisión son instrucciones dirigidas a los juzgadores de los hechos acerca de cómo arribar a juicios consistentes con los fines de las reglas. Una teoría de la prueba que pueda explicar formalmente cómo los estándares encajan con sus fines, pero no cómo los juzgadores de los hechos podrían implementar los estándares de

¹³ Puede verse una discusión acerca de estas políticas subyacentes a la regla de preponderancia en STEIN, 2005: 143-144.

¹⁴ Véase ZAMIR, 2010.

¹⁵ Véase *In re Winship*, 397 U. S. 358 (1970).

¹⁶ Véase ALLEN y LAUDAN, 2008.

¹⁷ Véase ALLEN y LETTER, 2001 (discutiendo la restricción del «debe implicar puede» en epistemología).

forma consistente con los fines, fracasa como teoría¹⁸. El objeto de una teoría exitosa, y aquello que requiere una explicación, son las aplicaciones concretas de los estándares en casos reales.

1.3. El requisito de integración

Finalmente, una teoría de la prueba exitosa debe ofrecer una explicación plausible de cómo sus propuestas en el nivel micro y en el macro encajan entre sí. Una teoría fracasará si sus propuestas no lo logran. Una teoría que ofrece propuestas plausibles en ambos niveles puede, no obstante, fracasar si tales propuestas son mutuamente inconsistentes. Por ejemplo, una teoría que, en el nivel micro, da cuenta del valor probatorio únicamente en términos de la creencia subjetiva de cada juzgador de los hechos (*i. e.*, el valor es simplemente lo que cualquier jurado cree que es) no puede luego, en el nivel macro, explicar los estándares de prueba en términos que impliquen que las creencias subjetivas acerca del valor de la prueba de los jurados individuales puedan estar equivocadas, o viceversa. Una de dos, o el valor de la prueba depende de algo más que de las creencias subjetivas, o no. Una teoría que quiera admitir las dos cosas debe al menos explicar cómo esto es posible y por qué resultaría deseable.¹⁹

Una teoría también puede ser problemática si no explica las relaciones entre estos dos niveles o si éstas resultan de alguna manera misteriosas. Por ejemplo, una teoría que explique la relevancia en términos de probabilidades cardinales y los estándares de prueba en términos de probabilidades ordinales (o viceversa), o una teoría que explique un nivel en términos de valoraciones probabilísticas subjetivas (*i. e.*, estados mentales de creencia) y el otro nivel en términos de frecuencias objetivas relativas, debe explicar cómo se trasladan las determinaciones del nivel micro al nivel macro y viceversa. En resumen, una teoría exitosa deberá cerrar la brecha que existe entre los dos niveles.

Antes de evaluar las teorías de la prueba a la luz de estos criterios, voy a mencionar algunos presupuestos con el fin de clarificar el alcance de mi análisis. Primero, la prueba estadística es diferente de una teoría (probabilística o de otro tipo) de los aspectos del proceso probatorio. En otras palabras, las cuestiones acerca de si la prueba estadística es relevante, cuánto valor probato-

¹⁸ Sin embargo, esto no es una crítica legítima si los estándares de prueba no producen por sí mismos resultados óptimos o una cierta distribución de errores deseada. Como terminarían operando en la práctica los estándares dependerá de su interacción con una multitud de otras cuestiones, que incluyen, entre otras, las demás reglas probatorias y de procedimiento, la calidad de las pruebas, la capacidad de los juzgadores de los hechos para evaluar las pruebas de forma adecuada y la proporción en que las presunciones de cada una de las partes son verdaderas.

¹⁹ Esa explicación puede ser construida, por ejemplo, alrededor de la idea de que debería haber más control judicial en un nivel que en el otro.

rio tiene y si es suficiente para satisfacer una regla decisoria son diferentes de si la relevancia, el valor probatorio o los estándares de decisión son probabilísticos o deben ser concebidos en términos probabilísticos. Segundo, las cuestiones prácticas relacionadas con la pregunta acerca de si la prueba estadística o las técnicas estadísticas deberían ser introducidas en el juicio oral (o en otras etapas procesales) son diferentes de las cuestiones teóricas relacionadas con las concepciones de la prueba. Las cuestiones prácticas relacionadas con la introducción de pruebas o técnicas estadísticas en el proceso judicial dependen de si éstas contribuirán o no a la corrección de la determinación de los hechos. Tercero, desde un punto de vista epistemológico, las pruebas estadísticas no son necesariamente inferiores ni superiores en comparación con las pruebas no-estadísticas. Como en el caso de las pruebas no-estadísticas, algunas veces serán buenas (relevantes y de alto valor probatorio) y algunas veces no, dependiendo del contexto. No hay una distinción cualitativa abstracta, a priori o de otro tipo, entre las pruebas expresadas en términos estadísticos y las pruebas expresadas en términos no-estadísticos²⁰. En cuarto lugar las pruebas no-estadísticas también se basan en generalizaciones, por lo que si hay una distinción entre prueba «específica» y «general», se trata a lo sumo de una cuestión gradual, no categórica. Finalmente, en lo que hace a estas cuestiones teóricas, a veces se presenta una falsa dicotomía entre las concepciones probabilísticas y el *Story Model*²¹. El *Story Model* es una concepción psicológica descriptiva del comportamiento del jurado, que lo describe particularmente bien²². Y sería acertado por parte de cualquier enfoque teórico tener en cuenta el mejor enfoque empírico del comportamiento de los juzgadores de los hechos. Sin embargo, si este enfoque es cierto y si los abogados presentan sus hipótesis bajo esta forma de narrativas o teorías, son cuestiones diferentes de las teorías acerca de la naturaleza y la estructura de la prueba.

2. TEORÍAS PROBABILÍSTICAS DE LA PRUEBA Y ESTÁNDARES DE PRUEBA PROBABILÍSTICOS

2.1. El nivel micro

Un elemento de prueba es «relevante» si «tiene una tendencia a hacer más o menos probable la existencia de cualquier hecho con consecuencias para la determinación de la acción en comparación con la ausencia de dicha prueba»²³. Las pruebas que no sean relevantes no son admisibles²⁴. El «valor probatorio»

²⁰ Por ejemplo, véase PUNDIK 2008. Esto no implica negar que puede haber razones morales o políticas para distinguir entre prueba estadística y no estadística.

²¹ Para una comparación entre estos dos modelos, véase BILZ 2010; KAPTEIN 2009.

²² PENNINGTON Y HASTIE, 1991; también VIDMAR Y HANS, 2007 (discutiendo la extensa literatura que confirma el *Story Model*).

²³ *Federal Rules of Evidence*, 401.

²⁴ *Federal Rules of Evidence*, 402.

se refiere a la fuerza de la prueba —o a cuán fuerte es su «tendencia»— a la hora de confirmar o refutar un hecho relevante, teniendo en cuenta otros factores contextuales como, por ejemplo, lo necesaria que resulta esa prueba para una de las partes, la importancia de la cuestión y la medida en que es disputada por la otra parte²⁵. Las pruebas pueden ser excluidas si su valor probatorio es superado sustancialmente por la presencia de uno o más «peligros»²⁶.

Estos términos doctrinarios naturalmente invitan a hacer interpretaciones probabilísticas. Y, de hecho, las concepciones teóricas dominantes de la relevancia y el valor probatorio se expresan en términos probabilísticos. Según estas concepciones, la relevancia se expresa en términos de «ratios de probabilidad»²⁷ o simplemente en términos de un aumento (o disminución) de la probabilidad de un enunciado fáctico respecto de su probabilidad anterior a la prueba en cuestión²⁸. Estas concepciones se presentan como una descripción o explicación del significado de la definición de «relevancia»²⁹ dada por la Regla 401 y del concepto de «valor probatorio»³⁰ señalado por la Regla 403 y también como un estándar normativo que ofrece un criterio para evaluar la potencial relevancia y el valor probatorio, y con el cual evaluar la decisión tomada sobre la relevancia y el valor probatorio³¹.

Según el enfoque de la «ratio de probabilidad» examinamos la probabilidad de recibir el elemento de prueba si el enunciado fáctico es verdadero, dividida entre la probabilidad de recibir el elemento de prueba si el enunciado fáctico es falso³². De acuerdo con el teorema de Bayes³³, podemos multiplicar la ratio de probabilidad por la probabilidad anterior para obtener una probabilidad posterior actualizada. Considérese el siguiente ejemplo:

Si sabemos que el autor de un delito tiene sangre tipo A y tenemos pruebas de que el acusado tiene sangre tipo A, la prueba es relevante para probar que el acusado es el autor si: 1) la probabilidad de esta prueba (él tiene sangre tipo A), dado que él es el autor, es mayor que 2) la probabilidad de esta prueba dado que él no es el autor.³⁴

El profesor LEMPERT, asumiendo que el 50 por 100 de la población tiene el mismo grupo sanguíneo, analiza la prueba como sigue: la probabilidad de la

²⁵ *Federal Rule of Evidence* 403.

²⁶ *Federal Rule of Evidence* 403.

²⁷ Véase LEMPERT, 1977: 1021.

²⁸ DAVIS y FOLLETTE, 2002.

²⁹ Véase LEMPERT, 1977: 1025 (describiendo la ratio de probabilidad como «equivalente matemático» de la «prueba relevante» en la *Federal Rule of Evidence* 401).

³⁰ Véase KAYE, 1995; KOEHLER, 2002 [«la ratio de estas probabilidades (conocida como la ratio de probabilidad) capta el valor probatorio de la prueba»].

³¹ Véase LEMPERT, 1977: 1023 («este es un modelo normativo»).

³² Véase *ibid.*

³³ La probabilidad de la culpabilidad dada la prueba = (Probabilidad de la prueba dada la culpabilidad/Probabilidad de la prueba dada la inocencia) x la probabilidad anterior de culpabilidad.

³⁴ Véase LEMPERT, 1977: 1024.

prueba, dado que él es el autor, es 1 y la probabilidad de esta prueba si él no es el autor es 0,5. Al dividir el primer número entre el segundo se observa un aumento en la posibilidad de culpabilidad (de chances iguales) a 2:1 (o 0,67)³⁵.

En términos generales, cualquier ratio de probabilidad diferente de 1 lleva a la conclusión de que la prueba es relevante. Si las posibilidades son mayores que 1:1, entonces la prueba resulta relevante para confirmar el enunciado fáctico en cuestión. Si la probabilidad de la prueba es menor cuando asumimos que el enunciado fáctico es verdadero que la probabilidad de la prueba cuando asumimos que el enunciado fáctico es falso, entonces la prueba es relevante para refutar el enunciado fáctico en cuestión. El valor probatorio bajo esta concepción se expresa sobre la base de cuán mayor (o menor) a 1 es la ratio de probabilidad. Una ratio de probabilidad de 1 significa que la prueba es irrelevante bajo esta concepción. Un ejemplo de prueba irrelevante bajo esta concepción sería el siguiente:

En un juicio penal hay pruebas de que el acusado apoya a candidatos políticos liberales.

Como explica LEMPERT, «en ausencia de alguna razón para creer que los liberales tienen una tendencia mayor a cometer el delito en cuestión, la probabilidad de que se haya podido mostrar al acusado como liberal siendo culpable es... la misma que la probabilidad de que se haya podido mostrar al acusado como liberal siendo no culpable»³⁶. Las probabilidades de que fuera culpable se mantienen iguales a las que había antes de considerar la prueba y, por tanto, según esta concepción, la prueba resulta irrelevante.

Ciertamente, esta noción de la ratio de probabilidad logra capturar algo intuitivo e importante acerca de la relevancia. Si la ratio es mayor que 1, entonces cuando la multiplicamos por la probabilidad anterior el número aumentará, y viceversa si la ratio es menor a 1. La ratio de probabilidad también sirve como una herramienta útil para llamar la atención de estudiantes y tribunales acerca de la cuestión de cuán probable sería una prueba si el enunciado fáctico para el cual se ofrece fuera falso —un aspecto importante, y a veces pasado por alto, de la relevancia y el valor probatorio—.

Sin embargo, cuando es presentada como una teoría acerca del significado de estos conceptos o como una teoría normativa para su aplicación, la teoría de la ratio de probabilidad no funciona. Esto es así porque una ratio diferente de 1, o un aumento de la probabilidad posterior más general, no son necesarios ni suficientes para que una prueba sea relevante para confirmar la verdad de un enunciado fáctico disputado en el proceso. Desde el punto de vista metodológico, si nos topamos con casos no claros de relevancia lógica (o «racional»)

³⁵ Véase *ibid.*

³⁶ Véase *ibid.*: 1025.

y de valor probatorio que no encajan con los dictados de la teoría, entonces la teoría no logra capturar el «significado» de estos conceptos (*i. e.*, no es una afirmación analíticamente verdadera), ni tampoco es un estándar normativo con el cual se puedan medir sus aplicaciones. En cambio, la teoría coincidiría algunas veces con la relevancia y con el valor probatorio y otras no.³⁷

En cualquier juicio puede haber pruebas relevantes que cada parte sostenga que apoya su teoría acerca de lo sucedido y que, desde la perspectiva de un jurado razonable o racional, no permite distinguir entre ambas hipótesis. El siguiente es un ejemplo extraído de la transcripción de un juicio penal.³⁸ El acusado, un recluso en una prisión de máxima seguridad, es acusado de dos hechos de agresión contra los guardias de la prisión. Los cargos surgen de un altercado entre el acusado y varios guardias luego de que el acusado rehusara devolver una bandeja de comida que estaba en su celda. La teoría de la fiscalía fue que el acusado embistió y agredió a los oficiales cuando abrieron la puerta de la celda para recuperar la bandeja. El acusado testificó que uno de los guardias entró y comenzó a pegarle primero; y su abogado además argumentó que incluso si el acusado hubiera hecho el primer contacto con el oficial, estaba actuando en legítima defensa.

Una pieza de información que un jurado razonable querría conocer a la hora de analizar estas explicaciones alternativas de lo sucedido sería qué fue lo que precipitó el altercado. ¿Por qué retenía el acusado la bandeja, en primer lugar? Y ambas partes ofrecen una cierta cantidad de detalles acerca de esta prueba aparentemente relevante. Ambos abogados mencionan durante los alegatos de apertura y clausura que el acusado no había recibido un paquete que le habían enviado y que luego de varios intentos de hablar con un sargento y recuperar su paquete, el acusado rehusó entregar su bandeja de la comida.³⁹

³⁷ Al señalar ejemplos de relevancia que no encajan en la teoría de la ratio de probabilidad, no intento *meramente* sugerir que hay otros aspectos que hacen a la relevancia (por ejemplo, que alguna prueba ilustrativa o demostrativa puede ser relevante sólo en la medida que ayuda a los jurados a entender otra prueba en el caso y no tiene una relevancia independiente o directa; antecedentes generales; *res gestae*). Más aún, si la prueba realmente no es relevante no debería ser admitida sobre esta base. Véase *Federal Rule of Evidence* 402.

³⁸ Se hace referencia a este caso como *People v. Johnson* y está reproducido (con los nombres cambiados) en ALLEN, KURNS, SWIFT Y SCHWARTZ, 2006: 2-75.

³⁹ Del discurso de apertura de la fiscalía: «Sin importar la razón, el recluso Johnson, acusado en este caso, estaba de pie en su celda con la bandeja en su mano y tuvo una discusión con los oficiales acerca del paquete. El quería el paquete», *ibid.*: 6.

Del discurso de apertura de la defensa: «el acusado fue notificado de que un paquete había llegado —que él conocía de alguna manera el hecho de que un paquete le había sido enviado por su familia—. Pasó un período de tiempo sustancial [...] Finalmente, el día 28 él retuvo su bandeja, lo que reconozco que es una actitud desobediente», *ibid.*: 39.

Del alegato final de la fiscalía: «Por alguna razón, el recluso Johnson decidió que ese día había llegado al límite acerca del tema del paquete que decía no le estaban entregando», *ibid.*: 65.

Del alegato final de la defensa: «Johnson decidió que había llegado al límite. Estoy de acuerdo. Él marcó ese límite. ¿quiere que se haga algo, quiere ver a un sargento...?» Y resultó obvio del testimonio que si él seguía trazando ese límite en algún punto habría podido ver a un sargento. Esto no significa que Johnson se decidiera a salir y comenzar a agredir al oficial o los oficiales», *ibid.*: 38.

Múltiples testigos de cada parte testifican acerca de los detalles del paquete y de los procedimientos que existen en la prisión para recibirlos. Cada parte usa estas pruebas para apoyar su teoría: 1) el acusado se sentía frustrado y enojado por no haber recibido el paquete, retuvo su bandeja y embistió al guardia,⁴⁰ o 2) el acusado se sentía frustrado por no recibir el paquete, retuvo su bandeja y en respuesta el guardia le embistió (para vengarse o castigarle por su comportamiento)⁴¹. Ahora bien, todas estas pruebas aparentemente relevantes no parecen distinguir entre las dos teorías. Como en el ejemplo mencionado anteriormente sobre el «apoyo a candidatos liberales»⁴², no hay una base razonable para concluir que estas pruebas apoyan a una teoría sobre la otra. En otras palabras, la ratio de probabilidad es 1. Según la concepción de la ratio de probabilidad, esto implica que la prueba es irrelevante (y a fortiori no tiene valor probatorio) y, por tanto, debería haber sido excluida.⁴³ Esto implicaría además que se debería haber obligado al jurado a decidir quién inició el altercado sin darle ninguna información acerca de los hechos que lo precipitaron. La relevancia de esta prueba —prueba que se superpone en ambas teorías pero que no las distingue— es un caso de *reductio ad absurdum* del enfoque de la ratio de probabilidad. La prevalencia de este tipo de pruebas puede ser motivo de discusión, pero sin duda existen al menos algunos ejemplos. Y el fracaso del enfoque de la ratio de probabilidad para explicar dichas pruebas resulta problemático para la teoría.

Un segundo problema para los enfoques probabilísticos en el nivel micro tiene que ver con los intentos de cuantificar el *valor probatorio*. Éste es el problema de las clases de referencia.⁴⁴ El uso de una ratio de probabilidad para cuantificar el valor probatorio implica típicamente ubicar a la prueba dentro de una clase particular (de la cual hay datos disponibles). Considérese este ejemplo del tratado sobre prueba de MCCORMICK acerca de «un patrón conductual que se suponía característico de los niños que habían sido maltratados»:

Sila investigación determina que el comportamiento es igualmente común entre niños maltratados y no maltratados, entonces la ratio de probabilidad sería uno y la prueba de tal patrón no serviría para probar el abuso [...]. Y si fuera mil veces más común en el caso de niños abusados, su valor probatorio sería mucho mayor.⁴⁵

Estas conclusiones dependen de varios presupuestos adicionales (que pueden o no ser verdaderos). La prueba en cuestión es un comportamiento determinado. Incluso si este comportamiento fuera igualmente común en ambos grupos de niños, serviría de todas formas como prueba en un caso determinado

⁴⁰ *Ibid.*

⁴¹ *Ibid.*

⁴² Véase LEMPERT, 1977: 1025.

⁴³ *Federal Rules of Evidence*, 402.

⁴⁴ ALLEN Y PARDO, 2007.

⁴⁵ MCCORMICK, 1999.

si, por ejemplo, los niños abusados que tienen este comportamiento poseyeran alguna característica adicional, los niños no abusados no tuvieran esta característica y el niño particular en cuestión la poseyera (o no). De forma similar, incluso si el comportamiento fuera mil veces más probable en el caso de niños abusados, el valor probatorio podría ser mínimo si el niño poseyera (o no) una característica adicional que le ubicara en el grupo de los no abusados que tienen la conducta. En estos ejemplos, la prueba y las ratios de probabilidad se mantienen constantes, pero el valor probatorio puede variar dramáticamente. El valor probatorio no es entonces la ratio de probabilidad.

Para ilustrar de forma más contundente el problema de la clase de referencia, considérese este famoso ejemplo de TVERSKY y KAHNEMAN:

Un taxi se ve involucrado en un accidente de tráfico y huye del lugar. Dos compañías de taxis, La Verde y La Azul, operan en la ciudad. La prueba: 85 por 100 de los taxis en la ciudad pertenecen a la compañía Verde y 15 por 100 pertenecen a la Azul. ¿Cuál es la probabilidad de que el taxi involucrado en el accidente sea de la compañía Azul?⁴⁶

La respuesta inicial sería 0,15. Pero, ¿por qué? La distribución de taxis en la ciudad es solamente una de las clases de las cuales este evento es miembro. El problema de las clases de referencia surge porque la información depende de ubicar en una clase un evento o un elemento de prueba en particular, pero el evento o la prueba pertenecen a un número prácticamente infinito de conjuntos, cada uno con diferentes probabilidades. Por ejemplo, supóngase que el 85 por 100 de los taxis en la calle pertenecen a la compañía Azul y que el 85 por 100 de los taxis circulando a esa hora del día pertenecen a la compañía Verde, 85 por 100 de los taxis en el área de servicio (que incluye varios distritos) pertenecen a la compañía Azul, etcétera. ¿Cuál es la probabilidad de que el taxi del accidente pertenezca a la compañía Azul? Las inferencias basadas en la prueba de una clase de referencia particular implican una cantidad de supuestos auxiliares como, por ejemplo, la conveniencia y homogeneidad de esa clase particular, la distribución probable en otras clases y muchas otras posibles cuestiones. El resultado, en resumen, es que no podemos tomar los datos de una clase de referencia como estándar normativo con el cual cuantificar y juzgar el valor probatorio, ni como una explicación teórica del contenido del concepto.

2.2. El nivel macro

La restricción de nivel macro sobre la teoría de la prueba postula que una teoría de la prueba debería explicar cuándo las pruebas son suficientes para satisfacer estándares de prueba. Las concepciones probabilísticas en el ni-

⁴⁶ TVERSKY y KAHNEMAN, 1982.

vel macro son un lugar común. Típicamente, se propone o presupone que la «preponderancia de la prueba» equivale a una probabilidad mayor a 0,5; que la «prueba clara y convincente» equivale a una probabilidad mayor a 0,75 aproximadamente; y que «más allá de toda duda razonable» equivale a una de 0,9 o más⁴⁷. Y las pruebas resultan suficientes para confirmar un enunciado fáctico cuando la probabilidad de éste, dadas esas pruebas, supera el umbral de decisión. Esta es la interpretación más o menos estándar acerca de la que se ha escrito mucho. Como en el nivel micro, esta concepción captura algunos aspectos importantes de la cuestión pero, nuevamente, es también profundamente problemática como propuesta teórica o estándar normativo para alcanzar los fines subyacentes de los estándares (i. e., corrección y distribución del riesgo de error). Me centraré aquí en dos de los problemas de las concepciones probabilísticas.

El primer problema de las concepciones probabilísticas se refiere a la falta de guía en la aplicación de estándares y a la falta de criterios con los cuales determinar qué aplicaciones resultan razonables e irrazonables. Para aplicar un estándar de decisión probabilístico, uno debe también cuantificar las pruebas⁴⁸. Esto podría hacerse sobre la base de datos *objetivos*⁴⁹ o de creencias *subjetivas*. Para la mayoría de los elementos de prueba de cualquier proceso, los datos objetivos no existen⁵⁰. Si los hubiera y tales estándares fueran aplicables, entonces la única decisión razonable sería la respuesta objetiva, lo que haría a los juzgadores de los hechos ampliamente superfluos. Por tanto, debemos confiar en gran medida en evaluaciones probabilísticas subjetivas. Desafortunadamente, los estándares de decisión probabilísticos ofrecen poca orientación acerca de cómo proceder. Suponga que usted es parte de un jurado en un caso civil donde se aplica el estándar de la preponderancia y usted está tratando de decidir, luego de escuchar a varios testigos, si la probabilidad de un enunciado fáctico supera el 0,5. ¿Cómo haría esto? ¿Cerraría sus ojos y pensaría en un número? Suponga que usted llega a la conclusión de que la probabilidad resultante es 0,6 y que otro de los miembros del jurado dice que es 0,2. ¿Son ambas razonables? ¿Lo es 0,99? ¿Y 0,01? ¿Cómo lo podemos saber?⁵¹.

El segundo problema surge de la naturaleza no comparativa de las concepciones probabilísticas que se centran en si las pruebas presentadas por la parte que tiene la carga de la prueba sobrepasa un determinado umbral. Si los estándares son entendidos de esta manera, se verían frustrados los fines

⁴⁷ Véase BELL, 1987.

⁴⁸ Esto se podría hacer de manera precisa o con juicios vagos acerca de si las pruebas sobrepasan cierto umbral.

⁴⁹ Pero nótese que los datos objetivos estarían todavía sujetos al problema de la clase de referencia.

⁵⁰ Aún si existieran, la complejidad computacional sería demasiado grande para la mayoría de los juicios. Véase ALLEN, 1993: 444.

⁵¹ Más aún, no hay razones para pensar que las creencias subjetivas serían conducentes a la verdad. Véase CLERMONT, 2009.

subyacentes de las reglas. Las reglas probabilísticas comparan la probabilidad de una proposición con su negación, pero deberían estar comparando los enunciados fácticos alternativos presentados por cada parte. Supóngase que la regla de la preponderancia intenta, como comúnmente se cree, minimizar los errores y distribuir el riesgo de error de forma más o menos pareja entre demandante y demandado. Ahora supóngase que un demandante ofrece una teoría de lo sucedido para la que un jurado razonable concluye que la probabilidad es 0,4 y que el demandado ofrece una teoría del caso para la que se concluye que la probabilidad es 0,2.⁵² El estándar de «mayor que 0,5» implica que el demandante debería perder—incluso cuando la probabilidad de que su versión sea verdadera es dos veces mayor que la versión alternativa del demandado—. Esto no tiene sentido si lo que el sistema jurídico está buscando con el estándar de la preponderancia es la minimización de errores (en cuyo caso debería preferir a la más probable) o la distribución equitativa del riesgo de error (el demandante no debería cargar con todo el riesgo de error en un espacio de probabilidad desconocido). El error está en presumir que cualquier posibilidad desconocida favorece al demandado (o a la parte que no tiene la parte de la prueba). Esto resulta inconsistente con la distribución equitativa del riesgo de error.⁵³

2.3. Integración

Los estándares de prueba probabilísticos también se enfrentan con un problema basado en el requisito de integración. En este nivel se encuentra la segunda limitación que tienen las concepciones probabilísticas subjetivas. Existe una inconsistencia entre los estándares de prueba probabilísticos del nivel macro, basados en probabilidades subjetivas y las concepciones probabilísticas convencionales del nivel micro de la relevancia y el valor probatorio. En el nivel micro, la relevancia y el valor probatorio dependen de algo más que de las evaluaciones subjetivas de cada juzgador de los hechos.⁵⁴ Cierta prueba puede resultar relevante incluso si los miembros del jurado creen que es irrelevante y las pruebas pueden tener un valor probatorio diferente al que creen los miembros del jurado. Más aún, la aplicación de cualquier estándar de decisión es esencialmente una evaluación del valor probatorio de las pruebas como un todo, lo que presupone la posibilidad de una aplicación razonable e irrazonable. Si es posible hablar del valor probatorio de las pruebas como un todo,

⁵² Para los fines de este ejemplo, estoy dejando de lado otros problemas de estas concepciones.

⁵³ Esto es así sin siquiera considerar los distintos problemas de *conjunction*, que podrían variar aún más el riesgo de error de maneras no deseadas. Los problemas de *conjunction* aparecen porque los estándares de decisión se aplican sobre elementos individuales y no sobre el caso como un todo. Véase ALLEN Y JERH, 2003.

⁵⁴ Y si no fuera así, entonces los jueces no tendrían bases para excluir pruebas por motivos de relevancia o por dilaciones indebidas, pérdida de tiempo o redundancia.—*Federal Rule of Evidence* 403.—

entonces las concepciones subjetivas de la probabilidad en el nivel macro no logran explicar el proceso probatorio en este nivel y no logran ofrecer la orientación normativa necesaria. Y la falta de datos objetivos para la mayoría de los elementos de prueba significa que las concepciones probabilísticas *objetivas* del estándar de prueba no pueden salvar esta brecha teórica. Es necesario algo adicional—y más plausible—. Si, por el contrario, un teórico de la probabilidad argumenta que de hecho no hay nada más allá de las evaluaciones subjetivas para el valor probatorio en el nivel macro, entonces, en el nivel micro, el teórico debe aceptar que la relevancia y el valor probatorio también dependen únicamente de evaluaciones subjetivas.⁵⁵ Esta concesión no lograría explicar la relevancia y el valor probatorio. En cualquier caso, el teórico probabilístico está en problemas.

3. UNA TEORÍA EXPLICATIVA DE LA PRUEBA

Cualquier concepción teórica acerca de una práctica social tan rica y compleja como la prueba jurídica enfrentará probablemente ciertas dificultades. Entonces, en ausencia de una alternativa robusta ¿quizás las concepciones probabilísticas son lo mejor que podemos ofrecer?⁵⁶ Concluiré presentando el esquema de una concepción alternativa y explicando cómo ésta satisface mejor los tres requisitos teóricos articulados en el apartado 1.⁵⁷

Esta concepción alternativa examina las relaciones explicativas que hay entre los enunciados fácticos y las pruebas. En lugar de enfocarse en las relaciones probabilísticas—sea la probabilidad de un enunciado fáctico (o su negación) dadas ciertas pruebas o la probabilidad de ciertas pruebas dado un enunciado fáctico (o su negación)—la concepción explicativa examina las explicaciones alternativas de las pruebas y los hechos presentadas por las partes y desarrolladas por los juzgadores de los hechos. Más específicamente, la cuestión crucial es si las explicaciones particulares presentadas por las partes o desarrolladas por los juzgadores de los hechos, de ser verdaderas, explicarían mejor o peor las pruebas y los hechos subyacentes (en comparación con las alternativas). Según esta concepción, la prueba jurídica es un proceso de abducción o inferencia a la mejor explicación.⁵⁸ Este proceso inferencial tiene dos etapas: generando explicaciones potenciales y luego seleccionando una como la mejor o como aceptable, dependiendo del estándar.

⁵⁵ O esta aparente inconsistencia debe explicarse.

⁵⁶ Y, como he concedido, las concepciones probabilísticas de hecho capturan e iluminan algunos aspectos importantes de la estructura y la naturaleza de la prueba jurídica.

⁵⁷ La concepción explicativa ha sido desarrollada por mí y por Ron ALLEN. Véase PARDO Y ALLEN, 2008. El foco en el criterio explicativo brinda una base epistemológica para la teoría de la «plausibilidad relativa» desarrollada previamente por el profesor ALLEN. Véase ALLEN, 1994.

⁵⁸ Para un resumen general, véase LITTON, 2004.

dar de decisión⁵⁹. Por ejemplo, la calidad de una explicación necesaria para satisfacer el estándar «más allá de toda duda razonable» o «prueba clara y convincente» será mayor que en el caso del estándar de la preponderancia de la prueba.

3.1. El nivel micro

Según esta concepción, la relevancia y el valor probatorio dependen de si las pruebas forman parte de la explicación presentada por una parte o si son un ataque a la explicación de la otra parte. A diferencia de la teoría de la ratio de probabilidad en particular o de las concepciones probabilísticas en general, la concepción explicativa da cuenta de aquellas pruebas que resultan relevantes pero que no distinguen entre las distintas hipótesis planteadas por las partes. Una prueba resulta relevante si conforma la explicación proporcionada por una parte (o presentada por la otra parte como desafío a esa explicación), independientemente de si también conforma la explicación presentada por la otra parte. Y su valor probatorio depende del papel que tenga en las explicaciones ofrecidas por las partes o en la distinción entre ellas.

Esta concepción explica, entonces, el ejemplo presentado en el apartado 2 respecto a la prueba *superpuesta* que sirve de apoyo a ambas partes pero no distingue entre las distintas hipótesis planteadas. Las pruebas sobre el paquete en el caso transcrito conforman la explicación ofrecida por cada parte procesal acerca de lo ocurrido y resultan por tanto relevantes. Esta concepción logra escapar a las limitaciones del problema de la clase de referencia al otorgar valor probatorio a la prueba de acuerdo con su pertenencia a una clase particular. La pregunta crítica respecto a qué explican las pruebas logra aislar lo importante de la prueba sin necesidad de cuantificar sobre la base de las clases de referencia. Por ejemplo, en el caso del abuso infantil la cuestión es qué explica mejor el comportamiento del niño (y la prueba estadística será relevante a la hora de responder esta pregunta). En el ejemplo del taxi de la compañía Azul, por el contrario, la información estadística tiene una conexión explicativa menor. Aunque resulta relevante, esta falta de conexión explicativa implica que la prueba no tiene un valor probatorio siquiera cercano al que ha sido sugerido⁶⁰.

⁵⁹ Véase PARDO y ALLEN, 2008. Las explicaciones involucran tanto «actos de explicar» como las explicaciones en sí, que se refieren a las relaciones epistémicas entre enunciados. Las explicaciones ofrecen típicamente respuestas o comprensiones acerca de otro enunciado; por ejemplo, por qué sucedió algo. Sobre la naturaleza de las explicaciones, véase LIPTON, 2004: 21-29; ACHINSTEIN, 1983: 74-102.

⁶⁰ Véase PARDO y ALLEN, 2008 (criticando la conclusión de TVERSKY y KARNEMAN de que los sujetos erran cuando descuentan esta prueba).

3.2. El nivel macro

Según esta concepción, las reglas de decisión se pueden articular sobre la base de criterios explicativos. Por ejemplo, de acuerdo con la regla de la preponderancia, los juzgadores de los hechos deberían inferir la mejor explicación y la parte que tiene la carga de la prueba debería ganar si esa explicación incluye los elementos formales y debería perder si la mejor explicación no incluye uno o más de estos elementos⁶¹. Esta podría ser una mejor manera de poner en práctica los fines de minimización de errores y de distribución igual del riesgo de error entre las partes —las mejores explicaciones son con un mayor grado de probabilidad verdaderas y cada una tiene un riesgo de error similar—. Más aún, como el estándar es comparativo, ninguna de las partes se beneficia del espacio probabilístico desconocido.

El estándar de la preponderancia explicativa resulta ampliamente consistente con el *Story Model* como modelo empírico, ya que varios de los criterios que hacen persuasiva a una *historia*, también hacen que una explicación sea mejor y viceversa (*i. e.*, consistencia, cobertura, completitud, simplicidad, ausencia de lagunas y encaje con el conocimiento y la experiencia acumulados). Sin embargo, existen diferencias importantes: las explicaciones pueden ser generales⁶² y disyuntivas⁶³, dependiendo de cómo las partes elijan contrastar lo sucedido, y no historias singulares, específicas e integradas como las postuladas por el *Story Model*⁶⁴. Por otra parte, bajo estándares de prueba más altos, la calidad de una explicación necesaria para satisfacer los estándares logra una ventaja simplemente siendo mejor que sus alternativas. Por ejemplo, bajo el estándar «prueba clara y convincente» la explicación debe ser sustancialmente mejor que sus alternativas y para el estándar «más allá de toda duda razonable» sería si existe una explicación plausible que incluya los elementos formales y no hay una explicación plausible compatible con la inocencia (o que no logre incluir uno o más elementos). El aumento en la calidad de una explicación necesaria para satisfacer estos estándares desplaza el riesgo de error de una manera consistente con los fines de las reglas y debería implementar mejor estos fines en comparación con las valoraciones subjetivas de la probabilidad⁶⁵.

⁶¹ Este enfoque también resuelve cualquier problema potencial acerca de la *conjunción*.

⁶² Por ejemplo, en un caso de *res ipsa loquitur* una explicación general puede ser suficiente.

⁶³ Por ejemplo, una explicación acerca de que «X, Y o Z ocurrieron» puede ser suficiente si esta explicación disyuntiva es mejor que la explicación alternativa (la cual puede ser o no disyuntiva).

⁶⁴ Cuando ambas divergen, el criterio explicativo provee una guía normativa.

⁶⁵ Aunque los estándares no son tan precisos como los umbrales de la probabilidad cardinal, estos pueden ser entendidos mejor por los juzgadores de los hechos y encajan mejor con los procesos de razonamiento de los jurados. Estas características podrían hacer que las instrucciones a los jurados basadas en explicaciones sean más efectivas que las actuales. Para profundizar acerca de este asunto, véase PARDO, 2009.

Los criterios explicativos que subyacen a esta concepción también ofrecen alguna guía y criterios para los juzgadores de los hechos y los tribunales de instancias superiores, más allá del *Story Model* como modelo empírico y de las concepciones teóricas probabilísticas. Existen maneras objetivas de examinar y discutir las relaciones explicativas que existen entre las pruebas y los enunciados fácticos alternativos, aunque sin duda habrá desacuerdos en algunos casos acerca de qué explicaciones son mejores o peores o qué hace que una explicación sea mejor o peor⁶⁶. En igualdad de condiciones, una explicación consistente es mejor que una que sea contradictoria; una explicación que dé cuenta de mayor cantidad de pruebas y de los elementos de prueba más importantes será mejor que aquella que no pueda explicarlas o dar cuenta de ellas; una explicación compatible con el conocimiento y la experiencia acumulados y que no requiera asunciones insólitas será mejor que aquella que es inconsistente con el conocimiento y la experiencia acumulados y necesita de estos supuestos irrealistas. Estos criterios explicativos están conectados con cuestiones doctrinales más profundas en los ámbitos civil y penal: ofrecen criterios para distinguir entre aplicaciones razonables e irrazonables y también pueden ofrecer una carga de explicación y un vocabulario para los tribunales revisores⁶⁷.

3.3. Integración

Los niveles micro y macro de la teoría explicativa son claramente coherentes entre sí, y esto proporciona otra virtud a esta concepción teórica. Las explicaciones como un todo animan al nivel macro, y la calidad de estas explicaciones depende de la relevancia y del valor probatorio de las pruebas en el nivel micro. De la misma forma, la relevancia y el valor probatorio en el nivel micro dependen de si una prueba forma parte de la explicación ofrecida por una parte en el nivel macro (o un ataque a la explicación de la otra parte), y la importancia de las pruebas en el nivel micro depende de la importancia de las pruebas para estas explicaciones del nivel macro. Los mismos criterios explicativos ofrecen una guía, los límites y una base para la evaluación en ambos niveles.

⁶⁶ Véase CLERMONT, 2009: 482 (a través de su estándar de prueba el derecho busca forzar al juzgador de los hechos, en el paso de la decisión final, a que relacione su estado mental interno con el «mundo real exterior»). Las consideraciones explicativas fuerzan a conectar estos estados mentales con características de la prueba en el «mundo real exterior». Además, algunos ámbitos del derecho —por ejemplo, la discriminación laboral— ya incluyen marcos para evaluar las explicaciones alternativas. Véase *Reeves v. Sanderson Plumbing Prods., Inc.*, 530 U. S. 133, 149-150 (2000).

⁶⁷ Un tribunal que dicta un fallo sumario a favor del demandado en un juicio civil o que le concede una sentencia favorable tras declarar la cuestión de puro derecho, por ejemplo, debería justificar su decisión acerca de por qué un jurado razonable podría no considerar a la explicación del demandante sobre la prueba y de los hechos como más plausible que la del demandado.

BIBLIOGRAFÍA

- ACHENSTEIN, P., 1983: *The Nature of Explanation*, Oxford: Oxford University Press.
- ALLEN, R. J., 1993: «Constitutional Adjudication, the Demands of Knowledge, and Epistemological Modesty», *Northwestern University Law Review*, 88.
- 1994: «Factual Ambiguity and a Theory of Evidence», *Northwestern University Law Review*, 88.
- ALLEN, R. J., y JEHU, S. A., 2003: «Burden of Proof in Civil Cases: Algorithms vs. Explanations», *Michigan State Law Review*, 2003.
- ALLEN, R. J.; KUHNS, R. B.; SWIFT, E., y SCHWARTZ, D. S., 2006: *Evidence: Text, Problems and Cases*, 4ª ed., New York: Aspen Publishers.
- ALLEN, R. J., y LAUDAN, L., 2008: «Deadly Dilemmas», *Texas Tech Law Review*, 41.
- ALLEN, R. J., y LETTER, B., 2001: «Naturalized Epistemology and the Law of Evidence», *Virginia Law Review*, 87.
- ALLEN, R. J., y PARDO, M. S., 2007: «The Problematic Value of Mathematical Models of Evidence», *Journal of Legal Studies*, 36.
- BEHL, R. S., 1987: «Decision Theory and Due Process: A Critique of the Supreme Court's Lawmaking for Burdens of Proof», *Journal of Criminal Law and Criminology*, 78.
- BURZ, K., 2010: «We Don't Want to Hear It: Psychology, Literature and The Narrative Model of Judging», *University of Illinois Law Review*, 2010.
- CLERMONT, K. M., 1987: «Procedure's Magical Number Three: Psychological Bases for Standards of Decisions», *Cornell Law Review*, 75.
- 2009: «Standards of Proof Revisited», *Vermont Law Review*, 33.
- COHEN, L. J., 1977: *The Probable and the Provable*, Oxford: Clarendon Press.
- DAVIS, D., y FOLLETTE, W. C., 2002: «Rethinking the Probative Value of Evidence: Base Rates, Innate Profiling, and the "Postdiction" of Behavior», *Law and Human Behavior*, 26.
- HAYEK, A., 2009: *Interpretations of Probability*, *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/probability-interpret/> (última consulta realizada el 7 de septiembre de 2011).
- KAPTEIN, H.; PRAKKEN, H., y VERHEIJ, B. (eds.), 2009: *Legal Evidence and Proof: Statistics, Stories, and Logic*, Farnham: Ashgate Publishing.
- KAYE, D. H., 1995: «The Relevance of "Marching" DNA: Is the Window Half Open or Half Shut?», *Journal of Criminal Law and Criminology*, 85.
- KOELER, J. J., 2002: «When Do Courts Think Base Rate Statistics are Relevant?», *Jurimetrics Journal*, 42.
- KOLMOGOROV, A., 1956: *Foundations of the Theory of Probability*, New York: Chelsea Publishing Company.
- LEMPERT, R. O., 1977: «Modeling Relevance», *Michigan Law Review*, 75.
- LIPSON, P., 1991: *Inference to the Best Explanation*, 2ª ed., London: Routledge, 2004.
- MCCORMACK, Ch. T., 1954: *On Evidence*, 5ª ed., J. W. STRONG (ed.), St. Paul, Minnesota: West Publishing Co., 1999.
- MCCNAUGHTON, J., 1955: «Burdens of Production of Evidence: A Function of the Burden of Persuasion», *Harvard Law Review*, 68.

- PARDO, M. S., 2009: «Second-Order Proof Rules», *Florida Law Review*, 61.
- PARDO, M. S., y ALLEN, R. J., 2008: «Juridical Proof and the Best Explanation», *Law and Philosophy*, 27.
- PARK, R. C., y SAKS, M. J., 2006: «Evidence Scholarship Reconsidered: Results of the Interdisciplinary Turn», *Boston College Law Review*, 47.
- PENNINGTON, N., y HASTIE, R., 1991: «A Cognitive Model of Juror Decision Making: The Story Model», *Carozo Law Review*, 13.
- PUNDIK, A., 2008, «What is Wrong with Statistical Evidence? The Attempts to Establish an Epistemic Deficiency», *Civil Justice Quarterly*, 27.
- RISINGER, D. M., 2010: «Introduction, Bayes Wars Redivivus. An Exchange», *International Commentary on Evidence*, 8 (1).
- STEIN, A., 2005: *Foundations of Evidence Law*, Oxford: Oxford University Press.
- THAYER, J. B. [1898]: *A Preliminary Treatise on Evidence at the Common Law*, Boston: Elibron Classics, 2005.
- TILLERS, P., 2011: «Trial by Mathematics. Reconsidered», *Law, Probability & Risk*, 10 (3).
- TVERSKY, A., y KAHNEMAN, D., 1982: «Evidential Impact of Base Rates», en KAHNEMAN, D.; SLOVIC, P., y TVERSKY, A. (eds.), *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press.
- TWINING, W., 1990: *Rethinking Evidence: Exploratory Essays*, 2.ª ed., Evanston, Illinois: Northwestern University Press, 1994.
- VIDMAR, N., y HANS, V. P., 2007: *American Juries: The Verdict*, New York: Prometheus Books.
- ZAMIR, E., 2010: *Loss Aversion and Law's Formation*, en <http://papers.ssrn.com/abstract=1638527> (última visita realizada el 7 de septiembre de 2011).

V

LA ELEMENTAL ARITMÉTICA EPISTÉMICA DEL DERECHO II: LOS INAPROPIADOS RECURSOS DE LA TEORÍA MORAL PARA ABORDAR EL DERECHO PENAL *

Larry LAUDAN

Universidad Nacional Autónoma de México y University of Texas

Quiero comenzar agradeciendo colectivamente a los participantes, a Jordi FERRER y a los demás organizadores de este prometedor evento. Es tan raro como importante, para los epistemólogos y para los teóricos de la prueba, reunirse bajo un mismo techo a discutir sobre epistemología jurídica. Por ello, estoy encantado de estar aquí, participando en este debate.

La instrucción que recibí de los organizadores fue concentrarme en cuestiones sobre los estándares de prueba. Aunque haré algunas anotaciones al respecto, ese no será el enfoque principal. Esto porque a lo largo de la última media docena de años he escrito una enorme cantidad de cosas sobre el estándar de prueba penal. Muchos de los que están aquí han oído o leído mis puntos de vista sobre ese tema, al punto de que probablemente los conocen tan bien como yo. Por tanto, no voy a aburrirles haciéndoles oír argumentos que ya les son familiares. Por el contrario, me referiré a la cuestión de si la filosofía moral tiene, como insisten muchos expertos en ética, recursos conceptuales que permitan explicar patrones de prueba y patrones de castigo en el derecho.

* Traducción de Maximiliano ARAMBURU.

Ustedes bien pueden suponer qué es lo que tienen en común, si es que tienen algo, la prueba judicial y la teoría de la pena. Trataré de persuadirles de que tienen mucho en común, y de que el hecho de que ambos requieran un sistema métrico para evaluar la magnitud de diferentes tipos de daño, es una de dichas coincidencias.

1. PERSPECTIVA GENERAL

Hay una idea habitual que consiste en que las teorías morales, y especialmente las teorías deontológicas de corte kantiano, pueden, deben y de hecho lo hacen, proveer la clave conceptual que apuntale al derecho penal. Este último suele verse, en efecto, como filosofía moral aplicada. Esta concepción no se limita a los expertos en ética. Los constitucionalistas y los filósofos del derecho rutinariamente buscan una justificación o fundamentación en la teoría moral para cada aspecto del derecho penal. Si se quieren ejemplos recientes de la tendencia a fundamentar al derecho probatorio en la teoría moral, basta con mirar los recientes libros de Alex STEIN y Ho HACK LAI, o los esfuerzos del penalista Paul ROBINSON por fundamentar al derecho penal en teorías morales sobre la causación de daños.

En cuanto a mí, soy escéptico con respecto a la idea de que podamos derivar las principales características del sistema jurídico, especialmente las de su parte penal, de la filosofía moral tradicional. No es simplemente porque los sistemas jurídicos existentes exhiban características incompatibles con la deontología convencional (de lo cual examinaremos ejemplos más adelante en este trabajo); aquéllos, después de todo, podrían cambiarse para poner los sistemas jurídicos a tono con nuestras intuiciones morales cuando se considerase necesario o deseable. Mucho más revelador es el hecho de que haya algunas partes de cualquier sistema jurídico imaginable que son al mismo tiempo *ineliminables e importantes*, y que no parecen ser subsumibles bajo una sombrilla deontológica.

En este trabajo identificaré dos áreas clave de la teoría del delito y de la pena que, como mostraré, *no* se derivan de una teoría moral deontológica o retributiva. En la primera sección, mostraré que las teorías morales deontológicas no sirven de nada frente al desafío de justificar o formular un estándar de prueba para los juicios penales. En la segunda sección, mostraré que las teorías morales de la pena que existen (especialmente el retribucionismo), de hecho no son tal cosa, en tanto carecen de recursos para justificar o deducir un escalafón de penas o al menos una especificación no arbitraria de merecimientos de pena para cada delito del que se trate. En otras palabras, aunque las teorías deontológicas subrayan la importancia de que los acusados condenados reciban su merecido, carecen de equipamiento para indicar cuál es ese merecido para un delito

determinado cualquiera. Estos dos temas constituyen una agenda ambiciosa, así que permítaseme acometerla de inmediato.

2. PARTE I: LOS FORCEJOS DE LA DEONTOLOGÍA CON EL ESTÁNDAR DE PRUEBA

Ningún procedimiento penal que merezca su nombre puede tener lugar en ausencia de un estándar de prueba. Dicho estándar sirve como regla de decisión para que el juzgador de los hechos alcance el veredicto del caso. Sin un estándar de prueba el veredicto mismo no estará justificado y cualquier declaración de culpabilidad será injusta, a menos que pueda mostrarse que las pruebas presentadas contra el acusado satisfacen el estándar probatorio preestablecido.

Las teorías deontológicas tienen una estructura curiosa: por una parte, son inflexibles en cuanto a que un acusado no puede ser castigado de forma justa a menos que estemos seguros de que cometió el delito. Por la otra, esas teorías carecen de recursos para decirnos qué tan seguros debemos estar antes de decir que una condena está justificada. En pocas palabras, su demanda moral de que estemos seguros de la culpabilidad les exige cumplir una promesa epistémica que han formulado de tiempo atrás, pero nunca han satisfecho. En esta sección quiero explorar esa tensión en algún detalle.

Los deontologistas suponen amplia y correctamente que el estándar de prueba debe depender (al menos en parte) de la determinación que se haga de los correspondientes daños causados por los errores que pueden cometerse en los veredictos judiciales. Una condena errónea inflige un grave daño al acusado inocente, a quien se le niega su libertad y cuya reputación queda indebidamente arruinada. Una falsa absolución, igualmente, inflige daño moral no sólo al acusado que (como diría KANT) tiene derecho a recibir su merecido, sino también a la víctima del delito, quien no obtiene un sentimiento de haber alcanzado justicia si el verdadero agresor es absuelto. Hay consenso universal entre los deontologistas (y probablemente están en lo correcto) acerca de que el primer daño es más atroz que el segundo. Y dado que lo es, sostiene, es razonable que pidamos algo cercano a la certeza antes de que consideremos una condena como justa. Por tanto, dicen, queremos un estándar de prueba que sea muy exigente.

El problema con este argumento, por supuesto, es que el hecho de que una falsa condena represente un daño moral mucho mayor que una falsa absolución, únicamente respalda la afirmación de que el estándar de prueba en un proceso penal debe ser mayor que el de la preponderancia de la prueba, pero nada más que eso. Si no se dice *cuidado peor* es la falsa condena sobre la falsa absolución, lo único que podemos inferir de la idea según la cual la primera es peor que la segunda, es que el estándar de prueba debe ser el de más probable

que su negación. Entonces, en este punto, no podemos adoptar un estándar que se aproxime a la cuasicertera.

Durante mucho tiempo los deontologistas han tratado de llenar este vacío con conjeturas acerca de los daños correspondientes a esos dos errores. Lo que generalmente se conoce como la ratio de Blackstone (que es una hipótesis acerca de la ratio admisible de falsas absoluciones frente a falsas condenas en un conjunto amplio de juicios) se establece a menudo como 10:1. Esto se interpreta generalmente como la garantía de un estándar de prueba de alrededor del 90 por 100, pues se cree —equivocadamente, por supuesto— que ese estándar produciría errores que ilustrarían esta ratio¹. Eso en sí mismo no es un disparate, como han sostenido repetidamente Ron ALLEN y otros. Sin embargo, ése no es el punto que quiero demostrar aquí. Lo que quiero señalar es la absoluta arbitrariedad de la idea de que una falsa condena es considerablemente más lesiva que una falsa absolución. En general, las teorías deontológicas tienen las herramientas conceptuales para establecer que un daño es más o menos lesivo que otro. Lo que generalmente no pueden hacer —excepto por un acto de pura y arbitraria intuición— es determinar cuánto más lesivo es un daño con respecto a otro.

Volvamos brevemente a la ratio de Blackstone. Aunque se ha debatido mucho desde la Edad Media, pensadores eminentes no han logrado alcanzar un consenso acerca de qué valor debe tener. Maimónides insistía en que era mejor absolver a mil acusados culpables que condenar a un inocente. Benjamin Franklin opinaba que la ratio debería ser de alrededor de cien a uno. Matthew Hale, famoso jurista inglés del siglo XVII, sugería que cinco a uno podría ser lo correcto. Voltaire afirmó que la ratio debería ser de dos a uno. Interesantemente, ninguno de estos valores propuestos para la ratio fue debatido ni defendido, excepto para decir que las falsas condenas eran más lesivas que las falsas absoluciones. Más preocupante aun, estas estimaciones difieren en un factor de 500, lo que sugiere que no hay ni siquiera uniformidad en las intuiciones, y ni hablar de argumentos razonados, para sostener una versión de la ratio de Blackstone por encima de otra. Cada una de estas versiones de la ratio de Blackstone fue defendida de manera simple por sus proponentes como autoevidente, tal como puede serlo una mera intuición.

Pese a su arbitrariedad, los valores propuestos para la ratio y, por tanto, para el estándar de prueba penal, están sobre la mesa. Si Voltaire tiene razón,

¹ Este método para el establecimiento del estándar de prueba, que se ha convertido en un lugar común, fue propuesto originalmente por John KAPLAN en 1968. Sostiene que el estándar de prueba está dado por

$$\text{SoP} = 1/[1 + \text{HFa}/\text{HFc}]$$

siendo HFa el daño producido con una falsa absolución y HFc el daño producido con una falsa condena (véase KAPLAN, 1967: 1071 y ss.). Un método mucho más sólido para generar el estándar, fue desarrollado poco después por Laurence TRIBE.

el estándar de prueba penal debería ser de alrededor del 65 por 100. Si Maimónides está en lo correcto, debería ser del 99.9 por 100. La cuestión es que, a falta de consenso sobre el valor preciso de la ratio de Blackstone y a falta de todo argumento posible para establecer esa ratio en un punto y no en otro, un deontologista no puede pretender fundamentar un estándar de prueba a partir de su opinión personal con respecto a los costos vinculados a los veredictos erróneos. Ciertamente, este enfoque particular de los deontologistas para definir un estándar de prueba nunca llega a cerrarse, pues depende de tener un sistema métrico para la determinación del grado de lesividad correspondiente y los deontologistas no tienen instrumentos para participar en ese juego, excepto sus ingenuas e inmotivadas intuiciones.

Un análisis muy diferente ha sido propuesto por DWORKIN y merece nuestra atención aunque sea sólo brevemente. DWORKIN sostiene que los acusados inocentes tienen un derecho moral a no ser condenados. Si son condenados, el daño causado con esa condena y la pena consiguiente, se agrava por el daño moral que se causa al negar a los acusados inocentes el derecho a ser hallados como no-culpables. De acuerdo con ello, dice DWORKIN, debemos reconocer un derecho moral que proteja al acusado de una falsa condena. Ese derecho, dice, es el derecho a ser juzgado según el estándar «más allá de toda duda razonable». ¿Por qué ese estándar en particular? Porque, dice DWORKIN, representa lo máximo que puede hacer el sistema jurídico para proteger al acusado de una falsa condena.

Esta idea, según la cual lo que justifica el uso del estándar «más allá de toda duda razonable» es que representa la máxima protección contra las falsas condenas que puede presentar el sistema jurídico, está ampliamente difundida tanto entre filósofos morales como entre la academia jurídica estadounidense. Rinat KITAI insiste: «La prueba más allá de toda duda razonable se asume como el más alto estándar de prueba realista dentro del ámbito del conocimiento humano»². Y añade:

El deber impuesto al Estado [por el contrato social] [...] origina el derecho individual, absoluto e inalienable [...] a no ser condenado a menos que la culpabilidad haya sido probada más allá de toda duda razonable.

Alex STEIN suscribe el mismo punto de vista:

Un sistema jurídico puede condenar justificadamente a una persona sólo si hizo su mejor esfuerzo para protegerle del riesgo de una condena errónea³.

Hace treinta años Laurence TRIBE cantó la misma canción:

El sistema [de justicia penal], en efecto, no autoriza la imposición de penas cuando el juzgador reconoce una duda cuantificable sobre la culpa-

² KITAI, 2003: 1163.

³ STEIN, 2005: 175.

bilidad del acusado. Por el contrario, el sistema [...] insiste en una aproximación tan cercana a la certeza cuando sea realizable atendidas las circunstancias.⁴

El estruillo también suena conocido: el Estado justo debe hacer todo lo que esté a su alcance para evitar la falsa condena de una persona inocente y una parte clave de ese proceso implica adoptar un estándar de prueba que sea tan exigente como lo permitan los esfuerzos humanos.

Desafortunadamente, este análisis está doblemente equivocado. El estándar «más allá de toda duda razonable» no representa lo mejor que podemos hacer para proteger a los acusados inocentes de una condena, si es que ese fuese nuestro deber. Y más importante aún, se puede demostrar que no es nuestro deber hacer todo lo que sea humanamente posible para evitar las falsas condenas. Miremos estos puntos.

Si estuviésemos inequívocamente obligados a proteger a toda costa al acusado inocente, podríamos, por ejemplo, aumentar el tamaño del jurado; insistir en que los miembros del jurado no tuvieran duda alguna —ni siquiera dudas irracionales— sobre la culpabilidad del acusado⁵; podríamos insistir en que, para una condena, el juez deba coincidir en la decisión con los jurados; podríamos excluir las confesiones, sobre la base de que algunas son falsas; y así podría continuarse. En pocas palabras, literalmente no hay límite en lo lejos que podríamos llegar a hacer para evitar una falsa condena. La razón por la que no hacemos ninguna de esas cosas es porque hacerlas impondría un costo inaceptable de falsas absoluciones, llevando a una situación en la que virtualmente ningún acusado en un juicio penal sería condenado. Y esto nos lleva, por supuesto, a la segunda razón por la cual la posición de DWORKIN y STEIN es equivocada. El Estado no tiene un deber de hacer todo lo humanamente posible para evitar una falsa condena. Si tomamos ese deber literalmente, tendríamos que imponer una regla según la cual los juicios penales no podrían terminar con una condena, en la medida en que, tratándose de condenas, algunas serán falsas. Es inevitable. En consecuencia, es un ingenuo consejo de perfección imaginar que nuestro objetivo debe ser el diseño de juicios de forma tal que hayamos hecho nuestro mejor esfuerzo por reducir el riesgo de falsas condenas. Por esas razones, ser juzgados de acuerdo con el estándar más allá de toda duda razonable no es, ciertamente, un derecho moral. Ni tiene el acusado inocente un derecho a que hagamos todo lo humanamente posible por evitar una falsa condena.

⁴ TRIBE, 1971: 1318. ¿Realmente cree Tribe que si el juez de los hechos está persuadido de que la probabilidad de que el acusado sea culpable es del 99,999 por 100, lo que se impone es la absolución? El hecho de que una duda sea cuantificable no tiene nada que ver con la cuestión de su razonabilidad o irrazonabilidad.

⁵ Esta disparatada propuesta ha sido defendida realmente por varios académicos como el estándar aplicable a los casos de pena de muerte. Erik LILLQUIST se ha dado juiciosamente a esta tarea. LILLQUIST, 2005.

De conformidad con lo anterior, fracasa cada una de las tres principales rutas argumentativas que los deontólogos han propuesto para llegar a un estándar de prueba. La meta afirmación de que las falsas condenas son peores que las falsas absoluciones no nos lleva a un estándar de prueba claro. La afirmación de que las falsas condenas son n veces peores que las falsas absoluciones, tampoco nos llevaría a dicho estándar, dado que cualquier valor que asignemos a n , sería arbitrario e inmotivado. Y, finalmente, el esfuerzo por justificar el estándar «más allá de toda duda razonable» aduciendo que representa el mejor esfuerzo que los mortales podemos hacer por minimizar las falsas condenas, es simplemente falso. Si no fuésemos tan cerrados, podríamos hacer mucho más de lo que hace el estándar «más allá de toda duda razonable» para proteger a los acusados inocentes.

He sostenido hasta aquí que las teorías deontológicas estándares carecen de los recursos conceptuales para llevarnos a deducir un estándar de prueba justo y razonable. Pero, tristemente, ese no es el final del problema sino apenas su comienzo. En cada procedimiento penal, desde el momento del arresto inicial, pasando por las decisiones sobre la caución, la admisibilidad de pruebas, la jurisdicción, la competencia técnica de los peritos, la decisión de si un acusado que alega una defensa afirmativa ha alcanzado la carga de producción de prueba, o la sentencia final (por mencionar apenas unos cuantos de las decenas de puntos a decidir más comunes), se requiere una regla de decisión. Esa regla normalmente es menos exigente que el estándar de prueba.

Si queremos tener una visión coherente del sistema de justicia penal, debemos ser capaces de mostrar que las reglas que rigen *cada una* de esas decisiones son adecuadas. Por la misma razón que los deontólogos no pueden suministrar un estándar de prueba para el veredicto al final del juicio, sus teorías no permiten ninguna otra regla de decisión a lo largo del mismo. De una manera que es absolutamente típica del *modus operandi* de los deontólogos, han construido sus teorías morales suponiendo que sabemos (por ejemplo) que X es una prueba admisible, o que ésta es la jurisdicción adecuada, o que la caución es justa, o que es desproporcionada. Pero generalmente no «conocemos» esas cosas, en el sentido técnico del término. Todas esas decisiones se adoptan bajo condiciones de incertidumbre y se sabe que la filosofía moral carece de una explicación acerca de cómo tiene lugar la adopción de decisiones bajo esas condiciones. Si los deontólogos no pueden decir cuándo está justificada la imposición de una caución, cuándo una prueba debe admitirse, cuándo debe concederse o rechazarse un recurso, y así sucesivamente ad nauseam, entonces no podemos tomarlos en serio su pretensión de ver el proceso judicial como un ejercicio de moral aplicada.

3. PARTE II: ¿PUEDE LA TEORÍA MORAL DEONTOLÓGICA EXPLICAR O JUSTIFICAR REGÍMENES PUNITIVOS ESPECÍFICOS?

Buena parte de la literatura sobre teoría moral en el ámbito jurídico ha planteado diversas cuestiones alrededor de la pena. Entre ellas: ¿cuándo y por qué puede el Estado aplicar penas; tienen o no los acusados un derecho a ser castigados?; ¿la pena pretende, legítimamente o no, la rehabilitación, la venganza, la disuasión o algo más?; ¿están justificadas o no la pena de muerte o las penas corporales? Éstas son preguntas interesantes e importantes. Pero la cuestión central que encaran quienes tienen que imponer la pena dentro un sistema jurídico, sean jueces o jurados, no es ninguna de las anteriores. Es, por el contrario, ¿cuál es la pena justa para alguien que ha sido condenado por un delito en concreto? Oficialmente, la teoría deontológica contemporánea tiene también una respuesta para esa pregunta, a saber: los acusados culpables deben recibir su merecido; o, más coloquialmente, los acusados culpables deben recibir una condena que se ajuste a su delito, es decir, que refleje con precisión el mal por el que acaban de ser condenados. Esta forma de venganza o (usando su descripción, más correcta políticamente) retribución, está muy de moda entre los teóricos de la moral, ahora que la rehabilitación como objetivo de la pena está ampliamente desacreditada.

Una cosa es decir que el acusado debe recibir su merecido y otra, muy diferente, determinar cuál es el merecido. Las teorías morales contemporáneas son categóricas con respecto a lo primero, pero callan extrañamente sobre lo segundo. En efecto, es difícil encontrar un solo teórico moral que diga específicamente cuál es la pena adecuada, lo merecido para un determinado tipo de delito. Este silencio dominante acerca de la condena apropiada para un delito determinado no es accidente ni descuido. Es, creo, un resultado directo del hecho de que la teoría deontológica tiene o muy pocas o ninguna herramienta para determinar si una pena «se ajusta» o no al delito.

Debo añadir por anticipado que no soy, de ninguna manera, el primero en señalar el hecho de que los retribucionistas, por su énfasis en dar a los criminales condenados su merecido, casi nunca se molestan en decirnos cuál es ese merecido⁶. Pero trataré de ofrecer una explicación al hecho de que ellos prácticamente nunca abordan lo que se esperaría fuese su preocupación más importante.

La razón de esta laguna no es difícil de imaginar. Los deontologistas tienen herramientas para determinar si un delito ha producido más daño que otro.

⁶ Véase especialmente a DOLINKO, quien observa que «desde hace mucho tiempo es una objeción de linaje al retribucionismo, que simplemente carece de una manera viable de determinar qué pena merece el delincuente». DOLINKO, 1992: 1636. Véase también BEDAU, 1978: 611-615.

Específicamente, miran los derechos de las víctimas violados por un delito y los comparan con los derechos violados por otro tipo de delito. Si algunos de esos derechos son más fundamentales, más básicos que otros, es posible decir que (por ejemplo) el asesinato es un daño más gravoso que la violación o que el robo a mano armada es peor que el allanamiento de morada. Este procedimiento, al menos en principio, permite a los deontologistas construir un escalafón que ordena la severidad de los distintos delitos. Cuando esto se combina con la regla de que la pena debe ajustarse al delito, se llega inexorablemente a la idea de que el escalafón que ordena la severidad de la pena debe reflejar una gradación de las condenas asociadas a esos delitos.

El problema irresoluble para los retribucionistas es simple: de un escalafón ordinal que dice cuáles daños, entre un catálogo de éstos, son más graves y cuáles son menos graves, no se puede derivar ninguna inferencia acerca de qué penas son adecuadas a qué daños. Podemos estar seguros de que, por ejemplo, el asesinato demanda un castigo más severo que la estafa o el robo, pero nada en esa gradación nos dice qué pena específica es el merecido para un delito independientemente de todos los demás (las mediciones ordinales describen el orden de un conjunto de objetos en relación con su ejemplificación de una propiedad o interés, pero no dicen nada sobre el tamaño o el grado correspondiente de diferencia entre los elementos clasificados). Las cosas serían diferentes si los retribucionistas nos dieran el criterio para hacer esa gradación. Eso nos diría no sólo qué daños son peores que otros, sino que además nos diría específicamente cuanto peor es un delito en comparación con otro. De esa forma, dado que sabríamos la pena apropiada para al menos un delito, podríamos calcular las penas adecuadas para los demás. Sin embargo, aun con ese criterio de gradación, los retribucionistas todavía tendrían que determinar cuál es la pena justa para al menos un delito, tarea que no está claro que sepan cómo llevar a cabo. Tal como están las cosas, sin tener un solo delito con su pena definida y sólo con un listado ordinal para trabajar a partir de él, es *lógicamente* imposible determinar cuáles penas serían las merecidas para qué delitos. En efecto, dadas las irrisorias herramientas que suministran los retribucionistas, puede demostrarse que es posible construir una cantidad infinita de escalafones de penas mutuamente incompatibles y sentenciar que son compatibles con cualquier gradación dada. Esta característica del discurso deontológico explica, en mi opinión, por qué se oye hablar mucho de acusados que reciben su merecido, pero casi nada acerca de cuál es ese merecido.

La crítica de que los deontologistas no han ofrecido un análisis de las penas adecuadas para delitos en concreto, por carecer de recursos conceptuales para hacerlo, fue explícitamente formulada por el deontologista VON HIRSCH en un breve artículo publicado en el *Journal of Philosophy*. Reconociendo que los deontologistas no han calculado una escala de ratios para medir la lesividad asociada a los diferentes delitos, VON HIRSCH presenta una explicación

diferente a la mía para esa falta. El hecho de que carezcamos de esa escala, insiste, simplemente «se debe a que nadie ha intentado seriamente construir dicha teoría»⁷.

El profesor VON HIRSCH, tomando él mismo las riendas, ha explorado —en una variedad de libros y artículos en las dos décadas posteriores a la citada publicación— un conjunto de medidas retribucionistas para deducir penas concretas a partir de escalafones ordinarios de la gravedad de diversos delitos.⁸ Cada una de esas medidas resultó finalmente ser una farsa. Por ejemplo, una de ellas implicaba tomar el escalafón de penas existente en un sistema jurídico y reajustarlo cuando se encontrase un delito cuya pena asociada fuese más (o menos) severa de lo que la gradación retribucionista permitiría. Sin embargo, ciertamente esos reajustes no le dan mucha holgura al proyecto retribucionista, en la medida en que la materia prima de HIRSCH (es decir, las condenas existentes) podría decirse que no tiene fundamento en instintos estrictamente retribucionistas. Después de todo, nadie niega que el principio según el cual la pena debe ajustarse al delito, puede servir como base para descartar ciertas penas actuales que resultan muy severas (o muy benévolas) en comparación con las penas previstas vigentes para otros delitos. Lo que los críticos del retribucionismo se preguntan es si éste tiene las herramientas para determinar de manera no comparativa cuál sería la pena justa para un delito determinado.

En otras correcciones propuestas por VON HIRSCH a la teoría, se le da amplia cabida al interés de permitir que la disuasión y la inhabilitación jueguen un papel en la formación de las condenas. Aunque la inclusión de factores tan progresistas es (en mi opinión) completamente razonable, su presencia socava el propósito del ejercicio de VON HIRSCH de mostrar que los factores puramente retribucionistas (y por tanto retrógrados) eran suficientes para fundamentar una teoría de la pena. VON HIRSCH efectivamente reconoce el fracaso del proyecto de deducir penas específicas de un catálogo basado en la retribución de los daños, cuando dice: «Un modelo de merecimiento es un esquema de condenas que observa el principio de proporcionalidad: las penas están escalonadas de acuerdo con la gravedad de los delitos [...] Una variedad de escalafones podría satisfacer las exigencias de este modelo»⁹. En pocas palabras, no hay una gradación o escalafón de penas que se siga del principio de proporcionalidad. Esto significa, aunque VON HIRSCH no lo concede tan explícitamente, que el proyecto retribucionista de deducir condenas a partir del principio de merecimiento es un fracaso.

Nos encontramos, pues, más de treinta años después de la insistencia de VON HIRSCH en que podría construirse, con la debida diligencia, un escalafón

⁷ VON HIRSCH, 1978: 623-624. Cursos originales.

⁸ Véase, especialmente, VON HIRSCH, 1985 y 1993.

⁹ VON HIRSCH, 1992: 79.

estrictamente retribucionista de penas y aún no lo tenemos pese a la abundancia de retribucionistas que han estado investigando activamente en las últimas tres décadas. ¿Cuánto más habrá que esperar para poder concluir que la inexistencia de dicho escalafón es un indicio poderoso de que los deontólogos no tienen las herramientas necesarias para construirlo?

Si mis críticas al retribucionismo parecen un poco abstractas, la siguiente tabla puede ayudar a concretarlas:

Tabla 1

Orden de lesividad	Delito	Escalafón 1	Escalafón 2	Escalafón 3	Escalafón 4	Escalafón 5
1	Estafa	1 año	1 día	10 días	20 años	Destierro
2	Robo	5 años	2 días	20 días	21 años	Amputación de manos
3	Secuestro	10 años	3 días	30 días	22 años	Esclavitud
4	Violación	15 años	4 días	60 días	23 años	Castración
5	Asesinato	25 años	5 días	25 años	25 años	Ejecución

Cada uno de estos escalafones ilustrativos cumple con el «orden de lesividad» supuestamente generado por los retribucionistas. ¿Pero cuál es el merecimiento para, digamos, la violación: 15 años, 4 días, 60 días, 23 años, o la castración? Si el lector cree que estoy simplemente jugando con los números, observe algunos casos de la vida real. Si miramos escalafones de condenas en Francia y en Estados Unidos por delitos similares, descubrimos que las condenas en Estados Unidos son, en general, entre 500 por 100 y 1.000 por 100 más severas que las condenas francesas¹⁰. Por ejemplo, mientras las directrices en Estados Unidos ordenan como máximo la cadena perpetua para homicidio en primer grado, las inglesas dan un máximo de treinta y cinco años y las francesas especifican ocho años. Mientras la condena máxima en Estados Unidos por estafa es de doce años, la inglesa es de siete años. Es interesante que los escalafones de penas de los tres países cumplan de manera general con el tipo de listado ordinal que un retribucionista aprobaría.

Sin embargo, estamos frente a una cuestión inevitable: ¿Cómo puede un retribucionista decidir cuál de los tres países, si es que alguno lo hace, da a

¹⁰ WITKIN, 2003. Véase también КЛЕМЕНД, «The Concept of Evil in American and German Criminal Punishment»: «Las condenas estadounidenses a penas de prisión son en promedio cinco a diez veces más prolongadas que en Francia, y también mucho más prolongadas que en Alemania, país del que se puede decir que es el más importante sistema jurídico continental y el punto de comparación para este ensayo».

los asesinos su merecido? ¿Los estadounidenses, los ingleses o los franceses? (Anoto entre paréntesis que esta confusión es justamente la contrapartida a la pregunta de la Parte I: «¿Quién tiene la razón en el valor de la ratio de Blackstone: Maimónides, Voltaire, Franklin o el propio Blackstone?»). Y si, como creo, él no puede responder esta pregunta, entonces su obstinada insistencia en que un delincuente condenado debe recibir su exacto merecido (y no más, ni menos) revela ser en sí misma un eslogan vacío y una farsa. En efecto, el retribucionista está en una posición en la que sólo puede decir: «El Estado tiene la obligación de darle su merecido al delincuente condenado; y, por cierto, no tenemos ni idea de cuál es ese merecido».

Las tesis retribucionistas, según las cuales la pena debería ajustarse al delito, se vuelven aún más caóticas cuando uno descubre que prácticamente todos los escalafones de penas en realidad no especifican una condena máxima sino un rango de penas. Estos son los rangos de penas vigentes actualmente en Estados Unidos para media docena de delitos:

Tabla 2

Delito	Rango de penas
Asesinato	De 10 meses a cadena perpetua
Intento de asesinato	De 6 a 14 años
Violación agravada	De 3 a 8 años
Violación	De 10 meses a 15 años
Posesión de drogas	De libertad condicional a cadena perpetua
Estafa	De libertad condicional a 12 años

En pocas palabras, de acuerdo con las autoridades estadounidenses, una condena de siete años podría ser la apropiada para un asesino, para quien ha intentado un asesinato, para un violador, para un expendedor de drogas y para alguien que ha cometido una estafa o una violación agravada. ¿Son comparables estas figuras con la doctrina de que la pena debe ajustarse al delito? ¿Hay formas de asesinato que sean tan benignas comparativamente como para que merezcan la misma o menor pena que una estafa? Si, como sospecho, rangos que se *superponen* como éstos no pueden ser acomodados dentro de los escalafones ordenados propuestos por los retribucionistas¹¹, entonces queda aun más claro que expresiones como «su merecido» y «la pena debe ajustarse al delito», son poco más que clichés vacíos, que ni siquiera están cerca de arrojar

¹¹ Claramente, el retribucionismo puede acomodar rangos de merecimiento justos para los delitos, de tal manera que no se superpongan con los rangos de delitos claramente desiguales.

luz sobre qué tipo de políticas respecto de las condenas penales son razonables y cuáles no lo son¹².

Un conocido retribucionista, Paul ROBINSON, ha sugerido que el problema de no tener un escalafón de condenas asociado a una perspectiva deontológica puede obviarse mediante la cuidadosa determinación de qué condenas cree el público general que son apropiadas para cada delito¹³. En sus tres estudios, Robinson y sus colegas describieron un conjunto de situaciones para sus entrevistados mediante las cuales sería posible ver si las preocupaciones por la pena merecida eran más o menos importantes que la disuasión, en la formación de las creencias de la gente acerca del castigo adecuado. Aunque los dos tipos de preocupaciones evidentemente se entremezclan en la opinión popular, ROBINSON afirma tener evidencia de que los juicios retrospectivos de la severidad del delito jugaron un papel mayor que los aspectos prospectivos de la disuasión delictiva general en la adopción de una decisión sobre la adecuación de una condena. Desafortunadamente, lo que estos estudios de investigación no hicieron fue prestar atención alguna al papel ni de los aspectos inhabilitantes ni rehabilitantes que puedan haber conducido las preferencias de la gente sobre las penas. Bajo esas circunstancias, no podemos asumir plausiblemente que los juicios de los ciudadanos coinciden estrictamente con lo que podría contener una teoría de los merecimientos de pena.

Por esas razones, hay que decir que los deontologistas carecen de los ingredientes necesarios para tener una teoría de la pena. Esta es una carencia grave en lo que pretende ser una teoría de la pena. No quiero sugerir que los deontologistas queden reducidos al silencio total, dado que pueden: a) criticar escalafones de condenas que no logren satisfacer su jerarquía de gravedad de delitos; así como, b) presentar objeciones a la imposición de penas basada en consideraciones sobre la peligrosidad o el comportamiento futuro del acusado. Pero ser capaz de entrometerse desde el margen no habilita para ser un jugador serio en el complejo mundo de la articulación de un sistema de penas satisfactorio. Hasta que, y a menos que, un deontologista pueda cuantificar la magnitud de las diferencias en los daños morales, no podrá decir, de ninguna pena ni de ningún delito, que es el merecido de alguien a quien se ha encontrado culpable de haberlo cometido.

¹² BEDAU plantea el punto sucintamente: «El nuevo retribucionismo fracasa precisamente en el mismo punto en el que lo hizo el viejo retribucionismo; todavía tiene que resolver el dilema de hacer que la pena se ajuste al delito», BEDAU, 1978: 615.

¹³ Véase ROBINSON *et al.*, 2002: 284. En un trabajo anterior, ROBINSON es muy explícito al decir que las penas que surgen de su versión del retribucionismo no son penas derivadas de una teoría moral sino del sentimiento popular: «Pedimos una distribución del merecimiento de responsabilidad, aunque es una forma particular e inusual de responsabilidad basada en merecimiento: una que se basa en los principios de justicia compartidos por la comunidad, más que en aquellos desarrollados por los filósofos» (ROBINSON y DARLEY, 1997: 456).

4. CONCLUSIÓN

Sin lugar a dudas, el lector ya habrá notado un enfoque común compartido por las dos secciones de este trabajo, una relacionada con el estándar de prueba y otra con las políticas sobre las condenas. En ambos casos, mi crítica a los deontologistas conlleva una observación sobre la incompletitud de su proyecto para suministrar fundamentos conceptuales al derecho penal. En ambos casos, los deontologistas están atados de manos porque carecen de toda vía plausible para graduar los daños relevantes. En el caso del estándar de prueba, sólo tienen meras intuiciones para decir cuánto más grave es una falsa condena que una falsa absolución. Sin una medida más específica y menos arbitraria de esa diferencia, no pueden ni deducir un estándar de prueba ni derivar ninguna otra regla de decisión para las numerosas decisiones que los jueces ordinariamente están llamados a adoptar.

Exactamente de la misma manera, su incapacidad para medir las diferencias entre los daños implicados en diferentes delitos hace imposible que identifiquen la pena merecida que, en su opinión, debería asociarse a todos y cada uno de los delitos. Sin saber cuál es el merecido para cada delito, su resis retribucionista central de que el merecido debe ser el motor que conduzca al conjunto de las penas, se hace poco más que un ejercicio de estlogan vacío. La línea de fondo, en mi opinión es simplemente ésta: hasta que y a menos que los deontologistas puedan expresar un sistema de medición para evaluar la magnitud de los daños morales, serán incapaces de decir cualquier cosa con contenido informativo acerca del estándar de prueba, acerca de las reglas de decisión que conducen los juicios penales y acerca de las penas que deben imponerse para delitos que tienen diferente grado de lesividad. Hace casi un siglo, A. C. EWING, al reconocer que el retribucionismo enfrenta serios problemas conceptuales, insistió en que no debía descartarse tan rápido¹⁴. Esto pudo haber sonado como un consejo en 1929, cuando el retribucionismo estaba surgiendo tras un siglo de desatención. Pero ahora no tiene sentido, dado que los retribucionistas contemporáneos, a pesar del impresionante crecimiento en su número, no parecen estar más cerca que EWING de resolver estos problemas.

Para ver lo empobrecida que está la teoría del merecimiento, imagínese una analogía cercana. Supóngase que algún epistemólogo anuncia que tiene una nueva teoría sobre la creencia racional. La teoría, en resumidas cuentas, es: nuestra confianza en una hipótesis debe ser proporcional al soporte empírico que el conjunto de pruebas disponible le da a esa hipótesis. Probablemente todos estaremos de acuerdo con que esto suena a un consejo sensato. Entonces

¹⁴ «Las ideas de merecimiento de "justicias" como buenas en sí mismas, y aun las de proporción entre culpabilidad y pena, están muy profundamente arraigadas en nuestro pensamiento ético como para descartarlas a la ligera, no importa lo mucho que puedan ser racionalizadas» (EWING, 1929: 45).

le preguntamos a nuestro filósofo amigo: «¿Cómo mides el grado de soporte empírico que el conjunto de pruebas le da a una hipótesis?». El responde, avergonzado, diciendo: «Bueno, no hemos llegado hasta allí todavía». La teoría del merecimiento viaja en un barco similar. Comienza con una premisa que suena plausible, según la cual los acusados culpables deben recibir la pena que merecen, pero después de ese punto no llega prácticamente a ningún lado.

Si hay o no teorías morales no-deontológicas (como el consecuencialismo) que tengan las herramientas para obtener esos fines, es una cuestión que dejo para otro momento.

Pero éstas no son las únicas dos áreas en las que la teoría moral no ilumina al derecho penal. Mencionaré una tercera, de pasada. Los deontologistas generalmente suscriben la idea de que causar un daño es más grave que permitir un daño. Esta cuestión es altamente pertinente para el derecho penal, en la medida en que tiene múltiples ramificaciones para decisiones sobre la prevención del delito y el uso de la peligrosidad como criterio de criminalidad. La anomalía, nuevamente, es la carencia de un sistema de medición deontológico para descubrir cuánto peor es causar un daño que simplemente permitirlo. El problema clásico que se le encadena es un ejemplo paradigmático de la incompletitud de la deontología como herramienta para orientar la decisión entre comprometer recursos para el castigo de la delincuencia y su prevención.

Podría continuar con esta discusión durante horas, pero no lo haré. En lugar de ello, terminaré citando una breve observación hecha por DWORKIN en su clásica obra *Taking Rights Seriously* (1977): los juristas, dijo, «temen a la contaminación de la filosofía moral» (p. 149). Lo que no añadió fue «¡Y deberían!».

BIBLIOGRAFÍA

- BEDAU, H., 1978: «Retributivism and the Theory of Punishment», *Journal of Philosophy*, 75.
- DOLINKO, D., 1992: «Three Mistakes of Retributivism», *UCLA Law Review*, 39.
- DUFF, R. A.; FARMER, L.; MARSHALL, S. E., y TADROS, V., 2007: *The Trial on Trial* (3) *Towards a Normative Theory of the Criminal Trial*, Oxford: Hart Publishing.
- DWORKIN, R., 1977: *Taking Rights Seriously*, Cambridge: Harvard University Press.
- EWING, A. C., 1929: *The Morality of Punishment*, London: Kegan Paul.
- KAPLAN, J., 1967: «Decision Theory and the Factfinding Process», *Stanford Law Review*, 20.
- KIRAL, R., 2003: «Protecting the Guilty», *Buffalo Criminal Law Review*, 6.
- KLEINFELD, J., «The Concept of Evil in American and German Criminal Punishment», en SSRN (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1667093).
- LILQUIST, E., 2005: «Absolute Certainty and the Death Penalty», *American Criminal Law Review*, 42 (1).

- ROBINSON, P. *et al.*, 2002: «Why Do We Punish? Deterrence and Just Deserts as Motivations for Punishment», *Journal of Personality and Social Psychology*, 83.
- ROBINSON, P., y DARLEY, J., 1997: «The Utility of Deserts», *Northwestern University Law Review*, 91.
- STEN, A., 2005: *Foundations of Evidence Law*, Oxford.
- TRIBE, L., 1971: «A Further Critique of Mathematical Proof», *Harvard Law Review*, 84.
- VON HIRSCH, A., 1978: «Proportionality and Desert: A reply to Bedau», *Journal of Philosophy*, 75 (11).
- 1985: *Fast or Future Crimes*, Rutgers: New Jersey.
- 1992: «Proportionality in the Philosophy of Punishment», *Crime and Justice*, 16.
- 1993: *Censure and Sanctions*, Oxford: Oxford University Press.
- WITTMAN, J. Q., 2003: *Harsh Justice: Criminal Punishment and the Widening Divide Between America and Europe*, Oxford: Oxford University Press.

VI

¿FUE EL BEBÉ SACUDIDO? PRUEBA, PERICIA Y EPISTEMOLOGÍA JURÍDICA EN EL PROCESO PENAL INGLÉS*

Paul ROBERTS **
University of Nottingham

1. INTRODUCCIÓN: LA PREOCUPACIÓN EPISTEMICA EN LA JURISDICCION PENAL

«Hay pocos tipos de casos en los que surja mayor preocupación y controversia —dijo MOSES, L. J., al pronunciar la sentencia de la Corte de Apelaciones en el caso *Henderson*— que aquellos en los que se sostiene que un bebé ha muerto como consecuencia de haber sido sacudido»¹. En las últimas dos décadas, se han fallado más de cien casos sobre el «síndrome del bebé sacudido» en los juzgados penales de Inglaterra y Gales. La exitosa apelación del caso *Cannings*² (en el que se alegó asfixia y no sacudidas) precipitó una investigación oficial urgente sobre otras 297 condenas, en casos en los que habían muerto repentinamente bebés e infantes en circunstancias sospechosas³.

* Traducción de Maximiliano ARAMBURU.

** Profesor de *Criminal Jurisprudence* en la University of Nottingham; profesor-visitante de la University of New South Wales.

¹ *R. v. Henderson*; *R. v. Butler*; *R. v. Oyediran* [2010], 2 Cr. App R. 24; [2010] EWCA Crim. [1].

² [2004] 2 Cr. App R. 7.

³ Véase ATTORNEY GENERAL'S OFFICE, 2004.

En casi la tercera parte de éstos se encontró que los padres o las niñeras habían infligido a los bebés y a los infantes lesiones por sacudidas, muchas veces fatales. *Cannings* se aseguró un lugar en la historia británica de los errores judiciales notorios, al lado de la casi contemporánea exoneración de Sally Clark⁴, otro caso que hablaba de una inexplicable «muerte en la cuna», en lugar de las presuntas sacudidas.

Hay varios fundamentos en virtud de los cuales podría considerarse que estos casos son la excepción, especialmente si el punto de partida de la comparación es la dosis diaria de hurtos, robos, vandalismo, agresiones menores y tráfico de drogas a menor escala, que consumen buena parte de los recursos y la atención de nuestros juzgados penales. El homicidio es estadísticamente raro y extraordinariamente grave; y es todavía más doloroso cuando la corta vida de la víctima apenas comenzaba. Los casos de bebés sacudidos son también excepcionales desde el punto de vista forense, dado que los juzgadores de los hechos a menudo tienen que elegir directa y dramáticamente entre asesinato (u homicidio voluntario) y muerte natural, en este último caso dando como resultado que no hay delito alguno de por medio. Y como si esta elección no fuese lo suficientemente difícil, el acusado terminará siendo un asesino o un ya turbado padre o niñera, para quien un veredicto erróneo de culpabilidad añadiría injusticia a una tragedia de por sí desgarradora. Un tercer aspecto en que los casos de bebés sacudidos son inusuales (aunque no únicos) estriba en la escasez de los testimonios fácticos de los que dispone el juzgador de los hechos. La víctima o bien ha fallecido o bien es un menor en estado prelingüístico, en cualquier caso incapaz de explicar cómo se produjeron las lesiones. Normalmente se alega que las lesiones se han producido mientras el menor se encontraba al cuidado únicamente del acusado, usualmente en casa o en otro establecimiento doméstico, eliminando así la posibilidad de testigos independientes. El único adulto testigo del supuesto delito es el acusado, es decir, la persona con mayores motivos para mentir si es que en realidad es culpable. Tomados en conjunto, estos tres factores son más que suficientes para preocupar a la jurisdicción penal. Como observó la Corte de Apelaciones en *Henderson*, «no sorprende la controversia que generan estos casos»⁵.

Dando por sentado que nos interesan las verdades generales —cuando no las universales— más que las biografías individuales, ¿qué podemos esperar aprender del análisis de la dinámica forense de estos procesos y sus apelaciones, que admitimos como excepcionales? *Henderson* es en la actualidad un importante precedente jurídico en la práctica penal inglesa, pero ¿por qué deberían prestarle alguna atención los juristas extranjeros o los filósofos? El estudio de casos puede ser ilustrador, justamente porque suministra un destacado punto de comparación con lo prosaico y rutinario. Igualmente, el estudio

de casos puede ser valioso porque ejemplifica la cotidianidad. Este trabajo demostrará por qué la sentencia de la Corte de Apelaciones en *Henderson* compensa su cuidadoso estudio desde esos dos puntos de vista.

A pesar de su evidente excepcionalidad, los casos de bebés sacudidos también ejemplifican muchas características ordinarias de los juicios penales controvertidos. En ellos hay, por definición, una disputa acerca de uno o más hechos clave que el tribunal debe resolver de alguna manera. Para hacerlo, los tribunales despliegan una variedad estándar de recursos institucionales, normativos y epistemológicos. Para los casos de delitos graves en Inglaterra y Gales, esto significa que se trata de procesos conducidos por las partes y la determinación de los hechos corresponde a jurados legos, todo dentro de un marco de reglas y principios procesales detallados (incluyendo las reglas que rigen la admisibilidad y los usos lícitos de las pruebas así como las teorías jurídicas para distribuir los déficits y los riesgos epistémicos, tales como las presunciones probatorias, las cargas y los estándares de prueba)⁶. Con pocas fuentes de información adicionales sobre los hechos en controversia, el testimonio de los peritos es inevitablemente crucial en los procesos de bebés sacudidos y éste es, también, un rasgo predominante del litigio penal moderno. Podría simplemente mencionar las acusaciones que se basan, por ejemplo, en perfiles de ADN, huellas dactilares, pruebas de balística, cintas de vídeo de circuitos cerrados de televisión o en cualquier otra «pista» proveniente de las ciencias forenses. Como dijo la Corte de Apelaciones acerca de las condenas por bebés sacudidos que fueron revisadas en *Henderson*, «expertos, autoridades de la acusación y jurados deben reconstruir lo que ha sucedido, de la mejor manera posible»⁷.

En concreto, la siguiente disección de la sentencia de *Henderson* pretende elucidar las características más destacadas de la «epistemología jurídica»⁸ propia de la determinación de los hechos al servicio de la justicia penal, al menos en el modo en que estas prácticas epistemológicas se interpretan, se institucionalizan y se recrean en el proceso penal inglés. Además de presentar un caso de estudio finamente construido en las prácticas epistémicas de la jurisdicción, este artículo pretende contribuir a los estudios de derecho comparado, enfatizando la importancia de las distintas tradiciones jurídicas⁹ y sus explicaciones justificativas, para darle sentido al proceso penal en términos de sus instituciones jurídicas locales, sus marcos normativos y sus culturas jurídicas profesionales¹⁰.

⁶ Para una elucidación sistemática, véase ROBERTS y ZUCKERMAN, 2010: especialmente capítulos 2 y 6.

⁷ [2010] 2 Cr. App R. 24 [1].

⁸ Con una venia respetuosa a LAUDAN, 2006.

⁹ Cf. GLENN, 2010.

¹⁰ Temas en desarrollo, previamente explorados en ROBERTS, 2006; 2008a; 2008b; 2010; 2011; ORCOT y NELKEN, 2007; NELKEN, 2000.

⁴ R. v. Clark [2003] EWCA Crim. 1020; KIRBY, 2003; ROZENBERG, 2003.

⁵ R. v. Henderson [2010] 2 Cr. App R. 24 [1].

2. HISTORIA DE LOS CASOS

Nuestra primera tarea, teniendo presente el argumento siempre vigente de William TWINING con respecto a «tomar los hechos en serio» en la enseñanza y la investigación del derecho probatorio¹¹, es mirar más de cerca los hechos del caso. La sentencia de la Corte de Apelaciones en *Henderson*, en realidad, se ocupó de las apelaciones conjuntas de tres procesos separados, cada uno de los cuales en ciertos aspectos ejemplifica y en otros tantos contradice mi bosquejo introductorio del caso «típico» de bebé sacudido.

En el propio caso *Henderson*, la acusada era una muy reputada y experimentada niñera. Un niño de once meses de edad que estaba bajo su cuidado se desplomó y dos días después murió en el hospital. Por lo menos doce peritos declararon en el juicio de Henderson: once de ellos apoyaban la versión de la acusación, según la cual Henderson había sacudido al menor hasta provocarle la muerte o bien había golpeado su cabeza contra una superficie blanda mientras lo sacudía. Henderson resistió vigorosamente las acusaciones y negó haber lesionado en forma alguna al niño. Presentó como pruebas de carácter a testigos que declararon sobre su experiencia y fiabilidad como niñera y destacó el hecho de que había criado a dos hijos propios. Después de un largo juicio de seis semanas, Henderson fue condenada por homicidio involuntario en un veredicto del jurado con una mayoría de 10-2, fundado ampliamente en el testimonio médico. No sin algunas dudas por parte de la Corte de Apelaciones, la condena de Henderson fue ratificada.

La segunda de las tres apelaciones conjuntas, *R. v. Butler*, implicaba a un padre cuya hija de siete semanas, Ellie, fue ingresada de urgencia en el hospital. El bebé no mostraba señales externas de traumatismos, pero resultó que tenía graves lesiones internas en la cabeza. Las sospechas aumentaron porque Ellie había sufrido también quemaduras en su frente y en un brazo que, según se dijo, habían ocurrido cuando se cayó de una almohada y entró en contacto con un radiador. Los médicos diagnosticaron una clásica lesión de cabeza no accidental (NAHI: *non accidental head injury*), con un pronóstico terminal. La «característica inusual de este caso»¹² consistió en que Ellie, inesperadamente, se recuperó por completo. Su padre, sin embargo, fue procesado y condenado por lesiones corporales graves¹³ y crueldad infantil¹⁴, tras un juicio de cuatro semanas en el que la acusación llamó a quince médicos expertos y la defensa presentó como contraprueba a tres más. La Corte de Apelaciones admitió subsecuentemente la apelación de Butler y anuló su condena de die-

¹¹ TWINING, 1982 (reimpr. 2006); 2007.

¹² [2010] 2 Cr. App R. 24 [87].

¹³ *Offences Against the Person Act* de 1861, art. 20.

¹⁴ *Children and Young Persons Act* de 1933, art. 1.

ciacho meses de prisión. La Corte temió que el conocimiento del jurado sobre las lesiones previas de Ellie lo hubiese predispuerto contra el acusado, a pesar de que no había ningún indicio de que las quemaduras de radiador hubiesen sido infligidas deliberadamente. En retrospectiva, dijo la Corte, este incidente pudo atribuirse a la «falta de cuidado de padres primerizos»¹⁵, lo cual no es de interés para el derecho penal.

En *R. v. Oyediran*, la tercera de las apelaciones conjuntas, el acusado fue condenado por matar a su hijo de diez semanas de edad, Femi, quien había sufrido daños cerebrales mortales. La hipótesis de la acusación en el proceso, basada en el testimonio de cinco patólogos y pediatras, era que Femi había sufrido dos lesiones cerebrales distintas y acumuladas, la primera cerca de dos semanas antes de su muerte y la segunda dos o tres días antes de morir, ambas atribuibles a sacudidas o empujones, o a una combinación de las dos. Después de su muerte se encontró que Femi también tenía una inusual y no explicada fractura en el brazo, que aparentemente habría ocurrido entre dos y cuatro semanas antes de su muerte. También había pruebas considerables sobre ciertos antecedentes del temperamento volátil del acusado y de su hostilidad hacia médicos y trabajadores sociales. Sin embargo, el acusado rechazó firmemente cualquier insinuación de haber lesionado deliberadamente a su hijo. Habría hecho cualquier cosa, decía, para proteger al pequeño Femi. No había notado ninguno de los síntomas descritos por los expertos de la acusación y no podría explicar cómo ocurrieron las lesiones de Femi. La madre de Femi, Sophia Rudder, quien aparentemente estaba dominada por el acusado, tenía esclerosis múltiple avanzada. Se sugirió en el juicio y se sostuvo de nuevo en la apelación, que las lesiones de Femi podrían haber sido causadas por la incapacidad de su madre para cargarle adecuadamente. Rudder también fue acusada, junto con su esposo, por haber causado o haber permitido la muerte del niño¹⁶. Ella fue absuelta, pero sucumbió a su propia enfermedad devastadora y murió poco después. La Corte de Apelaciones posteriormente rechazó la apelación de Oyediran y ratificó su condena por asesinato, con la recomendación de que purgara como mínimo un período de trece años de prisión.

Debido a sus similitudes fácticas y sus aspectos jurídicos convergentes, las tres apelaciones separadas *R. v. Henderson*, *R. v. Butler* y *R. v. Oyediran*, fueron acumuladas en un único fallo. Como observó la Corte de Apelaciones, «sería desafortunado que diferentes composiciones [de la Corte] redactaran sentencias distintas en momentos diversos, a pesar de que las apelaciones fueron escuchadas tan cerca unas de las otras»¹⁷. Los tres niños lesionados fueron diagnosticados con encefalopatía y hemorragia subdural; y dos de ellos (salvo Femi Oyediran) tenían también hemorragias retinianas, un trío de sínto-

¹⁵ [2010] Cr. App R. 24 [88].

¹⁶ *Domestic Violence, Crime and Victims Act* de 2004, art. 5.

¹⁷ [2010] Cr. App R. 24 [8].

mas conocido en la literatura médica, y posteriormente en la práctica forense, como «la triada». En cierto momento, algunos expertos consideraron que la triada constituía un diagnóstico concluyente de una *NAHI*, mientras que otros pensaban que era una consecuencia de alguna patología subyacente que no despertaba sospecha, pues no estaba aún detectada («la hipótesis unificada»). Esta combinación de síntomas hoy es ampliamente considerada como fuertemente indicativa de *NAHI*, aunque no es concluyente. Como veremos, el fallo *Henderson* se construye a partir de pronunciamientos previos de la Corte de Apelaciones en *Harris*¹⁸, otra apelación conjunta en la que por primera vez los jueces ingleses revisaron sistemáticamente el valor probatorio de la triada.

3. LA JURISDICCIÓN DE APELACIONES EN INGLATERRA Y GALES: EL OBJETO Y LA RAZÓN DE SER DE LAS APELACIONES PENALES EN UN SISTEMA PROCESAL ADVERSARIAL

Toda determinación de los hechos realizada por seres humanos es «contextual», en el elemental sentido de que tiene lugar en tiempos determinados, en lugares y espacios concretos y que se lleva a cabo por seres humanos con habilidades cognitivas, a menudo admirables pero ciertamente limitadas, y con particulares motivaciones y objetivos en mente. La determinación de los hechos en el terreno jurídico es especial no por ser contextual, sino porque tiene lugar en un tipo especial de contextos institucionales y porque está sometida a expectativas normativas muy particulares y altamente desarrolladas.¹⁹

Todos los sistemas jurídicos modernos se ven a sí mismos en el esfuerzo por la realización de la justicia y no digamos en el de llevar a cabo una «búsqueda exhaustiva del entendimiento cósmico».²⁰ Pero las nociones acerca de lo que la justicia exige, difieren muy considerablemente de un lugar a otro, especialmente cuando se trata del diseño de instituciones y procedimientos adecuados para el proceso penal. La primera tarea que se acomete para propiciar una mejor apreciación del proceso de toma de decisiones jurídicas en casos que involucran pruebas científicas debatidas o, según el caso, cualquier otro tipo de pruebas, consiste en prestar mucha atención al tipo de procedimiento jurídico en cuestión, su finalidad, su alcance, los poderes jurisdiccionales y sus limitaciones así como a sus costumbres culturales características.

¹⁸ *R. v. Harris*; *R. v. Roake*; *R. v. Cherry*; *R. v. Faulder* [2006] 1 Cr. App. R. 5 [2005] EWCA Crim. 1980.

¹⁹ ROBERTS y ZICKERMAN, 2010, cap. 1. Véase además HO, 2008; STEIN, 2005; 1999.

²⁰ *Dabbert v. Merrill Dow Pharmaceutical*, 509 US 579, 597; 113 S. Ct. 2786, 2799 (1993) (US Supreme Court).

A diferencia de las apelaciones penales en buena parte de Europa continental, donde un segundo juicio *ex novo* no es infrecuente²¹, las apelaciones penales de una condena en Inglaterra y Gales sólo proceden sobre asuntos muy restringidos. La competencia de la Corte de Apelaciones (Sala de lo Penal) se reduce básicamente a la revisión en segunda instancia de la *legalidad* de la condena y no se ocupa del tipo o de la calidad de las pruebas con las cuales el acusado fue condenado²². Además, la Corte ordinariamente respeta el ejercicio de la discrecionalidad del juez de primera instancia sobre las cuestiones de manejo del juicio y sobre las cuestiones de procedimiento, de manera que aunque se reconozca que ha habido irregularidades procesales menores, esto no necesariamente significa que la condena del acusado será declarada «insegura»²³. * y anulada. En general, «ser (fácticamente) inocente» no es motivo de apelación en Inglaterra y Gales²⁴, a menos que el acusado esté en capacidad de respaldar esa simple afirmación con «nuevas pruebas» exculpatorias, que lleven a considerar que desde el juicio original²⁵ o, excepcionalmente, en la propia Corte de Apelaciones, se alberga una «duda acechante» sobre la seguridad de la condena, aunque no se puedan atribuir esas reservas a un defecto procesal en particular o a un error en el proceso judicial²⁶.

Las razones para este enfoque deliberadamente reactivo acerca de las apelaciones penales, se basan parcialmente en consideraciones de eficiencia e irre-

²¹ Véase, por ejemplo, VOGLER y HUBER, 2008. Nótese la posibilidad de juicios de «segunda instancia» en Francia, Alemania, Holanda y España.

²² Véase, por ejemplo, *R. v. Szyn* [2010] EWCA Crim. 923.

²³ Ese es el criterio para admitir apelaciones de conformidad con la *Criminal Appeal Act* de 1968, tras la reforma de la *Criminal Appeal Act* de 1995.

²⁴ *N. del T.*: Desde la reforma de 1995 citada por el autor en la nota anterior, la *safty* o seguridad de la condena es el único estándar o criterio con el que debe examinarse la sentencia para saber si es, prima facie, susceptible de apelación. Antes de dicha reforma, la Corte aplicaba el criterio «*unsafe and unsatisfactory*». QUINN (2007: 762, 763) explica que la *unsafe* no tiene una definición precisa y tiene que ver tanto con las dudas fácticas que puedan existir (ninguna condena puede ser *safty* si hay dudas), como con los vicios jurídicos del proceso (aunque no haya dudas, una condena puede ser *unsafe* si existen vicios jurídicos). En adelante, para hablar de la posibilidad o imposibilidad de revisar la condena en apelación—esto es, para hablar de la *safty* de la condena—utilizaré los términos «segura» e «insegura».

²⁵ Otras implicaciones fueron aclaradas por QUINN, 2007: 759.

²⁶ El art. 23 del *Criminal Appeal Act* de 1968 dispone, en lo que nos interesa, que «la Corte de Apelaciones puede, si lo considera necesario u oportuno en beneficio de la justicia [...] *c)* recibir cualquier prueba que no haya sido aportada en el procedimiento al que se refiere la apelación». El numeral 23 (2) más adelante, dice que: «La Corte puede, al considerar si recibe o no alguna prueba, mirar en particular *a)* si la prueba le parece a la Corte creíble; *b)* si le parece a la Corte que la prueba puede suministrar algún fundamento para permitir la apelación; *c)* si la prueba habría sido admisible en el proceso al que se refiere la apelación, sobre un punto que sea objeto de la apelación, y *d)* si hay una explicación razonable para el hecho de no haber aportado esa prueba durante el proceso».

²⁷ «[E]n casos como este, el juzgador debe al final formularse a sí mismo una pregunta subjetiva: si está satisfecho de dejar el asunto tal como está, o si hay alguna duda acechante en su mente que le haga pensar que se ha cometido una injusticia. Esta es una reacción que puede no estar basada estrictamente en las pruebas; es una reacción que puede producirse por el sentimiento general del caso, tal como lo concibe el juzgador». *R. v. Cooper* [1969] 1 QB 267, 271 CA. Véase también *Stofford v. DPP* [1974] AC 578 HL; *R. v. B. (Brian S.)* [2003] 2 Cr. App. R. 197; [2003] EWCA Crim. 319, CA.

vocabilidad comunes a todos los sistemas jurídicos modernos²⁷ pero también, de manera más concreta, en el papel fundamental que juegan los jurados legos en la justicia penal en el mundo del *common law*²⁸. Pese a su disminución cuantitativa con respecto a la totalidad de los juicios penales y a su casi completa erradicación de la justicia civil²⁹, los jurados legos siguen siendo una parte vital de la administración de justicia penal en Inglaterra y Gales en los delitos graves juzgados por la *Crown Court*. El *Lord Chief Justice* (Magistrado Presidente de la Corte Suprema) describió recientemente el juicio por jurados como un «principio sagrado» del proceso penal inglés³⁰. Más allá de sus cometidos prácticos en la determinación de los hechos (desarrollados ampliamente más adelante), los jurados legos aseguran la legitimidad de los veredictos penales garantizando que los procedimientos judiciales y la aplicación de la ley penal en cada caso obedezcan siempre a los valores y las expectativas de los hombres y las mujeres medios y respondan a su sentimiento de justicia. Lo que sea que se haya significado «juzgamiento por pares» para la nobleza que dedujo esa concesión de la Carta Magna del rey Juan en el siglo XIII, éste es el sentimiento demoliberal que ha llegado a incorporarse al pueblo británico actual. Una apelación penal ante un juez único no logra replicar la autoridad práctica o simbólica del juicio ante jurados. Los procedimientos de apelación, en consecuencia, se limitan a la revisión de la legalidad y la firmeza del veredicto emitido por el jurado, sin buscar de ninguna manera retar, ni mucho menos usurpar, la indispensable función constitucional del jurado. Aun en los casos en los que prospera la apelación de una condena, es probable que el caso sea remitido nuevamente a la *Crown Court* para repetir el juicio ante un jurado, a menos que un nuevo juicio sea inútil de cara a los intereses de la justicia en el caso concreto³¹.

El fallo *Henderson* comienza reiterando estos límites fundacionales institucionales de las apelaciones penales. «[E]l juicio por jurados no significa juicio por jurados en la primera instancia y juicio por jueces ante la Corte de Apelaciones, en la segunda», recitaba Moses, L. J. Este precepto fundamental se «aplica tanto a los casos que dependen de pruebas periciales como a aquellos que no»³². La primacía de la determinación de los hechos por legos se

²⁷ NOBLES Y SCHIFF, 2002: 676.

²⁸ Véase ROBERTS Y ZUCKERMAN, 2010: 65-72; VOGLER, 2005: cap. 10; THORNTON, 2004: 683.

²⁹ LLOYD-BOSTOCK Y THOMAS, 2000.

³⁰ R. v. *Twomey* [2009] 2 Cr. App R. 25 [10] (Juez CJ).

³¹ *Criminal Appeal Act* 1968, art. 7.

³² [2010] 2 Cr. App R. 24 [4], citando a Lord Bingham en R. v. *Pendleton* [2002] 1 WLR 721 [2001] UKHL 66 [17]: «[E]l papel central del jurado en un juicio con acusación [...] es una característica importante y muy apreciada de nuestra Constitución. El juicio por jurados no significa juicio por jurados en la primera instancia y juicio por jueces de la Corte de Apelaciones en la segunda. A la Corte de Apelaciones le ha sido confiada una potestad de revisar los casos, que debe preservar la posibilidad de injusticias, pero es una potestad que debe ejercer con cautela, teniendo en cuenta que la Corte de Apelación está al margen de las deliberaciones del jurado y no debe entrometarse en un territorio que le pertenece justamente a aquél».

combina con la estructura adversarial del proceso penal inglés para producir una segunda y más técnica limitación al alcance de las apelaciones penales. Es un principio general del litigio en procesos adversariales que ninguna parte pueda aportar un elemento probatorio, o argumentar sobre él, si lo tenía a su disposición desde antes de comenzar el juicio oral pero no lo presentó por razones estratégicas³³. Esta regla procesal genérica se destaca particularmente en la prueba pericial, y especialmente en las *opiniones* * de los peritos.

Se esperaría normalmente que las dos partes den lo mejor de sí en el juicio y la acusación está obligada legalmente a hacerlo³⁴. El proceso adversarial se justifica en parte precisamente porque, de forma natural, genera incentivos positivos para que cada parte procesal «gane el juicio» mediante la producción de la mejor prueba posible para persuadir al juzgador de los hechos sobre la justicia de su pretensión³⁵. Por supuesto, es perfectamente posible que surjan nuevas pruebas relevantes después de que el juicio oral ha tenido lugar, debido a un descuido, por incompetencia o falta de corrección, o simplemente porque la vida a veces funciona así y de ello no puede culparse a nadie. El derecho inglés ha desarrollado varios mecanismos para revisar los veredictos cuando genuinamente aparecen nuevas pruebas después del juicio oral, incluyendo, por supuesto, las apelaciones de condenas penales. Pero la situación es sustancialmente diferente cuando la información en cuestión estuvo disponible a lo largo del proceso y la parte interesada simplemente eligió no usarla.

Desde un punto de vista puramente epistemológico podría decirse que si los procedimientos previos al inicio del juicio oral penal operasen en la forma en que deberían hacerlo³⁶, de manera que el acusado sea informado adecuadamente tanto de la naturaleza de las afirmaciones de la acusación como de las pruebas existentes para demostrarlas³⁷, no hay espacio legítimo para que la estrategia defensiva entre en estas consideraciones tácticas. Si es inocente, el acusado sólo necesita decir la verdad. Si es culpable, el procedimiento penal bien puede proporcionarle oportunidades para ser evasivo y ocultarla (en particular guardando silencio), pero difícilmente podrá quejarse si sus estrategias fracasan o salen mal y termina siendo condenado. *Ex hypothesi*, ése es justamente el resultado procesal que se merece.

³³ «[S]i en circunstancias muy excepcionales admirará esta Corte nuevas pruebas en apelación, cuando el material probatorio haya sido deliberadamente abstraído del juicio por razones tácticas»: R. v. *Solomon* [2007] EWCA Crim. 2635 [29].

³⁴ *N. del T.*: «Opinión» es una traducción literal de «*opinion*». Sin embargo, incluso en el *common law* se discute si lo que emiten los peritos es mera opinión o un tipo de declaración más elaborada y comprometida. En este texto se preferirán expresiones como «testimonio» y similares, salvo cuando la literalidad lo impida.

³⁵ R. v. *Phillipson* (1989) 91 Cr. App R. 226, CA.

³⁶ Cf. POSNER, 1999: 1477.

³⁷ No debe subestimarse la importancia práctica del condicional. Los abogados defensores usualmente se quejan de que el descubrimiento por parte de la acusación es tardío, incompleto o ambos.

³⁸ *Criminal Procedure and Investigations Act* de 1996, Parte 1 (después de la reforma); *Criminal Procedure Rules* de 2011, Partes 21 y 22.

Este es uno de los muchos aspectos importantes en los que una epistemología descontextualizada no logra tener en cuenta las realidades empíricas del proceso penal. Los abogados defensores están habitualmente obligados a hacer razonamientos estratégicos, aun en los casos simples y aun cuando el acusado sea (fácilmente) inocente.³⁸ Un tipo importante de decisión táctica puede implicar elegir entre múltiples líneas de defensa o un simple contrarrelato. La noción psicologista, según la cual los juzgadoras de los hechos legos procesan la información en términos de «historias» narrativas holísticas³⁹, tiene fuerte eco en la experiencia de muchos abogados litigantes.⁴⁰ Teniendo esto en mente, podría ser más efectiva una estrategia defensiva que consista en presentar una narración simple, coherente y ordenada sobre la inocencia, en lugar de disparar hacia todos los lados de la acusación. Un ataque multifacético corre el riesgo de ser visto por el jurado como un intento más bien desesperado de detenerse en nimiedades, ante la ausencia de una defensa fuerte y probatoriamente creíble, y siempre se corre el riesgo de que los jurados legos no logren captar los matices de la posición de la defensa o de relatar la importancia global de una serie de distintos desafíos a la prueba de la acusación, ante la carga general que pesa sobre la acusación de probar la culpabilidad más allá de toda duda razonable. Otra decisión táctica relacionada tiene que ver con la elección entre aportar o no elementos de prueba particulares, que bien podrían ser relevantes para el caso considerado globalmente, pero que necesitan ser valorados en términos de si realmente contribuyen al relato elegido por la defensa, o suministran suficiente «valor añadido» a las pruebas como para que se justifique tomar el riesgo de que su aportación confunda las cuestiones, lleve a los jurados a un callejón sin salida o de alguna otra manera reste fuerza a los argumentos principales de la defensa.

El derecho inglés aborda estas decisiones estratégicas desde un punto de vista diferente del que emplea en aquellos casos en que realmente hay pruebas nuevas descubiertas después del juicio oral. La lógica del procedimiento adversarial consiste en que el acusado se responsabiliza de las elecciones estratégicas que toma—o que, en la mayoría de los casos, hacen sus abogados— a

³⁸ Tal vez la decisión táctica que más consecuencias tiene es que el acusado declare o no como testigo en su propia defensa. Sería simplista decir que un acusado será absuelto si dice la verdad. Muchos acusados rinden testimonios deficientes en su propia defensa: pueden tener dificultad para expresarse, ser olvidadizos, estar drogados, ser agresivos o desconfiados. Algunos acusados pueden verse —y de hecho lo son— como individuos completamente desagraviados que inmediatamente lograrían la antipatía del jurado tan pronto abran su boca. Estas son consideraciones que pueden llevar a los abogados defensores a no llamar al acusado como testigo en su propia defensa, aunque saben muy bien que el jurado podría reprochárselo, de acuerdo con la expectativa de sentido común según la cual la inocencia habla por sí misma—sin tomar en consideración el hecho de que, como formalidad jurídica, al juez se le permite invitar al jurado a desoír toda influencia derivada del silencio del acusado (tal como lo pueden hacer los jueces de primera instancia en Inglaterra y Gales). Véase *R. v. Beccuani* [2006] 1 Cr. App. R. 2 [2005] UKHL 55; cit. *R. v. Friend (No. 2)* [2004] EWCA Crim. 2661.

³⁹ PENNINGTON Y HASTIE, 1993; PARDO, 2000: 399; TURNING, 2006: caps. 9-13.

⁴⁰ BLUME, JOHANSON Y PAAVOLA, 2007: 1069; ALLEN, 1994: 604.

lo largo del proceso. A la defensa se le impide desconocer posteriormente esas elecciones tácticas, simplemente por el hecho de que el resultado final del juicio no le haya favorecido. Así como no es posible plantarse y arrepentirse en el juego de cartas del veintinueve, no se puede desarrollar una defensa en el juicio y luego esperar que se permita desarrollar una diferente en la apelación sobre la base de que la primera no tuvo éxito y se quiere otra oportunidad. A menos que el acusado tenga una asesoría jurídica incompetente (un argumento al que los jueces ingleses no son muy receptivos)⁴¹, está obligado a aceptar las consecuencias de sus decisiones estratégicas. Aunque ciertamente tiene sentido preocuparse por el hecho de que este riguroso diseño institucional pueda crear injusticias en casos concretos, la lógica del sistema adversarial es férrea en el nivel institucional del diseño del sistema. Ningún sistema jurídico permitiría al acusado exigir que el juicio que resulta en condena sea repetido tantas veces como le tome agotar cada línea alternativa de defensa que pueda concebir o hasta tanto sea absuelto.

Actualmente, la oportunidad para lograr que un proceso se repita es mucho mayor cuando hay peritos de por medio. Si la acusación está equipada de testigos presenciales que observaron todo el evento o si la defensa tiene la suficiente suerte de tener un testigo irrefutable de la coartada para presentarlo como prueba en el juicio oral, suele ser imposible obtener repetición alguna si, por cualquier razón, no logran conseguirse estos testigos estrella o no se desempeñan como se esperaba antes del juicio. Las partes juegan sus mejores cartas en el día señalado y, pase lo que pase, el resultado es definitivo. Sin embargo, con la posible excepción de muy pocas especialidades forenses⁴², los peritos son un bien esencialmente fungible, en especial cuando ofrecen una opinión profesional en lugar de un testimonio relacionado con algún aspecto fáctico del proceso en juego (como sería, por ejemplo, el perito que siendo el médico de urgencias examinó primero a la víctima y registró sus heridas). En un sistema adversarial con peritos que son instruidos por las partes, la prueba científica será aportada al proceso penal sólo cuando respalde el caso de quien la está proponiendo⁴³. Este tipo de disposiciones procesales causan una tentación obvia para que un abogado intente repetir la prueba pericial en apelación. Si el desempeño del perito en el juicio fue decepcionante, ¿por qué no asegurarse uno mejor (es decir, uno más favorable a la defensa) para que declare en la apelación? Igualmente, por las razones que hemos visto, la Corte de Apelaciones debe ser astuta para impedir repeticiones estratégicas del juicio. Tal como lo dijo la Corte en *Henderson*:

Subvertiría el proceso el hecho de que un acusado tuviera libertad general para llevar a la apelación una versión de los peritos que, si resulta estar bien

⁴¹ Véase, por ejemplo, *R. v. Mchkar* [2005] EWCA Crim. 1483; *R. v. Exsor* [1989] 1 WLR 497, CA.

⁴² Duran, 2003.

⁴³ Para una discusión crítica, véase EASTON, 2006: 675; HARRIS, 2000: 1.

fundamentada, pueda y deba presentarse al jurado [...]. El juicio oral no debe ser un «ensayo» de peritos. Las audiencias de apelación no deben presentar una oportunidad para llamar a nuevos peritos con la esperanza de que lo hagan mejor que aquellos cuya prueba fue previamente rechazada⁴⁴.

En muchos casos, la defensa ya se habrá dedicado a conseguir un dictamen pericial a su medida, al momento de concebir la estrategia procesal original. Esto fue, de hecho, lo que sucedió en *Henderson*. Por otro lado, si la Corte de Apelaciones hubiese permitido a la defensa aportar las pruebas que tenía en su haber —o que podría haber obtenido fácilmente— antes del juicio oral, pero que no aportó por razones tácticas, habría perversamente dejado en desventaja a aquellos acusados que pusieron todas sus cartas probatorias sobre la mesa en el juicio original.

[L]os informes periciales fueron obtenidos para el juicio oral por parte de la defensa en *Henderson* y la defensa eligió no llamar a todos esos peritos porque, en parte, le daban la razón a la acusación. En tal caso, un apelante no debe estar en mejor posición que otro apelante que ha llevado sus pruebas al juicio⁴⁵.

Por supuesto, los abogados defensores son muy conscientes de los límites institucionales respecto a las repeticiones estratégicas del proceso y son muy sofisticados al intentar pedir la apelación con el desesperado argumento de que procurarán hacerlo mejor en un segundo intento. Por el contrario, los abogados defensores se aprovechan de la literalidad del art. 23 de la *Criminal Appeal Act* de 1968, intentando alegar que un segundo (o tercero, o más) testimonio pericial constituye «una prueba que no fue aportada en el juicio al que se refiere la apelación»; en otras palabras, que se trata de una «nueva prueba». Ésta es una estrategia prometedora al menos por dos buenas razones. Primero, un perito que no ha sido escuchado en el juicio oral literalmente sí ofrece pruebas «nuevas», aunque en un sentido trivial —es, al menos, un *testigo* * nuevo—. Segundo, cualquier elemento de prueba puede ser individualizado y re-descrito de infinitas maneras diversas, de tal forma que siempre es posible sostener que la prueba pericial contiene algo *substancialmente* nuevo, cambiando el enfoque de su testimonio o cambiando las palabras de sus conclusiones, por ejemplo, para expresar mayor respaldo a los puntos de vista de la defensa o mayor escepticismo acerca de las pruebas de la acusación. De manera alternativa, un defensor entusiasta podría instruir a nuevos peritos de disciplinas o subespecialidades diferentes a las de aquellos a los que se llamó durante el proceso, resultando casi inevitablemente en variaciones de la me-

⁴⁴ [2010] 2 Cr. App R. 24 [3].

⁴⁵ *Ibid.*

* *N. del T.*: En el sistema procesal adversarial, el perito es un testigo de parte. Por eso, el autor utiliza expresiones como *expert witness* y similares. Para la traducción se ha preferido la expresión «perito» salvo cuando, por literalidad o para evitar repeticiones, sea necesario utilizar «testigo», aunque se ha empleado también, donde correspondiera, la expresión «testimonio pericial».

todoología, los datos, el énfasis o las expresiones, que pueden ser presentados como «nuevas pruebas» y que potencialmente socavan la seguridad de la condena. Nuevamente, las cortes de apelaciones necesitan estar dispuestas a todo para develar y repeler esas maniobras, dado que «la dificultad para aplicar el art. 23 en los casos que dependen completamente de la prueba pericial es mucho mayor»⁴⁶.

En la apelación de *Henderson* se decidió finalmente que la prueba que se presentaba no era realmente nueva, aunque la Corte mantuvo su decisión original de escuchar otros peritos en la apelación en virtud de la complejidad de los temas y su exposición a la controversia científica.

[L]as pruebas nuevas, en esencia, no eran nada nuevas. Los peritos eran «nuevos», pero sus pruebas no. Eran pruebas de peritos en las mismas disciplinas a las que pertenecían aquellos informes que la defensa había obtenido previamente y decidió no llamar. Los representantes anteriores del acusado eligieron no aportar la prueba pericial porque favorecía a la acusación. Ésa fue una elección dentro de los límites de lo razonable. Si hubieran aportado esa prueba y el jurado hubiese condenado, probablemente no se habría permitido la aportación de nuevas pruebas producidas por expertos en las mismas disciplinas. No deben estar en mejor posición por el hecho de haber realizado una prudente elección previa, consistente en no llamar a los peritos⁴⁷.

La posición en *Oyediran* fue más sencilla. Aquí, el acusado simplemente intentó volver a ventilar puntos previamente recogidos en el juicio o que se escucharan nuevamente argumentos acerca del estado de la ciencia médica que ya habían sido escuchados y rechazados, en el anterior caso *Harris*⁴⁸. En la apelación, el equipo de la defensa vio la posibilidad de aportar una nueva prueba pericial señalando la posibilidad de que Femi podría haber caído accidentalmente y luego haber sido atrapado por su madre, sin embargo, la Corte de Apelaciones consideró que esto era básicamente reformular el argumento ya presentado ante el jurado. Otro perito de la defensa, la Dra. Squier, especuló que Femi podría haber padecido VIH, pero esta posibilidad fue absolutamente rechazada por un patólogo especializado en VIH pediátrico que fue consultado justamente por sugerencia de la Dra. Squier. En consecuencia, la Corte de Apelaciones no pudo distinguir una prueba verdaderamente nueva en la estrategia de la defensa en la apelación ni alguna que pudiera «suministrar algún fundamento para conceder la apelación»⁴⁹.

⁴⁶ *Ibid.* [4].

⁴⁷ *Ibid.* [81].

⁴⁸ *R. v. Harris*; *R. v. Roek*; *R. v. Cherry*; *R. v. Faulder* [2006] 1 Cr. App R. 5 [2005] EWCA Crim. 1980.

⁴⁹ CAA 1968, s. 23 (2)(b).

4. ERRORES DE DERECHO Y DETERMINACIÓN DE LOS HECHOS EN LAS APELACIONES PENALES DE CONDENAS DICHTADAS POR JURADOS

En el sistema procesal penal inglés se pueden esquivar totalmente cuestiones epistemológicas difíciles si se determina que el proceso original y el veredicto están afectados por un error sustancial de derecho. En estos casos la atención se dirige particularmente a aquello que el juez de primera instancia le dijo al jurado al darle «instrucciones» como parte de la recapitulación que hace al cerrar la fase probatoria del juicio oral⁵³, antes de que el jurado se retire y antes de las deliberaciones secretas para dictar su veredicto. Los jueces de primera instancia en Inglaterra y Gales recapitulaban tanto los hechos como el derecho⁵⁴. Esto incluye un resumen de las pruebas presentadas en el juicio y una explicación de su importancia: según los criterios jurídicos que el jurado debe tener en cuenta para determinar si el acusado es responsable de los hechos que se le imputan. Las recapitulaciones judiciales deben ajustarse a los hechos y a las circunstancias del caso presente. Este requisito supone que la recapitulación en los procedimientos complejos ocupe a veces varios días del proceso y pueda conllevar la transcripción de cientos de páginas. Si se demuestra que el juez de primera instancia desorientó al jurado con respecto a la ley penal sustantiva aplicable o bien con respecto a un aspecto fundamental del procedimiento penal, como la carga o el estándar de prueba⁵⁵, la condena será casi con toda certeza considerada insegura y efectivamente revocada en la apelación. La Corte de Apelaciones puede intervenir en este escenario sin miedo a interferir con la sagrada función del jurado como juzgador de los hechos, dado que el error fatal está en cómo se le presenta el caso al jurado y no en ningún aspecto del desempeño de éste. Revocar condenas en estas circunstancias no implica una crítica al estatus institucional o a la competencia epistemológica del jurado, podría incluso considerarse que sirve para reforzar la legitimación del jurado como juzgador de los hechos mediante la descalificación de la conducta judicial que hubiese podido —consciente o inconscientemente— desviar ilegítimamente sus deliberaciones y su veredicto.

Las apelaciones que se enfocan en las instrucciones judiciales al jurado son características de la preocupación del *common law* por el «control de la información que entra» en las deliberaciones del jurado, lo que a su vez refleja el muy limitado alcance que tiene el «control del resultado»⁵⁶, es decir, el

⁵³ ROBERTS Y ZUCKERMAN, 2010: 82-86.

⁵⁴ Véase JUDICIAL STUDIES BOARD, 2010.

⁵⁵ Véase, por ejemplo, R. v. *Alison* (2004) EWCA Crim. 1646; cfr. R. v. *Majid* (2009) EWCA Crim. 2563 (los jueces de primera instancia no están obligados a usar una fórmula exacta de palabras, siempre que se comuniquen clara y efectivamente lo esencial de las normas jurídicas).

⁵⁶ Véase DORAN, JACKSON Y SEIGEL, 1995: 1 y ss., esp. 18 (atribuyéndole esta terminología a Milan DAMASKA); FRIEDMAN, 1998: 1921, 1951-1952; DAMASKA, 1997: cap. 2.

control de los veredictos que se dictan «en la caja negra» (irracional), sean de «culpabilidad» o de «inocencia»⁵⁴. Aunque el control judicial «de la información que entra» está asociado, comprensiblemente, a las reglas de exclusión de pruebas, el mismo mecanismo institucional asegura reglas procesales que regulan la recapitulación judicial del juicio oral y el alcance de las inferencias jurídicamente admisibles. Estas «reglas de razonamiento forense»⁵⁵ necesariamente concierne a la información que ya ha sido presentada en el juicio, buena parte de la cual se traducirá en pruebas jurídicamente admisibles⁵⁶. Algunas características contrastantes de las apelaciones de *Henderson*, *Butler* y *Oyediran* ilustran con precisión la libertad que tienen los tribunales de apelación cuando se concentran en errores de derecho y, de ese modo, consciente o inconscientemente, eluden profundas controversias epistemológicas.

El caso *Butler* representó una especie de enigma respecto a la forma usual de los incidentes de bebés sacudidos y sus desalentadoramente predecibles consecuencias fatales; y esta característica tan excepcional, inmediatamente puso en duda la seguridad de la condena. No sólo Ellie Butler no murió por el daño cerebral sufrido, sino que además tuvo una recuperación completa. Esto no prueba necesariamente la inocencia de su padre. La acusación estaba preparada para llevar adelante el caso sobre la base de un amplio arsenal de pruebas periciales y el jurado estaba, obviamente, persuadido más allá de toda duda razonable de que Butler había infringido las lesiones intencional o imprudentemente a su hija⁵⁷. Pero la defensa también fue capaz de aportar testimonios de peritos médicos igualmente eminentes afirmando, por el contrario, que las lesiones de Ellie pudieron haberse producido de forma natural o mediante actividades completamente inocentes, como haberla llevado de paso en un cochecito. La Corte de Apelaciones confesó hallarse frente a un hercúleo conflicto de testimonios expertos:

Hubo tres especialistas en neuroradiología pediátrica que rindieron prueba, pero la esencia de la discusión estuvo entre dos de ellos: el Dr. Stoodley, presentado por la acusación, y el Dr. Anslow, presentado por la defensa. Ambos peritos tienen una considerable reputación en el campo de la neuroradiología pediátrica y ambos tienen considerable experiencia forense, tanto en la jurisdicción penal como en la de familia. La nada envidiable tarea del jurado

⁵⁴ R. v. *Mirza* (2004) 1 AC 1118 HL. Para la discusión, véase FERGUSON, 2006: 180; JACKSON, 2002: 477; SMITH, 1998: 98.

⁵⁵ ROBERTS Y ZUCKERMAN, 2010: cap. 15.

⁵⁶ En algunos casos el juzgador de los hechos será instruido para que ignore pruebas jurídicamente inadmisibles que han sido aportadas al juicio, o para que use la información sólo para ciertos propósitos jurídicamente permitidos y no para otros jurídicamente prohibidos. La eficacia práctica de instrucciones como éstas es notoriamente dudosa; cfr. WISTRICH, GUTRIE y RACOURNSKI, 2005: 1251. Pero es posible que dichas reglas tengan propósitos normativos o ideológicos más amplios, adicionales o incluso independientes de sus objetivos epistémicos visibles.

⁵⁷ El *mens rea* según el art. 20 OADA (*Offences Against the Person Act*) de 1861, para quien cause GBH (*grievous bodily harm*, lesión corporal grave), corresponde a falta de cuidado intencional o subintencional —es decir, voluntaria—. R. v. *Savage*, R. v. *Parmenter* [1992] 1 AC 699, HL.

no sólo consistió en tener que decidir entre estos dos puntos de vista, sino además hacerlo sin ninguna prueba diversa que les ayudase [...].⁵⁸

Afortunadamente, la Corte de Apelaciones relevó al jurado en la nada enviable tarea, dada la existencia de errores de derecho que podrían ser considerados, a los efectos de esa apelación, como decisivos.

En efecto, pese a tratarse de un caso difícil en primera instancia, *Butler* fue un caso fácil en apelación porque la recapitulación del juez de primera instancia tuvo múltiples defectos. La Corte de Apelaciones con frecuencia está preparada para perdonar pequeños fallos de expresión del juez e incluso errores jurídicos aislados, en tanto sean desviaciones inocuas del debido proceso; pero no extiende esa indulgencia cuando se trata de una serie de errores que se agravan mutuamente y que, en conjunto, sólo pueden verse como peligrosos para la seguridad de la condena. El error principal del juez en *Butler* tuvo que ver predominantemente con su resumen fáctico de las pruebas relacionadas con «la triada», en particular con la importancia de lo reconocido por los peritos de la acusación sobre que era muy raro que un niño cuyas hemorragias retinianas hubiesen sido causadas por una sacudida violenta, pudiese recuperarse completamente, tal como había sucedido con Ellie. La recapitulación del juez de primera instancia naturalmente se refirió al testimonio experto que hablaba de sangrado retiniano, pero no en términos que dejaran suficientemente clara su verdadera importancia probatoria frente al jurado, según la Corte de Apelaciones:

El juez correctamente observó que el jurado tendría que mirar las pruebas en su conjunto. Reconoció que la prueba oftalmológica debilitaba la triada. Pareció asumir la opinión según la cual ese debilitamiento se compensaba con la prueba de la hemorragia subdural. Esto, en nuestra opinión, no fue correcto. La recuperación demostró que las hemorragias retinianas no podrían considerarse prueba de las sacudidas. Por el contrario, eran prueba de una causa desconocida.⁵⁹

Una vez se aprecia la verdadera importancia de la prueba oftalmológica, resulta poco claro cómo el jurado pudo haber descartado racionalmente la posibilidad de que las lesiones de Ellie tuviesen origen en una causa desconocida:

[D]ebe haber una base racional a partir de la cual el jurado pueda concluir que el desplome no se debió a una causa desconocida y así llegar a una conclusión que rechace el peso probatorio que se asigna a la característica inusual en este caso, a saber, la completa recuperación.⁶⁰

La Corte de Apelaciones concedió la apelación en *Butler*, no con fundamento en discutibles (e inherentemente no verificables) especulaciones acerca

⁵⁸ [2010] 2 Cr. App R. 24 [95].

⁵⁹ *Ibid.* [106]-[107].

⁶⁰ *Ibid.* [107].

del razonamiento del jurado, sino sobre la base de la defectuosa forma (demostrable) en que el juez de primera instancia planteó el caso al jurado:

En ningún lugar de su decisión el juez reconoció completamente el peso a otorgar a la prueba oftalmológica no controvertida. En ningún lugar identificó la base sobre la cual el jurado podría haber rechazado la posibilidad de una causa desconocida [...]. El juez no instruyó de forma suficiente al jurado sobre la importancia de la prueba oftalmológica y sus efectos para socavar la triada como prueba de MAHI y como respaldo de una causa desconocida. Le correspondía hacerlo. No se dio una instrucción adecuada al jurado para que pudiese considerar la posibilidad de una causa desconocida, particularmente a la luz de la prueba oftalmológica, y así condenar únicamente si rechazaban esta posibilidad.⁶¹

A la luz de los estándares imperantes en apelación, la microscópica revisión de la Corte de Apelaciones al análisis fáctico del juez de primera instancia parece decididamente intervencionista. No se espera que las recapitulaciones judiciales deban entrar en minucias y valorar cada una de las inferencias fácticas no triviales que un jurado pueda potencialmente realizar a partir de las pruebas presentadas en el juicio oral. En el juicio por jurados se asume que éstos son expertos bien calificados en el razonamiento inferencial de sentido común.⁶² Dado lo anterior, los jurados legos pueden necesitar mayor asistencia del juez en el análisis de la lógica inferencial de la prueba pericial compleja, especialmente cuando ésta es contradictoria, como en *Butler*. Pero la insistencia de la Corte de Apelaciones en que el juez de primera instancia debió haber sido más explícito y haber dado mayores instrucciones sobre el «peso a otorgar a la prueba oftalmológica no controvertida», así como acerca de «la base sobre la cual el jurado podría haber rechazado la posibilidad de una causa desconocida»⁶³, es una instrucción inusual. La debatida prueba del sangrado retiniano había sido completamente examinada en el juicio y seguramente el jurado era plenamente consciente de que Ellie se había recuperado. De hecho, si la Corte de Apelaciones estaba tan segura de su propia conclusión, que no había «base racional a partir de la cual el jurado pueda concluir que el desplome no se debió a una causa desconocida»⁶⁴, es difícil imaginar cómo el juez de primera instancia pudo haber dejado el caso completamente en manos del jurado. En otras palabras, la situación comenzó a parecerse a una acusación muy similar a un «*no case to answer*»⁶⁵.

⁶¹ *Ibid.* [108]-[113].

⁶² Véase *R. v. Crowley* [2009] EWCA Crim. 1219; *R. v. Saleem* [2007] EWCA Crim. 1923 [44]-[49]; *R. v. Adams (No. 2)* [1998] 1 Cr. App R. 377, CA.

⁶³ [2010] 2 Cr. App R. 24 [108].

⁶⁴ *Ibid.* [110].

⁶⁵ *R. v. Galbraith* [1981] 1 WLR 1039, CA. Véase ROBERTS y ZUCKERMAN, 2010: 79-82.

* *N. del T.*: «*No case to answer*» es una figura del procedimiento penal inglés que permite al acusado pedir la cesación del proceso (con absolución), cuando considera que la acusación carece de pruebas que le incriminen, o éstas son tan débiles que no tiene sentido oponerse.

Aunque la crítica de la apelación sobre el análisis fáctico realizado por el juez en *Butler* se hubiese apartado más de la jurisprudencia de lo que la Corte misma admitió, también hubo otras «graves instrucciones equivocadas» que dieron un amplio «soporte convincente»⁶⁶ a las conclusiones de la Corte de Apelaciones. Destaca en la lista de los defectos el hecho de que «la recapitulación fue también defectuosa en su estructura»⁶⁷:

El juez recibió cuidadosa y conscientemente largos extractos de la prueba pericial al jurado. Pero no le «instruyó» en absoluto. Recitó aquellos extractos en el orden cronológico en el que fueron presentados. Esto fue de poca ayuda al jurado. El jurado necesitaba una cuidadosa instrucción en cuanto a los aspectos esenciales que debía determinar y un recordatorio de las pruebas y sólo de las pruebas relacionadas con dichos aspectos esenciales. Era necesario pormenorizar esas pruebas, elemento por elemento.⁶⁸

Recitar repetitivamente o recitar resúmenes de los relatos, en lugar de contextualizar la prueba y relacionarla con el derecho, es un pecado capital en las recapitulaciones. Este pecado se agravó más en *Butler* por la negativa del juez de primera instancia a eliminar la mención del art. 20 del cargo de crueldad infantil relacionado con las quemaduras de radiador de Ellie. Las decisiones judiciales sobre las solicitudes de la defensa de atenuar la acusación difícilmente son revocadas en apelación, pues se considera que las cuestiones sobre acumulación y atenuación entran dentro de la discrecionalidad del manejo del proceso por parte del juez de primera instancia.⁶⁹ En este caso, sin embargo, la Corte de Apelaciones procedió con su persistente «sentido de incomformidad»⁷⁰ y anuló también la condena de *Butler* por crueldad infantil.⁷¹ La Corte claramente pensó que el jurado había tratado las lesiones internas en la cabeza de Ellie y las quemaduras del radiador como mutuamente corroborativas e incriminadoras, cuando en realidad la prueba pericial indicaba que el daño cerebral de Ellie había sido muy probablemente producido por causas naturales desconocidas, mientras que su contacto con el radiador había sido posiblemente un accidente («descuido de padres primerizos»). No hizo diferencia alguna el hecho de que el juez hubiera elaborado un conjunto de instrucciones advirtiendo al jurado que debía tratar las lesiones por separado, pues el incidente del radiador era «el único ejemplo concreto de una conducta impropia» por parte del acusado y que «era mucho pedir al jurado que lo ignorase al considerar las lesiones en la cabeza».⁷² El efecto acumulativo de

⁶⁶ [2010] Cr. App R. 24 [112]. La Corte anotó correctamente que el juez de primera instancia no había tenido la ventaja, al formular su recapitulación, de la posterior sentencia de la Corte en *R. v. Schmidt* [2009] EWCA 2 Cr. App R. 24 [114].

⁶⁷ [2010] 2 Cr. App R. 24 [114].

⁶⁸ *Ibid.* [114]-[115].

⁶⁹ *R. v. Christie* [1997] AC [117, HL].

⁷⁰ [2010] 2 Cr. App R. 24 [83].

⁷¹ *Ibid.* [117].

⁷² *Ibid.* [88].

estos discretos errores teóricos permitió a la Corte de Apelaciones esquivar los proyectiles epistemológicos en *Butler*, tal como la propia Corte insinuó en el párrafo final de la sentencia que rechazó la apelación:

No necesitamos prestar atención a la nueva prueba patológica que buscaba aportar el apelante. Esa prueba es altamente controvertible. No nos ocupamos expresamente de la solicitud de nuevas pruebas, ni expresamos ninguna otra opinión al respecto.⁷³

La apelación en *Oyediran* también objetó las bases jurídicas de las decisiones sobre la admisibilidad adoptadas por el juez de primera instancia, pero esta vez con el resultado contrario —pese a que el apelante formuló los mejores de muchos argumentos jurídicos posibles—. La objeción se realizó sobre dos elementos de prueba en concreto, aportados al proceso por la acusación. La primera pieza probatoria se refería a «un incidente extremadamente desagradable» que ocurrió cuando el acusado fue visitado en su casa por un funcionario de la asociación para la vivienda *, quien declaró que el acusado había tratado con brusquedad a su compañera, Sophia Rudder, y luego había injuriado al propio testigo amenazándole con un bastón de caminar e instándole a que abandonara el apartamento. El mal rato había continuado hasta que Rudder intervino y exigió que al desventurado funcionario de viviendas se le permitiera irse. El funcionario declaró que había informado inmediatamente del incidente a la policía y que simultáneamente había realizado un procedimiento administrativo afirmando que la compañera del acusado y su hijo estaban en riesgo de violencia y necesitaban protección inmediata de los Servicios Sociales.⁷⁴

Esta prueba fue aportada por la acusación buscando la aplicación del art. 101 (1)(c) de la *Criminal Justice Act* de 2003 [«excepción (c)»], para el propósito específico y muy limitado de demostrar que Sophia Rudder, la coacusada en el proceso, era capaz de actuar autónomamente y valerse por sí misma contra la personalidad dominante del acusado. El objetivo de la acusación era demostrar que Rudder tenía suficiente capacidad mental y libertad de elección como para ser considerada penalmente responsable por causar o permitir la muerte de Ferni. En la apelación se sostuvo, muy plausiblemente, que esta prueba nunca debió ser admitida dado que su escaso valor probatorio en los cargos contra Rudder era absolutamente desproporcionado en relación con el prejuicio injusto que se causaba a Oyediran. O la prueba no era realmente una «prueba explicativa importante» dentro del significado de la excepción (c)⁷⁵ o, si lo era, de todas formas debía ser excluida en razón del art. 78

⁷³ *Ibid.* [118].

* *N. del T.*: Las *Housing Associations* o asociaciones para la vivienda, son organizaciones de naturaleza privada encargadas de suministrar soluciones de acceso a la vivienda a precios bajos.

⁷⁴ *Ibid.* [164].

⁷⁵ CIA 2003, art. 102. Más adelante aclara que «la prueba es prueba explicativa importante si: a) sin ella, al juez o al jurado le resultaría imposible o difícil entender adecuadamente otras pruebas del caso, y b) su valor es sustancial para entender el caso en conjunto».

del PACE con fundamento general en la injusticia. La Corte de Apelaciones rechazó estas consideraciones y mantuvo la decisión del juez de primera instancia, afirmando:

Acceptamos que la prueba no tiene relevancia en cuanto a si el apelante ejerció o no violencia contra su hijo [...]. En nuestra opinión, la decisión de admitir pruebas que demuestran la capacidad de la señora Rudder para reaccionar a la conducta del apelante era relevante para un aspecto importante relacionado con ella y con la acusación. Puesto que era la única prueba directa, estaba dentro de los límites de una conclusión razonable del juez sostener que no se podría causar perjuicio alguno al apelante al admitir esta prueba.⁷⁶

En otras palabras, sin dar necesariamente una aprobación rotunda a la decisión del juez de primera instancia, la admisibilidad de la prueba fue considerada dentro de los parámetros aceptables de la discrecionalidad de éste para determinar puntos probatorios y procedimientos a lo largo del proceso. De acuerdo con la Corte de Apelaciones, «debió haber sido obvio para el jurado que el citado incidente [con el funcionario de la asociación para la vivienda] no mostraba violencia alguna o agresión de parte del apelante contra su propio hijo»; y en esta ocasión se consideró importante que el juez de primera instancia hubiera instruido al jurado diligentemente acerca de que «el incidente no tenía ninguna relevancia en cuanto a si el apelante tenía o no propensión a ser violento»⁷⁷. Esto es difícil de encuadrar con la observación de la Corte en *Butler*, con respecto a que era mucho pedir a un jurado mantener separados en su mente dos incidentes no relacionados, pese a que el juez de primera instancia hubiese dado una instrucción impecable a esos efectos. Un mejor fundamento para la ratificación de la decisión del juez de primera instancia puede encontrarse en un aparte de la Corte de Apelaciones donde se dice que «hay un cuerpo de pruebas sobre la conducta agresiva y enérgica del apelante, particularmente cuando se enfrenta a quienes él podría considerar que estaban en posición oficial»⁷⁸. Tal vez la prueba era admisible, posiblemente en el *common law*⁷⁹, y no susceptible de ser objetada por prejuicio, porque no le estaba diciendo al jurado nada que no supiera ya acerca de la actitud violenta de Oyediran.

Un segundo fundamento de la apelación tenía relación con una de las formas probatorias más controvertidas y más usualmente empleada en el procedimiento penal inglés, como es la supuesta confesión al compañero de celda en la prisión. De acuerdo con el sujeto con quien el acusado compartió una celda de detención preventiva en la prisión de Wandsworth, aquél confesó haber ma-

⁷⁶ [2010] Cr. App R. 24 [167], [168].

⁷⁷ *Ibid.* [167].

⁷⁸ *Ibid.*

⁷⁹ La Corte de Apelaciones, *ibid.* [169], expresamente hace esta salvedad, pero la prueba relacionada con la capacidad de un coacusado normalmente se consideraría directamente vinculada a la acusación, sin tener que apoyarse en el mal carácter «subyacente».

tado a su hijo mediante una violenta sacudida. El acusado supuestamente admitió que a menudo se estresaba y tomaba a su hijo sosteniéndolo boca abajo por los tobillos, sacudiéndolo hacia atrás y hacia adelante y que esto lo habría hecho el día de la muerte de Femi. Hubo muchas dificultades probatorias con esta supuesta confesión, además de los problemas de credibilidad inherentes que casi siempre acompañan a las profesiones hechas al compañero de celda.⁸⁰ El testigo tenía antecedentes penales por delitos graves; su afirmación contradecía hechos básicos conocidos (por ejemplo, sostenía que el acusado había dicho que Femi tenía tres o cuatro años de edad); y era aparentemente incapaz de estructurar la historia claramente, incluyendo los detalles acerca de los tiempos de la supuesta confesión del acusado que podían ser rotundamente contrarios por pruebas independientes. Al final del proceso, la credibilidad del testigo estaba destrozada y la acusación efectivamente lo rechazó*. Sin embargo, pese a haber escuchado de nuevo la letanía de defectos en la prueba, el juez de primera instancia indicó al jurado que debería otorgarle credibilidad a la supuesta confesión sólo si estaban seguros de que era cierta.

En la apelación se sostuvo que la supuesta confesión al compañero de celda ya no era susceptible de credibilidad y que debería ser desechada de la consideración del jurado. La Corte de Apelaciones estuvo más o menos de acuerdo con esta afirmación, pero no pensó que hubiera existido ningún daño colateral a la seguridad de la condena:

Si nosotros hubiésemos estado siguiendo el juicio oral, probablemente habríamos instruido al jurado para desechar la prueba. Era tan defectuosa que tenía poco valor. Pero el hecho cierto de que fuese tan evidentemente poco fiable sólo podía beneficiar al apelante. Era una prueba admisible y no podemos decir que el juez se equivocó al permitir que la prueba se presentase o por no advertir más fuertemente al jurado, ni tampoco podemos decir que dicha prueba hace que el veredicto del jurado sea inseguro⁸¹.

Con todo respeto, es difícil ver cómo la no fiabilidad de la confesión «sólo podía beneficiar al apelante», a menos que el jurado hubiese tenido que concluir que el caso de la acusación era tan débil que tuvo que recurrir a pedirle a condenados por delitos sexuales que cometiesen perjurio! En realidad, la acusación presentó un caso circunstancial contundente contra el acusado, donde la confesión al compañero de celda serviría apenas como corroboración incidental. Resulta más plausible pensar que el jurado o bien ignoró comple-

⁸⁰ Véase, por ejemplo, *R. v. Nidds* [2008] EWCA Crim. 148; *Labrador v. R.* [2003] 2 Cr. App R. 390, PC.

⁸¹ *N. del T.*: Que la acusación haya rechazado una prueba de cargo tiene notable importancia en la lógica del sistema adversarial. El hecho de que una parte rechace como suyo a un testigo significa que podrá «atacarlo», por ejemplo, presentando pruebas para desvirtuar su confiabilidad. lo que implica que podrá tener un interrogatorio cruzado con éste y ya no un interrogatorio directo como se hace con los testigos propios.

⁸¹ [2010] 2 Cr. App R. 24 [177].

tamente la confesión hecha en la prisión (como afirmó la defensa en la apelación, el juez de primera instancia debió haber instruido al jurado a hacerlo), o bien la asimiló vagamente en el marco de sus impresiones generales sobre el acusado, viéndolo como alguien con personalidad desconfiada, con muy mal carácter y mucho que explicar.

La observación más bien débil de la Corte de Apelaciones acerca de que la prueba «era admisible» y su conclusión optimista acerca de la seguridad de la condena del acusado, sólo se pueden aceptar si se comparte la evidente confianza de la Corte en que el acusado era obviamente culpable. En términos claros, estas decisiones jurídicas sólo son sostenibles (si es que son sostenibles) sobre la base de algunas asunciones epistemológicas muy sólidas.

Si *Bulter* demuestra cómo la Corte de Apelaciones pudo descartar antipadadamente cuestiones epistemológicas difíciles tratando los errores de derecho como determinantes en la apelación de una condena, *Oyediran* ilustra los límites de cualquier análisis teórico meticuloso, revelando cómo a menudo las conclusiones sobre puntos de derecho, implícitamente si no es que explícitamente, descansan en conclusiones sobre los hechos. La deferencia algo elástica de las cortes de apelaciones hacia la discrecionalidad del juez de primera instancia en cuestiones de procedimiento penal y pruebas puede verse, en un análisis más cercano, como una estira y afloja para calibrar la fuerza probatoria del caso presentado por la acusación. La fuerza del caso de la acusación es, a su vez, un factor muy importante en la valoración de la «seguridad» de una condena en apelación.

La otra cara de la preocupación de la Corte de Apelaciones por la rectitud jurídica del procedimiento penal es que las «dudas que acechan» al veredicto del jurado generalmente son pasadas por alto si el juicio se ha llevado a cabo de manera justa, sin errores de derecho y si no hay (genuinamente) nuevas pruebas que el acusado pueda aportar para apoyar su pretensión de inocencia. *Henderson* es un ejemplo clásico de un caso en que la Corte de Apelaciones abrigaba dudas acerca de la culpabilidad del acusado, pero sin una intensidad suficiente como para impelerla a sustituir el veredicto del jurado por su propia valoración de la prueba. El resultado, incluso donde subsisten «misterios no resueltos», es que debe ratificarse el veredicto del jurado:

[Q]ueda por resolver el misterio de cómo una niñera tan reputada como la apelante pudo haber sido responsable del uso excesivo de fuerza, aunque fuese momentánea, al cuidar a este bebé. Pero ése era un problema con el que el jurado tenía que lidiar. No hay fundamentos para que esta Corte pueda decir que el jurado no estaba autorizado a [...] concluir que la prueba pericial demostraba más allá de toda duda razonable que la acusada había sacudido a Maeve con fuerza excesiva.⁸²

⁸² *Ibid.* [74].

Este razonamiento puede parecer austero e incluso indiferente. Sin embargo, la preocupación por un apelante que posiblemente había sido condenado erróneamente no es el único interés importante en juego al evaluar la seguridad de una condena penal. La Corte de Apelaciones en *Henderson* fue sumamente consciente de los intereses del bebé fallecido y de sus apenados padres (por no mencionar el amplio interés público en llevar ante la justicia a los asesinos de menores, incluyendo a aquellos que matan sin intención):

Seguimos preocupados sobre cómo la apelante se ve en la nada enviable posición de intentar establecer que el veredicto del jurado es inseguro, habiendo purgado la dura condena de tres años. Pero nuestra simpatía por todos los implicados, los padres de Maeve y la apelante, no puede subvertir nuestra función. No podemos sustituir el veredicto del jurado por una conclusión basada en la perplejidad de cómo la apelante pudo haber tratado a Maeve con fuerza injustificada, aunque fuese momentáneamente. Éste es un aspecto que fue resuelto por el jurado.⁸³

En términos epistémicos, la decisión de la Corte de negar la apelación en *Henderson* fue apuntalada por algunas generalizaciones banales acerca de la conducta humana.⁸⁴ Sabemos que a veces la gente buena hace cosas malas —actúan «fuera de sí», como se dice de forma reveladora— y más específicamente, sabemos que aun los padres amorosos a veces dañan a sus hijos en la desesperación de un momento de pérdida de autocontrol, precipitado por la fatiga, frustración o por sentimientos de incompetencia o incapacidad para sobrellevarlo —todas éstas son reacciones perfectamente normales frente a las exigencias físicas y psicológicas que plantea el cuidado de un recién nacido—. Los bebés y los niños son frágiles y obviamente están completamente desvalidos y son incapaces de defenderse por sí mismos ante un ataque. Otra pieza de la sabiduría del sentido común aplicable a este escenario es que la historia nunca predice completamente eventos futuros: hay «siempre una primera vez para todo». El hecho de que *Henderson* fuese una experimentada y respetada niñera con hijos propios, ciertamente hace relativamente poco probable que de repente cambiara y violentamente sacudiera y lesionara fatalmente a un niño que estaba a su cuidado, pero esto no excluye la posibilidad de que, justo esta vez, lo hiciera con trágicas consecuencias. Y el llanto constante de los bebés (como la conducta de los niños en general) puede ser extremadamente provocador.⁸⁵

En ausencia de generalizaciones pertinentes sobre la conducta humana y la vida familiar para soportar el peso de las inferencias fácticas específicas del

⁸³ *Ibid.* [80].

⁸⁴ Véase adicionalmente, ANDERSON, SCHON y TWINING, 2005: cap. 10; ANDERSON, 1999: 455.

⁸⁵ Un dicho de los padres, muy común en mi infancia, decía «*It'll swing for you!*», que significaba que un padre o una madre llevados al límite, podrían matar a su hijo y ser llevados a la horca por ello. El hecho de que la pena de muerte haya sido abolida recientemente en Inglaterra y Gales no cambió para nada el significado ni redujo la efectividad de este clamor por un mejor comportamiento.

caso, el veredicto de culpabilidad del jurado (por homicidio involuntario) habría sido ininteligible, presumiblemente irracional e insostenible en apelación. Pero con el lastre epistemológico de las generalizaciones de sentido común, el jurado tenía fundamentos para concluir, a partir de las pruebas, que un evento que, pese a su improbabilidad, *podría haber* ocurrido, *ocurrió* en el caso presente. Dado que el procedimiento penal inglés otorga al jurado, y no a un juez profesional o a un tribunal, la responsabilidad de determinar los hechos de la acusación, la Corte de Apelaciones no tiene autoridad institucional (ni epistémica) para sustituir sus propias conclusiones inferenciales en ningún «aspecto que el jurado ya haya resuelto».

Además, la posibilidad de permitir la apelación con fundamento en un error jurídico, que demostró ser tan útil en *Butler*, no existía en *Henderson*. La recapitulación del juez de primera instancia en *Henderson* no cometió el pecado capital de ser una perorata repetitiva que descontextualizara al jurado:

Eran «instrucciones», no un recorrido desestructurado a través de la prueba pericial. Le aclaró al jurado los aspectos que tenía que decidir y los fundamentos rivales sobre los cuales podría rechazar o aceptar las pruebas. La ruta que tomó el jurado para la condena fue bastante clara, a partir de la lectura de la recapitulación.⁸⁶

La Corte de Apelaciones se mostró particularmente impresionada por la forma en que el juez de primera instancia manejó la posibilidad de que hubiera podido haber alguna causa desconocida de muerte no planteada por ninguna de las partes en el proceso.

[E]l juez instruyó directamente al jurado sobre la necesidad de tener en cuenta que la ciencia médica en relación con las lesiones en la cabeza no occidentales está en desarrollo y es, como lo sostuvo la defensa, incierta, y es posible que no se pueda identificar la causa de la muerte. Dio una instrucción positiva y correcta respecto a que el jurado debía considerar si la muerte podría ser inexplicable. Les instruyó que debían estar seguros de que la muerte inexplicable podría ser excluida.⁸⁷

La última formulación, al menos en su versión en inglés*, suena extraña en un lenguaje coloquial, pero es jurídicamente correcta. El jurado debe estar «seguro de que la muerte inexplicable podría ser excluida» porque el deber de la acusación, de acuerdo con la presunción de inocencia, es *desvirtuar* la posibilidad de causas inexplicables según los criterios penales. Así, el juez de primera instancia correctamente «enfaticó que la defensa no tenía obligación alguna de proponer una teoría particular sobre la causa de

⁸⁶ [2010] 2 Cr. App R. 24 [76].

⁸⁷ *Ibid.* [77].

* *N. del T.*: Dice el texto original: «He gave a positive and correct direction that the jury had to consider whether the death may be unexplained. He directed them that they had to be sure that unexplained death could be excluded».

la muerte»⁸⁸. Sin embargo, en la medida en que la defensa no se apoyó en ninguna teoría afirmativa de la causa inexplicada de la muerte, sino que más bien se dedicó a negar completamente el ilícito, el jurado fue adecuadamente instruido respecto a que «el rechazo de la teoría del desplome espontáneo en sí mismo no significa que el jurado esté obligado a aceptar el caso de la acusación y condenar»⁸⁹. Así, el análisis fáctico del juez de primera instancia en *Henderson* fue consistente con la carga y el estándar de prueba ortodoxos y no podría ser objetado a partir de la base (en cierta forma atenuada) sobre la que se criticó la recapitulación en la apelación del caso *Butler*. «De acuerdo con ello», dijo la Corte de Apelaciones, «éste es un caso en el que el asunto de la causa inexplicada, en un área donde la ciencia médica se encuentra en desarrollo, fue adecuadamente expuesto ante el jurado. La justificación del rechazo de esa posibilidad y de la aceptación del caso de la acusación es clara, dada la recapitulación»⁹⁰.

5. PRECEDENTES JURÍDICOS Y HECHOS CIENTÍFICOS

Los sistemas jurídicos de la familia del *common law* adoptan la doctrina del *stare decisis*, que significa —en términos generales— que los jueces crean derecho vinculante mediante la decisión de casos «precedentes» que los tribunales inferiores (incluyendo los juzgados penales de primera instancia) deben seguir y aplicar⁹¹. Aparentemente, los peritos son testigos sobre los hechos y emiten dictámenes en relación con su área de conocimiento. Ciertamente no declaran sobre cuestiones del derecho inglés⁹², de manera que no se les puede plantear ninguna pregunta sobre los precedentes. Sin embargo, los teóricos han apuntado la tendencia de algunos tribunales a tratar anteriores decisiones de admisibilidad como precedentes que autorizan la admisión de ciertos tipos de pruebas científicas o de otros tipos de pruebas periciales en casos subsecuentes. Este enfoque «que reconoce autoridad a lo precedente» sobre la admisibilidad de la prueba pericial ha sido criticado por Simon COLE⁹³ y otros teóricos estadounidenses⁹⁴, quienes sostienen que un enfoque formalista de la admisibilidad no logra someter a la prueba científica a los estándares adecuados

⁸⁸ *Ibid.* [78].

⁸⁹ *Ibid.*

⁹⁰ *Ibid.* [79].

⁹¹ Tal como observó Lord Bingham en *Key v. Lambeth LBC* [2006] 2 AC 465 [2006] UKHL 10 [42]: «Aunque la adhesión al precedente ha sido ridiculizada por algunos, aun desde los tiempos de BENTHAM, como una receta para la perpetuación del error, ha sido la piedra angular de nuestro sistema jurídico [...] [y] "un fundamento indispensable a partir del cual decidir cuál es el derecho y su aplicación a casos individuales"».

⁹² El derecho extranjero, sin embargo, es una cuestión de hecho en el procedimiento inglés, y en consecuencia un tema susceptible de prueba pericial: por ejemplo, *Attorney-General of New Zealand v. Ortiz* [1984] AC 1 45, HL.

⁹³ COLE, 2004: 1189.

⁹⁴ Cf. RISINGER, DENBEAUX y SAKS, 1989: 731, nota 182. También, ANÓNIMO, 2010: 2021.

dos de validez y de demostración de su fiabilidad. El hecho de que otro juez haya admitido previamente una prueba no garantiza su fiabilidad, si el primer juez tampoco logró evaluar satisfactoriamente si la prueba estaba basada en principios científicos sólidos. Al mismo tiempo, el hecho de que nuevas teorías y técnicas científicas hayan recibido un reconocimiento jurídico por ser admitidas como prueba en procesos penales, puede ser visto por la sociedad en general como el otorgamiento de cierta credibilidad social a una ciencia novedosa que escasamente la merecería mediante un criterio científico objetivo. De esta forma, la «coproducción»⁹⁵ de la prueba científica, a falta de una validación científica adecuada y de un sometimiento a comprobación, puede engendrar errores judiciales atribuibles a una prueba pericial superficialmente persuasiva pero no fiable.⁹⁶

Las implicaciones diagnósticas de «la triada» ejemplifican la controversia ciencia forense novedosa. Antes de *Henderson*, el asunto había sido discutido en *Harris*⁹⁷. En ese caso, la Corte de Apelaciones hizo un recuento de los orígenes de la «hipótesis unificada» en un artículo científico conocido como «Geddes III» y de las aclaraciones subsiguientes de su autora en su testimonio ante la Corte. La hipótesis unificada postulaba que la triada putativa de síntomas podía ser atribuible a una causa única, en cuyo caso el valor probatorio de la triada para diagnosticar las sacudidas deliberadas estaría significativamente disminuido:

Quando se publicó Geddes III fue, y sigue siendo, muy controvertido. No es exagerado decir que este artículo generó un intensísimo debate en la profesión médica, tanto nacional como internacionalmente. En el curso de la audiencia de estas apelaciones hemos escuchado testimonios de una gran cantidad de muy distinguidos peritos médicos de diferentes especialidades, la mayoría de los cuales han expresado en sus testimonios puntos de vista en uno u otro extremo del debate. Sin embargo, desde muy pronto en la audiencia se hizo evidente que la parte sustancial del fundamento de la hipótesis unificada no podría sostenerse. La Dra. Geddes, al comienzo de su interrogatorio cruzado, aceptó que la hipótesis unificada nunca fue desarrollada para ser demostrada judicialmente. Dijo que estaba destinada a generar debate. Adicionalmente, aceptó que la hipótesis podía no ser muy correcta.⁹⁸

La Corte de Apelaciones, citó el siguiente revelador pasaje del interrogatorio cruzado de la Dra. Geddes, en el que el *Crown Counsel* llamó la atención sobre el hecho de que los abogados de la defensa hubieran invocado regularmente su teoría para minar la credibilidad de la prueba pericial de la acusación sobre (supuestos) casos de bebés sacudidos:

P.: Dra. Geddes, se están presentando casos a lo largo y ancho del país en los que Geddes III se cita por la defensa una y otra vez como demostración de que la teoría establecida está equivocada.

R.: Lo lamento mucho. No es un hecho; es una hipótesis pero, como ya lo he dicho, así es la explicación tradicional [...] Me haría sentir muy triste pensar que los casos han sido desechados sobre la base de que mi teoría es un hecho. Le preguntamos al editor si podíamos poner en la parte superior «Publicación de una de hipótesis» y no lo hizo, pero nosotros usamos la palabra «hipótesis» a lo largo de todo el artículo.

En este tenso intercambio sobre la «coproducción» jurídico-científica de pruebas científicas con potenciales efectos perniciosos en el litigio penal, la Dra. Geddes insistió en que ella nunca pretendió que su «hipótesis» fuese tratada como un hecho establecido y que intentó reivindicar su carácter provisional en el título de la publicación original (intento que fue frustrado por el editor). La consecuencia fue que los abogados defensores tomaron el artículo fuera de contexto y tergiversaron su significado. Pero igualmente, la hipótesis unificada no fue inventada por abogados. Si no hubiese sido desarrollada por la Dra. Geddes y sus colegas, ni publicada como Geddes III, no hubiese estado disponible para el (mal) uso por parte de los abogados defensores y sus peritos. Por otro lado, dado que la carga de la prueba en el estándar penal recae de lleno sobre la acusación, la defensa puede perfectamente invocar hipótesis científicamente plausibles que sean consistentes con la inocencia del acusado, para sugerir que la acusación no ha logrado probar la NAHI, es decir, que el bebé fue sacudido.

Después de *Harris*, la hipótesis unificada planteada en Geddes III ya no podía considerarse una base autosuficiente para una absolución. Sin embargo, la Corte de Apelaciones no volvió por defecto a la asunción opuesta de que la triada es un diagnóstico infalible de NAHI causada por sacudidas. Observando que «sigue habiendo un cuerpo de opiniones médicas que no aceptan que la triada es una herramienta infalible para este diagnóstico»⁹⁹, la Corte del caso *Harris* concluyó:

Aunque en sí misma es un fuerte indicio de NAHI, no creemos que sea posible decidir que deba automática y necesariamente conducir a un diagnóstico de NAHI. Todas las circunstancias, incluyendo el cuadro clínico, deben ser tenidas en cuenta.¹⁰⁰

En las consideraciones subsiguientes en *Henderson*, la Corte de Apelaciones se vio en apuros no sólo para confirmar que el valor probatorio de la triada seguía siendo el establecido en *Harris*, sino también para reiterar la ortodoxa posición jurídica de que los hallazgos fácticos no pueden crear precedentes vinculantes para casos posteriores:

⁹⁵ JASANOFF, 2006; 1997.

⁹⁶ ROACH, 2009; 67; MANOOKIN, 2008; 1009; WALKER Y STOCKDALE, 1999; JONES, 1994; cap. 10.

⁹⁷ R. v. *Harris*; R. v. *Rock*; R. v. *Cherry*; R. v. *Faulder* [2006] 1 Cr. App. R. 5 [2005] EWCA Crim.

1980.

⁹⁸ *Ibid.* [58].

⁹⁹ *Ibid.* [69].

¹⁰⁰ *Ibid.* [70].

Con frecuencia se observa que la conclusión de cualquier tribunal sobre la prueba médica, tanto en primera instancia como en apelación, depende de la prueba presentada ante ese juez. Ninguna jurisprudencia sobre las apelaciones podría dotar de autoridad a una proposición médica. La fuerza de una proposición en medicina depende de la fuerza de la evidencia médica en la que se basa. La calidad y alcance de la prueba inevitablemente variará caso por caso. Aunque ahora está comúnmente aceptado que la tríada es, prima facie, una prueba fuerte de las sacudidas, esto depende de la aceptación común de los expertos en la materia y no de las conclusiones de los jueces que sólo son capaces de pesar las pruebas que se les presentan. Una autoridad jurídica previa no puede determinar si la conclusión de un informe médico debe ser aceptada o rechazada. Lo máximo que puede hacer una autoridad jurídica es presentar un registro preciso de lo que fue o no aceptado o propuesto.¹⁰¹

Éste es un mensaje que pretende llegar a múltiples audiencias. Los jueces de primera instancia y los abogados quedan notificados de que las decisiones sobre la admisibilidad y las instrucciones judiciales que informan el razonamiento inferencial del jurado, no deben tomarse asumiendo perezosamente que hay un enfoque particular garantizado sólo porque otro juez lo ha aceptado en el pasado. Este autorreferencial «reconocimiento de autoridad a lo precedente», para emplear la terminología de COLE, no debe emplearse como un sucedáneo aceptable de la validez científica o del valor probatorio.¹⁰²

La reiteración de principios jurídicos elementales por parte de la Corte, fue también dirigida de forma muy consciente a los peritos médicos: «Hacemos hincapié en este problema», explicó MOSES, L. J., «porque tememos que la profesión médica pueda dirigirse a los tribunales para resolver controversias médicas».¹⁰³ La Corte de Apelaciones claramente quiso abandonar una noción fuerte de «coproducción» del conocimiento médico en este contexto particular. Un punto focal de estas preocupaciones fue un informe producido por el Royal College of Pathologists que buscaba identificar áreas de consenso y de desacuerdo entre especialistas sobre el controvertido punto de los traumas craneales en los niños. La autoridad del informe estaba algo comprometida por el hecho de que el borrador de éste no había sido acordado por todos los que contribuyeron en él. Pero en cualquier caso, observó la Corte de Apelaciones, el informe en sí mismo, cualquiera que sea su estatus para los profesionales de la medicina, no fue directamente una prueba en el proceso actual y no podría haber preguntado alguna de un tribunal judicial que pueda interferir en la sustancia de las disputas científicas entre peritos médicos:

No hemos podido tener en cuenta ese informe excepto hasta que nos fue presentado como prueba [...] debemos enfatizar que este razonamiento consti-

¹⁰¹ [2010] 2 Ct. App R. 24, 6.

¹⁰² Véase *supra*, nota 93.

¹⁰³ [2010] 2 Ct. App R. 24 [7].

tuye una autoridad jurídica, no médica, y nada le suma ni le resta a la fortaleza de la prueba suministrada por la tríada.¹⁰⁴

Tal vez no es sorprendente que los patólogos y otros médicos y científicos forenses, especialmente aquellos cercanos a la idea (o a la terminología) de «coproducción», puedan albergar concepciones equivocadas del papel de los tribunales en la resolución de disputas científicas surgidas en procedimientos jurídicos. Tanto los juristas como los científicos están mutuamente desinformados acerca de los marcos teóricos del otro, sus prácticas institucionales, sus principios metodológicos, sus asunciones culturales, sus normas y estándares profesionales. Pero ésa es sólo una parte de la historia. Hay una diferencia por trazar, sutil pero importante, entre i) las conclusiones inferenciales justificadas por una determinación de los hechos propia del sentido común, y ii) si esas conclusiones dan suficiente respaldo epistémico a una condena penal. La primera es —jurídicamente— una cuestión de hecho, pura y simple.¹⁰⁵ La segunda es una cuestión acerca de si las inferencias particulares son suficientes para satisfacer un estándar jurídico y esto, en la teoría jurídica (del *common law*), es una cuestión de derecho.¹⁰⁶ Así, la Corte de Apelaciones puede decir que, *dado el conocimiento científico actualmente disponible* (tal como lo entiende la Corte, basada en las pruebas aportadas al caso en cuestión), la tríada es una prueba sólida pero no concluyente de una sacudida no accidental; y la hipótesis unificada es incapaz, por sí misma, de establecer una duda razonable. Pero la Corte de Apelaciones no dice, y no puede decirlo consistentemente con las premisas epistemológicas de base de la prueba pericial admisible, que esto equivale a hacer algún tipo de contribución a la resolución de las disputas científicas subyacentes. En consecuencia, subsiste tanto para la acusación como para la defensa la posibilidad de reabrir cualquier controversia científica previamente ventilada ante un tribunal, sin importar qué tan recientemente o qué tan frecuentemente haya sido tratada, sobre la base de que entretanto ese conocimiento científico ha avanzado y los tribunales necesitan reconsiderar el asunto nuevamente.

La jurisdicción penal inglesa está tan comprometida con el principio fundamental de que las cuestiones de hecho deben ser determinadas por el jurado mediante deliberaciones libres sobre las pruebas, que los tribunales penales no están vinculados de forma alguna por ninguno de los *veredictos* dictados previamente, sean de culpabilidad o de inocencia, que tengan relación con los aspectos fácticos del proceso en cuestión.¹⁰⁷ A la inversa, a menos que el juez

¹⁰⁴ *Ibid.* [7].

¹⁰⁵ El concepto de hechos «puros y simples» levantará una sonrisa en los epistemólogos de línea dura. Lo reitero: este juego de palabras es, consciente y marcadamente, jurídico-institucional.

¹⁰⁶ Hay otras complejidades adicionales a esta distinción pero no se plantearán aquí. Véase ROBERTS y ZUCKERMAN, 2010: 130-137; ZUCKERMAN, 1986: 487.

¹⁰⁷ En otras palabras, no hay una teoría general para los impedimentos legales en el proceso penal inglés: R. v. Z. (*Abolución previa*) [2000] 2 AC 483, HL; DPP v. Humphrys [1977] AC 1, HL.

tenga información sobre algún cambio sustancial en el conocimiento científico o tecnológico subyacente que sirvió como fundamento, está autorizado a asumir como cuestión de derecho precedente que si la Corte de Apelación ha dicho que un tipo particular de prueba científica es suficientemente fiable para ser admisible en el proceso penal o que, una vez admitida, puede respaldar las inferencias indicadas, seguirá siéndolo hasta que se tenga nueva información.¹⁰⁸

6. INCERTIDUMBRE CIENTÍFICA Y DESACUERDO ENTRE PERITOS

A menudo se cree que la incertidumbre científica plantea graves dificultades para la confianza del sistema jurídico en la prueba pericial. Sobre esta creencia hay una corriente general y otra más específica.

En una versión bastante estándar de la filosofía «popperiana» de la investigación científica, la ciencia nunca prueba nada de manera concluyente. Simplemente somete las hipótesis teóricas a prueba a través de métodos experimentales, hasta que una hipótesis en concreto pueda considerarse como suficientemente respaldada, por su capacidad para soportar intentos genuinos de falsación y, entonces, puede calificarse como conocimiento científico. La ciencia, desde este punto de vista, es inherentemente incierta y está siempre abierta a mayor refinamiento y revisión y, ocasionalmente, incluso a reconsideraciones revolucionarias a la luz de descubrimientos científicos posteriores. Se dice que el característico método científico¹⁰⁹ es hostil a la determinación judicial de los hechos, que busca respuestas claras, inequívocas y razonablemente verdaderas, para cuestiones fácticas discutidas, con el fin de asegurar la irrevocabilidad de los veredictos penales.¹¹⁰ La corriente más específica sobre este tema, sostiene que la incertidumbre científica plantea dificultades importantes a la función jurisdiccional en relación con las áreas novedosas de la ciencia o con las aplicaciones tecnológicas innovadoras, o cuando expertos competentes en su campo expresan diferencias de opinión sustantivas acerca de disputas científicas no resueltas.¹¹¹ Buena parte de la ciencia novel exhibe ambas características. Es, por definición, incapaz de suministrar al derecho el tipo de respaldo epistémico para la determinación judicial de los hechos que se le pide y que se espera suministre la prueba pericial.¹¹²

¹⁰⁸ Véase, por ejemplo, *R. v. Reed and Reed*; *R. v. Gormson* [2010] 1 Cr App R 22; [2009] EWCA Crim. 2698; *R. v. Kempster* (No. 2) [2008] 2 Cr App R 19 [2008] EWCA Crim. 975; *R. v. Dallagher* [2003] 1 Cr App R 12 [2002] EWCA Crim. 1903; *R. v. Robb* (1991) 93 Cr App R 161, CA.

¹⁰⁹ BEECHER-MONAS, 2000: 1563.

¹¹⁰ HAACK, 2009: 1; WONDER, 1989: 75.

¹¹¹ Para ejemplos, véase RENDWANE, 1997: 1027; THOMPSON, 1993; KEE, 1986: cap. 14.

¹¹² En forma general, véase GOLDMAN, 2002: cap. 7; WARD, 2004: 369. La caracterización propiamente jurídica que más tiene los pies en la tierra sobre la epistemología de la prueba pericial, se encuentra en las conocidas observaciones de LAWTON, L. J., al caso fundacional *R. v. Turner* [1975] 1 QB 834,

Por supuesto, hay mucho de cierto en esta caricatura aproximada de la compleja relación entre el derecho y la ciencia, pero he llegado a pensar que las versiones fuertes de la historia de «dos culturas irreconciliables» son en buena parte obras de ficción.¹¹³ Es cierto que la determinación de los hechos en el proceso penal debe basarse en pruebas sólidas para sustentar la irrevocabilidad y legitimidad de los veredictos penales en un periodo prolongado de tiempo (al menos mientras el condenado está en prisión, por ejemplo, que podría ser treinta años o más en los delitos muy graves). Sería completamente inadecuado sustentar una condena penal —o, para el caso, una absolución— sobre la base de pruebas científicas frágiles y, por tanto, susceptibles de ser refutadas en cuestión de semanas o meses: éste era el defecto epistémico fundamental de la «hipótesis unificada» expuesta en *Harris*.

Las pruebas pueden tener un alto valor probatorio en términos de su contribución lógica para alcanzar conclusiones inferenciales y, sin embargo, ser *constitutivamente débiles* en cuanto a su fiabilidad y capacidad de suministrar respaldo epistémico. Sin embargo, debe enfatizarse que la evaluación de la idoneidad para fines sociales es una característica genérica de todas las aplicaciones científicas y tecnológicas. Un ingeniero civil probablemente no querría construir un puente en una carretera con materiales nuevos no probados, ni una aerolínea pondría en servicio comercial el prototipo de un avión con motores experimentales, como tampoco una compañía farmacéutica sacaría al mercado una nueva y milagrosa medicina no sometida a ningún tipo de pruebas, del mismo modo que los jueces no deben conformarse con dar el siguiente paso, que tiene consecuencias sociales, de imponer una censura y castigo penal sobre la base de pruebas científicas inaceptablemente dudosas o frágiles. Como he dicho antes, el derecho es especial no porque esté obligado a hacer estos cálculos y compensaciones inherentemente arriesgados¹¹⁴, sino porque lo hace en ambientes institucionales particulares, con objetivos y valores característicamente jurídicos, bajo un conjunto específico de limitaciones prácticas, etcétera.

Resulta muy ilustrativo ver cómo la Corte de Apelaciones en *Henderson* afrontó los desafíos de la incertidumbre científica y el desacuerdo entre peritos a la luz de consideraciones más amplias y debates prolongados en el tiempo. La triada y su reciente desafío por parte de la hipótesis unificada, pueden ser justamente vistas como ciencia (médica) novel y los desacuerdos entre peritos fueron una característica de las tres apelaciones conjuntas en *Henderson*. Sin

841. CA: «La opinión de un perito es aceptable para suministrar al tribunal información científica que probablemente esté fuera de la experiencia y el conocimiento de un juez o jurado. Si sobre los hechos probados el juez o el jurado pueden formarse una opinión propia sin ayuda, entonces la opinión del experto no es necesaria».

¹¹³ Una opinión se encuentra desarrollada en ROBERTS, 2012.

¹¹⁴ ¿Se habrá cobrado menos vidas la aviación civil si hubiéramos esperado a la invención del Boeing 747? ¿Tal vez habrían colapsado menos puentes si Telford hubiese nacido cien años antes?

embargo, como hemos visto, la Corte fue capaz de descartar anticipadamente cualquier investigación seria sobre los hechos y los testimonios científicos debatidos en *Butler*, mientras que el diagnóstico de la triada no era central para el argumento de la acusación en *Oyediran*, donde la prueba pericial de las sacudidas estaba respaldada por las lesiones adicionales¹¹⁵ y por la demostrada (mediante pruebas circunstanciales) predisposición violenta e intencional del acusado. Sólo en el caso de *Henderson* la Corte tuvo que lidiar directamente con las incertidumbres inherentes al conocimiento científico.

Lejos de soslayar el aspecto de la incertidumbre científica o de expresar frustración por la incapacidad de la ciencia para suministrar respuestas concluyentes a cada cuestión jurídica susceptible de ser probada científicamente, tal como la caricatura de la prueba pericial podría llevar a esperar, la Corte de Apelaciones en *Henderson* reconoció abiertamente las limitaciones epistémicas de la actual ciencia médica respecto al examen del trauma craneal en bebés e infantes:

Enfatizamos que dentro del ámbito del conocimiento médico actual, nos estamos enfrentando a causas naturales de muerte, porque en esta apelación se nos recuerda claramente que en ningún momento en un área como ésta se puede considerar que el conocimiento es completo y comprensivo. Hay límites para el alcance del conocimiento y no se puede obtener ninguna conclusión sin reconocer la posibilidad de que surja una causa desconocida a la luz de la percepción médica y que la mera exclusión de toda posible causa conocida no prueba que se haya infringido violencia en forma deliberada¹¹⁶.

La Corte entendió perfectamente bien que la investigación científica, en principio, no es concluyente y que el actual conocimiento científico siempre es, en un sentido importante, provisional y sujeto a revisión a la luz de experimentos posteriores, de grandes avances teóricos o de alguna innovación tecnológica. Aun cuando el testimonio del perito se aventuró cerca de la vanguardia de la ciencia experimental moderna, el papel del juez no es contravenir ni negar el hecho irremediable de la incertidumbre científica, sino manejar ese déficit epistémico de una manera institucionalmente aceptable.

La Corte de Apelaciones en *Henderson* fue invitada por la defensa a tener en cuenta «un espectacular desarrollo posterior al juicio oral»¹¹⁷, es decir, un caso posteriormente reportado en la literatura médica en el cual se descubrió que un componente de la triada —«pliegues perimaculares asociados a hemorragias retinianas abundantes», esto es, sangrado tras los ojos—¹¹⁸, había sido causado por una enfermedad y no por un trauma externo:

¹¹⁵ «[E]n *Oyediran* la acusación alegó una lesión diferente y separada, a saber: la fractura del brazo del bebé» [2010] 2 Cr. App R. 24 [7].

¹¹⁶ [2010] 2 Cr. App 24 [21].

¹¹⁷ *Ibid.* [39].

¹¹⁸ *Ibid.*

El descubrimiento de un caso de pliegues retinianos debido a leucemia y, por tanto, sin causa traumática alguna, enfatiza la importancia de reconocer los límites del conocimiento médico en un tiempo determinado y la necesidad de apreciar que aquello que no ha sido previamente contemplado puede, no obstante, ocurrir. Debemos reconocer los límites de la ciencia médica y en particular que puede haber eventos, muertes o síntomas que no tienen explicación y son imprevisibles¹¹⁹.

Sin embargo, la Corte no se detuvo simplemente en reconocer la posibilidad teórica de que uno de los síntomas del pequeño *Maeve* pudiera ser atribuible a una causa no traumática. Para satisfacerse a sí misma, de acuerdo con la teoría ortodoxa de la revisión en apelación en Inglaterra y Gales, según la cual el veredicto del jurado está sustentado en conclusiones inferenciales lógicas y racionales que pueden derivarse de las pruebas presentadas en el juicio oral, la Corte intentó determinar si en los hechos del caso hubo algún fundamento para atribuir el sangrado retiniano de *Maeve* a una leucemia o a alguna otra etiología no traumática compatible con la inocencia del acusado. Esta tarea exigió que la Corte escuchara entre los folios de las pruebas periciales aportadas al proceso y en la apelación, prestando cuidadosa atención particularmente a los términos exactos en los que cada experto había declarado y, también, expresando una preferencia en relación con determinadas conclusiones inferenciales cuidadosamente delineadas, expresando una preferencia por el testimonio de un perito sobre otro.

La siguiente cita resume la cadena de inferencias que respaldan la conclusión de la Corte de que el sangrado retiniano atribuible a una enfermedad o a otra causa no traumática, no era una explicación viable en el caso *Henderson*:

Aceptamos, entre otras razones porque los peritos estuvieron de acuerdo, que el mecanismo de la formación de pliegues retinianos no se conoce con certeza. Pero rechazamos la sugerencia del profesor Luther de que la existencia de pliegues retinianos no es una característica adicional de la triada. La triada brinda un fuerte apoyo, aunque no es concluyente, a las sacudidas o a las sacudidas y lesiones debidas a un impacto, cuando no hay pliegues retinianos. Dado que no se conoce cómo se forman los pliegues retinianos, no hay una base probatoria sólida para decir que se producen simplemente como resultado de una hemorragia. El testimonio del señor Elston, como médico que en veinticinco años de oftalmología pediátrica sólo ha visto pliegues retinianos en casos de sacudidas y de sacudidas más golpes [...], nos persuade de que los pliegues son rasgos que apuntan hacia una sacudida o a una sacudida con golpes, además de la aparición de la hemorragia. [...] En esas circunstancias, el testimonio del profesor Luther sólo equivale a una expresión de mayor duda y mayor precaución que las que suscribiría el señor Elston¹²⁰.

¹¹⁹ *Ibid.* [43]-[44].

¹²⁰ *Ibid.* [41], [43].

Nótese que la Corte no dice que Elston estaba en lo correcto y Luthert equivocado. El análisis es más sutil de lo que podría sugerir el paradigma adversarial en una disputa entre peritos instruidos por las partes. La conclusión de la Corte es que, aun en su mejor versión, el testimonio del profesor Luthert simplemente pone en tela de juicio los mecanismos que causan los pliegues retinianos: sin embargo, dichos pliegues no son parte de la triada, que incluso por sí sola es altamente probatoria de *NAHI*. Por otra parte, la amplia experiencia clínica del Dr. Elston sugiera que los pliegues retinianos siempre indican *NAHI*, conclusión que el testimonio del profesor Luthert atemperaba pero no podía refutar. El Dr. Elston había visto lo que había visto a lo largo de veinticinco años, aunque bien es cierto que la experiencia clínica no es lo mismo que los estudios médicos controlados adecuadamente u otros experimentos científicos.¹²¹ Así la Corte concluyó, de manera suficientemente razonable, que la presencia de pliegues retinianos no es prueba concluyente de las sacudidas pero tampoco socava integralmente su importancia diagnóstica. Los pliegues retinianos de Maewe podrían, en consecuencia, ser presentados legítimamente ante el jurado como prueba médica adicional capaz de respaldar la triada y, por extensión, también a la teoría de la acusación sobre las sacudidas delberadas.

Al final, expertos y jueces estaban de acuerdo en que subsistirían algunas incertidumbres. Nadie está realmente muy seguro de qué causa los pliegues retinianos. Pero el trabajo de la Corte de Apelaciones estaba hecho cuando se cercioró de que todas las fuentes sustanciales de incertidumbre científica habían sido descartadas por las pruebas presentadas en el juicio oral y adecuadamente explicadas al jurado en la recapitulación del juez. Si un juez penal estuviese obligado a eliminar cada duda imaginable y toda posibilidad teórica de inocencia, nunca nadie sería condenado y la justicia penal nunca llegaría a realizarse. Las hipótesis experimentales, como Geddes III, se convertirían en órdenes de libertad para que los maltratadores infantiles saliesen inmediatamente de prisión. El jurado está metódicamente instruido sobre el significado social de la tarea que está llamado a desempeñar en la administración de justicia penal¹²² y está privilegiadamente bien posicionado (de acuerdo con el diseño institucional) para decidir si queda alguna duda residual que impida una condena penal. Esta es la traducción institucional práctica de la prueba más allá de toda duda *razonable* como estándar jurídico normativo.¹²³ La Corte de Apelaciones se cercioró de que no había otras explicaciones posibles para los pliegues retinianos de Maewe a las que debieran dársele mayor

¹²¹ Una consideración que tiene que ver con el peso, más que con la admisibilidad, en el derecho inglés: *R. v. Adams* [2010] 1 Cr. App R. 8 [2009] EWCA Crim. 1876; *R. v. Thomas* [2011] EWCA Crim. 1295.

¹²² Véase JUNICRA, STRONGS BOASE, 2010.

¹²³ De hecho, a los jurados en Inglaterra y Gales se les instruye sobre que deben estar «seguros» de la culpabilidad antes de condenar: véase, p. ej., *R. v. Mijid* [2009] EWCA Crim. 1876; *R. v. Thomas* [2011] EWCA Crim. 1295.

relevancia durante el juicio o en la recapitulación hecha por el juez de primera instancia al jurado: «La triada en sí misma suministró una base sólida para la conclusión de que las sacudidas o bien las sacudidas más golpes, fueron la causa del desplome repentino de Maewe. Ni una infección ni una repetición, como sugirieron las pruebas no debatidas de la primera instancia y ante esa corte, explicaban las hemorragias retinianas, ni los pliegues, ni el hematoma subdural».¹²⁴ Dado que la recapitulación del juez fue impecable desde todo punto de vista y escapa a cualquier crítica desde la perspectiva jurídica, la competencia de la Corte de Apelaciones para revisar la seguridad de la condena de Henderson se agota. La incertidumbre científica había sido confrontada y manejada en el proceso exactamente como debió hacerse, desde el propio punto de vista institucional del derecho.

7. CONFLICTOS SOBRE LA AUTORIDAD DEL PERITO Y SU CREDIBILIDAD

Generalmente, se considera que la prueba pericial es particularmente problemática cuando los peritos instruidos por las partes están en desacuerdo.¹²⁵ ¿Cómo se supone que el juzgador de los hechos debe resolver con alguna base racional las disputas entre los peritos? Por definición, el juzgador de los hechos carece del conocimiento científico relevante o de otro conocimiento especializado o pericial. De manera que la preocupación es que, a falta de alguna base racional para distinguir entre el testimonio de peritos en contradicción, los jurados legos recurran a estrategias cognitivas irracionales, como preferir el testimonio del experto con mayor antigüedad profesional, o el que habla con mayor seguridad o elocuencia en el estrado, o incluso al perito que se considere mejor vestido o más atractivo físicamente.¹²⁶ Alternativamente, el jurado puede optar por «mediar la diferencia» existente entre las pruebas periciales contrapuestas (p. ej., ajustando el veredicto a homicidio imprudente, si la prueba pericial de la acusación indica homicidio doloso y el perito de la defensa presenta una teoría de accidente imprudente), o quizá simplemente ignorar por completo la prueba pericial y concentrarse exclusivamente en los aspectos no científicos del caso.

Éstas no son cuestiones triviales y algunas de ellas están indirectamente respaldadas por investigaciones empíricas (algo desactualizadas) sobre el uso de las ciencias forenses en el procedimiento penal inglés.¹²⁷ Por otro lado, no debería soslayarse el riesgo de que un jurado pudiese ser acriticamente deficiente y estar excesivamente influenciado en sus deliberaciones por un testimo-

¹²⁴ [2010] 2 Cr. App R. 24 [72].

¹²⁵ Véase *supra*, n. 113.

¹²⁶ SEHMAN, CHAMPAGNE y WITAKER, 1996: 371; TANTON, 1979: 681.

¹²⁷ ROBERTS y WILLMORE, 1993: cap. 5.

nio experto no cuestionado¹²⁸. Además, no es necesariamente irracional que el jurado eche mano de sucedáneos de fiabilidad probatoria cuando no está calificado para evaluar directamente los méritos sustantivos en una disputa científica. Es cuestión de elegir sucedáneos epistémicamente defendibles para las conclusiones inferenciales concretas de la prueba pericial para autorizar y legitimar esas conclusiones en el caso en cuestión. En ciertos escenarios, la relativa antigüedad y experiencia de los expertos en contradicción bien podría ser un sucedáneo legítimo de la fiabilidad de su testimonio, mientras que —por supuesto— su atractivo físico nunca podría serlo. Finalmente, vale la pena insistir en que la determinación de los hechos por el jurado tiene lugar en un contexto institucional en el que el riesgo de error en la decisión judicial —condenar erróneamente al inocente o absolver erróneamente al culpable— se distribuye entre las partes sobre una base considerablemente asimétrica¹²⁹. El acusado tiene derecho a beneficiarse de cualquier duda razonable, de acuerdo con la presunción de inocencia. En consecuencia, la prueba pericial de cargo debe hacer que el jurado esté seguro de la culpabilidad del acusado¹³⁰, mientras que los peritos de la defensa sólo necesitan hacer surgir una duda razonable para asegurar una absolución.

Las tres apelaciones en *Henderson* presentaban puntos de desacuerdo entre los peritos, pero es nuevamente el propio caso *Henderson* el ejemplo más ilustrador. La sección previa resumía la disputa sobre la causa de los pliegues retinianos entre el Dr. Elston de la acusación y el profesor Luthert de la defensa. Una vertiente adicional del análisis de la Corte tenía que ver con el alcance de la pericia de cada perito:

El testimonio de cualquier experto en un área particular está inevitablemente limitado al área en que profesa su pericia. Esta, por supuesto, es la razón por la cual ningún experto puede estar en la posición del jurado ni de esta Corte, pudiendo suministrar pruebas concretas en el contexto de la totalidad¹³¹.

El propio profesor Luthert estuvo de acuerdo en que la pregunta acerca de si la lesión traumática causaba daños a las fibras nerviosas estaba fuera de su área de especialidad médica y, en consecuencia, era un factor para cuya valoración él no tenía conocimientos, a fin de determinar si la presencia de los pliegues retinianos podía corroborar la triada.

Los límites legítimos de la pericia de un perito deben ser supervisados diligentemente por los jueces de primera instancia no sólo para asegurarse, como lo recuerda la Corte de Apelaciones, de que el jurado siempre debe tener la última palabra en la decisiva cuestión de determinar la culpabilidad o la inocencia. La tendencia de genuinos expertos a sobrepasar los límites de su

¹²⁸ Cf. VAN KAMPEN y NUBOER, 1997: 951; HOWARD, 1991: 98.

¹²⁹ ROBERTS y ZUCKERMAN, 2010: 247-252.

¹³⁰ Véase *supra*, n. 125.

¹³¹ [2010] 2 Cr. App R. 24 [47].

pericia legítima es probablemente un problema que predomina en el proceso penal inglés y es pasado por alto con tanta facilidad que se termina presentando como pruebas periciales la charlatanería de los fraudes científicos. Pero hay un mundo de diferencia entre el testimonio pericial y las meras opiniones personales de un perito¹³² y no se debe inducir al jurado al error de confundir las segundas con el primero.

La causa del daño al nervio que sufrió Maeve se convirtió en un aspecto importante en el proceso porque si éste era prueba de una lesión traumática independiente de la triada, el daño del nervio corroboraría fuertemente la afirmación de la acusación de las sacudidas deliberadas. En la apelación, la defensa presentó el testimonio experto de un distinguido neuropatólogo de Chicago, el Dr. Leestma, quien sugirió que el daño al nervio sufrido por Maeve no podía respaldar ninguna inferencia que llevará a *NAHI*. La Corte de Apelaciones no quedó nada convencida, en primer lugar porque el testimonio del Dr. Leestma era contradictorio por el peso de la literatura médica relevante y por el testimonio del experto de la acusación, Dr. Al-Sarraj:

[S]u testimonio fue fundamentalmente flojo. Su insistencia en afirmar que no era posible atribuir importancia al daño del axón [nervio], va en contra de las referencias tanto en los casos Geddes I, II y Reichard, como de la prueba que resulta de la experiencia actualizada del Dr. Al-Sarraj. Rechazamos este testimonio en la medida en que no logra darle importancia a lo que reveló la prueba del Beta amiloide¹³³.

En general, a la Corte de Apelaciones le preocupaba que el Dr. Leestma carecía de experiencia clínica reciente y que parecía no estar informado de las últimas publicaciones y desarrollos en la materia:

En nuestra opinión, la experiencia del Dr. Leestma es sobre todo histórica y mucho más limitada que la del Dr. Al-Sarraj. No ha hecho autopsias ni ha ofrecido su testimonio en casos relacionados con bebés sacudidos en muchos años¹³⁴.

De forma no sorprendente, la Corte de Apelaciones concluyó que el testimonio del Dr. Leestma no atacaba de forma alguna al argumento presentado por la acusación, es decir, que el daño al nervio podía corroborar la triada de síntomas de sacudidas que presentaba Maeve. De hecho, habiendo escuchado lo que él tenía por decir *de bene esse*¹³⁵, la Corte incluso dudó si el Dr. Leestma era un perito competente en los asuntos médicos sobre los que testificó:

¹³² Véase, por ejemplo, *Meadow v. General Medical Council* [2007] QB 462 [2006] EWCA Civ 1390.

¹³³ [2010] 2 Cr. App R. 24 [61].

¹³⁴ *Ibid.*

¹³⁵ Esto es, sin previa decisión sobre su admisibilidad. La Corte incluso concluyó que el testimonio del Dr. Leestma era inadmisiblemente en el procedimiento de apelación, pero que había estado bien escucharlo, debido a la importancia y dificultad de las cuestiones en juego.

El profesor Luthert había reconocido la importancia de los signos de lesión traumática en los axones. En nuestra opinión, nada en el testimonio del Dr. Leestma sugiere que la importancia otorgada a estos signos de lesión estuviera infundada. De hecho, la insistencia del Dr. Leestma en adelantar proposiciones que subsecuentemente tenía que retirar cuando conocía más sobre el caso, a la par de su falta de experiencia actualizada, dañó seriamente y minó el efecto de su testimonio. Habríamos tenido serias dudas acerca de si estaba adecuadamente calificado para rendir un testimonio con el objetivo de refutar el testimonio del Dr. Al-Sarraj o de la Dra. Squier en el proceso. No concluimos nada en ese sentido. Es suficiente concluir que su testimonio no pone en duda la seguridad del veredicto, en la medida en que se basa en los signos de lesión traumática en los axones del tracto corticoespinal.¹³⁶

En otros casos recientes hay indicaciones de que la Corte de Apelaciones a veces puede ser demasiado propensa a dar más crédito a la experiencia clínica o de otro tipo, que al conocimiento teórico, el cual puede llegar a ser marginado como «meramente teórico» y, en consecuencia inadaptable e irrelevante para la cuestión práctica de resolver debatibles cuestiones de hecho en las sentencias penales.¹³⁷ En *Henderson*, la Corte se atrevió a dictar la siguiente guía general:

El hecho de que un perito esté en la práctica clínica cuando presenta su informe, es importante. La práctica clínica da a los peritos la oportunidad de mantener y desarrollar su experiencia. Estos peritos adquieren una experiencia que continúa y se desarrolla. Su observación continua, su experiencia tanto en lo previsto como en lo imprevisto, lo reconocido y lo no reconocido, forma una poderosa base para su testimonio. Los clínicos aprenden de cada caso en el que se ven involucrados. Cada caso les hace pensar y, a medida que se desarrolla su experiencia también lo hace su comprensión. La experiencia continua les da la oportunidad de ajustar las opiniones previas, de modificar sus puntos de vista. Están mejor situados para reconocer que lo que era desconocido un día, puede ser reconocido al día siguiente. Esa experiencia clínica [...] puede suministrar una fuente mucho más fiable de prueba que aquella aportada por quienes han dejado de practicar sus conocimientos en un ambiente clínico permanente y se han retirado de dicha práctica. Tales expertos están, en general, ocupados sólo en la revisión de las opiniones de otros. Han perdido la oportunidad diaria de aprender y desarrollarse a partir de la experiencia continua.¹³⁸

Es un comentario razonable hasta cierto grado. Al mismo tiempo, es fácil olvidar que la experiencia clínica típicamente *no* es científica desde un punto de vista riguroso, sea experimental o metodológico; y que las apelaciones que hacen caso omiso de las prácticas, culturas o métodos clínicos establecidos (¿un tipo de sucedáneo médico?) no constituyen validación científica. Sin embargo estas consideraciones no parecen haber afectado el análisis de la Corte y

¿FUE EL BEBÉ SACUDIDO? PRUEBA, PERICIA Y EPISTEMOLOGÍA JURÍDICA...

sus conclusiones en *Henderson*. La eminencia del Dr. Leestma no fue, ciertamente, tratada como un sucedáneo de la fiabilidad de su testimonio. Tampoco la Corte de Apelaciones se vio influenciada por los modales encantadores del Dr. Leestma en la sala de audiencias, que dejaron «una impresión más seductora, cortés y discreta tal como fue».¹³⁹ Adicionalmente, su falta de experiencia clínica reciente y relevante y su ignorancia sobre las recientes publicaciones médicas no fueron meros pretextos formales para restarle valor probatorio al testimonio del Dr. Leestma, sino vacíos demostrables en la sustancia de su testimonio, acentuados por las concesiones que fue obligado a hacer en el interrogatorio cruzado. Considerando todos estos factores prominentes, la decisión de la Corte de Apelaciones, desestimar el testimonio del Dr. Leestma, parece eminentemente racional y facilitó el camino a la conclusión, respaldada epistémicamente, de que la condena de Henderson era segura:

Las consecuencias de nuestra conclusión con respecto al testimonio del Dr. Leestma consisten en que hay al menos una característica de los síntomas de Maewe, adicional a aquellas que conforman la trada. Aunque hubiésemos aceptado el testimonio del profesor Luthert acerca de que los pliegues perimucosulares pueden ser atribuibles a la hemorragia y que fueron, por tanto, sólo un aspecto de los síntomas oftalmológicos de la trada, la evidencia de que hubo daño traumático en los axones suministra una poderosa prueba adicional del trauma. Conforma, con los demás síntomas no refutados, un sólido fundamento para el veredicto de culpabilidad.¹⁴⁰

Una ilustración final de la valoración judicial de la credibilidad del perito es suministrada por la apelación del caso *Oyediran*. Uno de los peritos que asistía a la defensa, la Dra. Squier, había especulado que las lesiones fatales de Ferni podrían haber sido causadas o estar materialmente afectadas por un VIH no diagnosticado. No había ningún diagnóstico de VIH en la familia: las especulaciones de la Dra. Squier estaban motivadas exclusivamente por el origen africano de Ferni. Aun así, a partir de la recomendación de la Dra. Squier, la defensa consultó al profesor Bell, un reconocido experto en la materia, para eliminar la sospecha de la Dra. Squier. La respuesta llevó a que no había absolutamente ninguna base para atribuir las lesiones de Ferni a un VIH no diagnosticado. Sin embargo, la Dra. Squier aparentemente mantuvo su versión y rechazó la posibilidad de descartar un VIH no diagnosticado como factor causal que contribuyó a las lesiones de Ferni. Incluso la Dra. Foster insistió —a pesar de existir prueba contraria del médico de cabecera de la familia— que la madre de Ferni, Sophia Rudder, podría padecer de un VIH no diagnosticado. La Corte de Apelaciones quedó perpleja ante esta aparente intransigencia:

[N]o hay prueba alguna de que el estado de Ferni haya tenido algo que ver con el VIH. No hay base probatoria para sugerir algo semejante. Por tanto, es

¹³⁶ [2010] 2 Cr. App. R. 24 [63].

¹³⁷ *R. v. Reed and Reed*; *R. v. Gammson* [2010] 1 Cr. App. R. 23 [2009] EWCA Crim. 2698 [101]-[110]; *R. v. Weller* [2010] EWCA Crim. 1085 [24]-[25], [38].

¹³⁸ [2010] 2 Cr. App. R. 24 [2008].

¹³⁹ *Ibid.* [218].

¹⁴⁰ *Ibid.* [64].

lamentable y sorprendente que debamos registrar el hecho de que, pese a la ausencia de toda prueba de encefalitis por VIH y pese al testimonio del profesor Bell, quien contaba con los elogios de la Dra. Squier que están registrados, ésta no estuviese preparada, ante nosotros, para rechazar la posibilidad de encefalitis por VIH. Repitió que el VIH no había sido descartado e hizo referencia a que el padre de Femi proviniese de un país endémico. La Dra. Squier no debió haber insistido en esa sugerencia [...]. Lo inapropiado de su testimonio tiene el agravante de su insistencia en sugerir que Sophia Rudder podría sufrir de VIH y no de esclerosis múltiple [...]. La Dra. Squier [...] creyó conveniente sugerir el VIH, pese a la reputación de la Dra. Foster y en aparente ignorancia del hecho real de que Sophia Rudder estaba siendo tratada por aquella. El enfoque dado por la Dra. Squier a dicho aspecto de la prueba respalda nuestra conclusión sobre la no fiabilidad de su testimonio. No debió haber sugerido esa causa de la condición y muerte de Sophia Rudder sin tomar cuidadosa consideración del tratamiento de la Dra. Foster y sin informarse acerca de las cualificaciones y reputación de la Dra. Foster.¹⁴¹

La Corte extrajo conclusiones inevitables:

La posición de la Dra. Squier, en su declaración oral ante nosotros, plantea serias dudas sobre la fiabilidad del resto de su testimonio y su enfoque sobre este caso. Demuestra, en nuestra opinión, que estaba preparada para sostener una teoría no confirmada e insostenible, en un intento de respaldar esta apelación [...]. A la luz de nuestro punto de vista sobre la calidad del testimonio de la Dra. Squier ante nosotros, concluimos que no logra minar la seguridad del veredicto.¹⁴²

El comportamiento de la Dra. Squier en *Henderson* (según lo transmitió la Corte de Apelaciones), resulta excéntrico, cuanto menos. ¿Por qué un perito habría de recomendar consultar a un conocido especialista en VIH pediátrico y luego negarse a aceptar sus conclusiones? ¿Por qué un perito persistiría en una teoría totalmente especulativa sobre la enfermedad de la madre de Femi cuando el médico de cabecera de Sophia Rudder estaba en mejor posición para dar un diagnóstico clínico informado? Al menos, como lo dice la Corte de Apelaciones, se tiene la clara impresión de un perito obsesionado con una idea o posiblemente un médico forense práctico que, presa de una equivocada concepción de la ética adversarial, rechaza por principio hacer cualquier concesión a la acusación. De manera alternativa, la Corte de Apelaciones podría no haberle entendido correctamente. La Dra. Squier pudo simplemente haber reafirmado la ineliminable incertidumbre de toda proposición científica, en el sentido de que la falsabilidad nunca descarta totalmente una hipótesis. Si es así, *Oyediran* demuestra que argumentar perogrulladas epistemológicas y pretender tomar cierta ventaja de un escepticismo global, probablemente no son estrategias forenses efectivas para los peritos. Por el

¹⁴¹ *Ibid.*, [188], [189].

¹⁴² *Ibid.*, [188], [190].

contrario, tener como contrincante a la Corte en temas de filosofía de la ciencia puede dar rial resultado y potencialmente socavar toda la prueba pericial (que podría ser perfectamente sólida en otros aspectos) y dañar severamente la reputación del perito, afectando su implicación en procesos futuros. En todo caso, puede entenderse fácilmente por qué la Corte de Apelación rechazó el testimonio de la Dra. Squier en *Oyediran*. Es un ejemplo en el que la credibilidad del testimonio experto estaba irremediablemente comprometida y podía ser rechazado con fundamentos completamente racionales y epistémicamente justificables.

8. CONCLUSIÓN: HACIENDO JUSTICIA A LA PRUEBA PERICIAL Y A LA PRUEBA CIENTÍFICA

Este trabajo ha presentado un pomenorizado caso de estudio de tres apelaciones conjuntas en las que pruebas periciales opuestas jugaron un papel central. En cierta forma, los casos de bebés sacudidos son altamente atípicos en el funcionamiento general de las acusaciones penales en Inglaterra y Gales; en otros aspectos, presentan problemas probatorios similares a los que se presentan con frecuencia en los juicios penales. Sin embargo, el valor del ejercicio no consiste en una pretensión formal de representatividad estadística. Por el contrario, he buscado explorar el modo en que en las apelaciones en *Henderson* se proyectan los mecanismos institucionales a través de los cuales la prueba pericial y la prueba científica se integran en los procedimientos penales: un sistema completo de decisión judicial en materia penal reflejado en un caso individual.

Habiendo repasado los hechos principales del caso de estudio en la sección 1, las secciones 2 a 4 destacan las características distintivas del contexto institucional y normativo de la jurisdicción penal en Inglaterra y Gales. Para esos arreglos procedimentales es fundamental la división institucional de trabajo entre el juez y el jurado en primera instancia y entre la determinación de los hechos por el jurado y la revisión del caso en apelación. Examinar las dinámicas epistemológicas de la prueba judicial, en mi opinión, no es simplemente una cuestión de entender lo que es característicamente «jurídico» en la determinación forense de los hechos, sino que requiere también una atención más próxima y receptiva a las tradiciones procesales y a las culturas locales. Ésta es una tarea tanto para los estudios del derecho comparado como un ejercicio de epistemología jurídica, además, conforma un entorno esencial para la discusión de aspectos familiares sobre la incertidumbre científica y el desacuerdo entre expertos, revisados en los apartados 5 y 6.

A menudo se afirma, explícita o implícitamente, que el tratamiento jurídico de la prueba pericial es fundamentalmente irracional y que carece de

respaldo epistemico. Esta es una crítica seria—y seriamente problemática—que este trabajo no pretende responder de forma exhaustiva. Por el contrario, busca equilibrar las que a veces parecen críticas exageradas y equivocadas mediante la descripción de algunos aspectos de los análisis y las conclusiones de la Corte de Apelaciones en *Henderson* que resultan—al menos, para mí—perfectamente racionales y epistémicamente justificables. La prueba científica nunca es una prueba suficiente para determinar la culpabilidad (o la inocencia)¹⁴³ y siempre debe ser evaluada contextualmente, en relación con el resto de pruebas del caso. Tal como lo subrayó la Corte de Apelaciones en *Henderson*, «la prueba pericial se utilizó para probar que el acusado está mintiendo en lo que ha dicho, bien sea inmediatamente o bien a lo largo del proceso»¹⁴⁴. La Corte llegó a ejemplificar enfoques judiciales prácticos de la incertidumbre científica y de la evaluación racional de la credibilidad de los peritos.

El punto de vista de este trabajo sobre la práctica judicial inglesa ha sido predominantemente descriptivo y crítico. Si bien *Henderson* es un caso importante, no quiero parecer complaciente con los antecedentes del derecho inglés respecto a su forma de abordar la prueba científica y otras pruebas periciales (que están innegablemente marcados por errores judiciales demostrados), ni quiero sugerir que el actual sistema no requiere reformas deseables. La filosofía del manejo judicial proactivo de los casos se ha arraigado en el proceso penal inglés a lo largo de la última década¹⁴⁵ y las ventajas de una preparación previa de dichos casos antes del proceso, más rigurosa y colaborativa, podrían considerarse especialmente evidentes en los casos de pruebas periciales en controversia. Al reiterar el mensaje de numerosos casos recientes, la Corte de Apelaciones en *Henderson* dijo:

¿[C]ómo debe abordar un jurado las pruebas periciales en contradicción? Sugierimos que sólo puede hacerlo si las pruebas están adecuadamente preparadas y controladas antes de presentarse ante el jurado. A menos que las pruebas estén adecuadamente preparadas antes de que el jurado tome posesión, es poco probable que se puedan dar instrucciones adecuadas sobre cómo el jurado debe abordarias. Así el jurado tendrá obstáculos para considerar esas pruebas de manera tal que le permita alcanzar una conclusión lógicamente justificable [...]. La justicia en esos casos depende de una adecuada preparación anticipada y del control de las pruebas, desde el inicio de la etapa de la investigación en adelante [...]. Es esencial una adecuada y sólida gestión previa al juicio oral. Sin ella, los aspectos médicos auténticos no pueden ser identificados. A falta de esa identificación, es poco probable que un juez sea capaz de advertir a los expertos que no divaguen en detalles innecesarios, complicados y confusos. A menos que los aspectos médicos auténticos sean identificados con antelación, los detalles evitables no serán evitados¹⁴⁶.

¿FUE EL BEBÉ SACUDIDO? PRUEBA, PERICIA Y EPISTEMOLOGÍA JURÍDICA...

Otro objetivo para una reforma podría encontrarse en las reformadas reglas de admisibilidad [los trabajos recientes de la *Law Commission* han generado el proyecto de la *Criminal Evidence (Experts) Bill 2011*]¹⁴⁷, para el desarrollo de nuevos códigos de ética y de responsabilidad profesional que regulen la práctica de los testimonios periciales¹⁴⁸, mejoras en las estructuras regulatorias y de acreditación¹⁴⁹ y continuos programas de formación y entrenamiento en áreas de conocida dificultad, tales como la prueba estadística y el razonamiento probabilístico¹⁵⁰. Sin embargo, no está de más insistir en que los procedimientos reformados deben ser compatibles con los preciados valores procesales y con las tradiciones, dado que éstos constituyen el suelo fértil en el que se arraiga (o no, según sea el caso) la legitimidad de los veredictos penales. El poder de la ciencia debe ser aprovechado por la administración de justicia, sin erosionar los fundamentos epistémicos y normativos del derecho. *Henderson* demuestra estrategias prácticas mediante las cuales esta síntesis algo elusiva podría ser alcanzada en un sistema procesal penal adversarial en el *common law*.

BIBLIOGRAFÍA

- ALLEN, R. J., 1994: «Factual Ambiguity and a Theory of Evidence», *Northwestern University Law Review*, 88.
- ANDERSON, T., 1999: «On Generalizations I: A Preliminary Exploration», *South Texas Law Review*, 40.
- ANDERSON, T.; SCHUM, D., y TWINING, W., 2005: *Analysis of Evidence*, 2.ª ed., Cambridge: Cambridge University Press.
- ANÓNIMO, 2010: Note «Admitting Doubt: A New Standard for Scientific Evidence», *Harvard Law Review*, 123.
- ATKIN, C.; ROBERTS, P., y JACKSON, G., 2010: *Fundamentals of Probability and Statistical Evidence in Criminal Proceedings: Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses*, London: Royal Statistical Society. Disponible en internet: www.rss.org.uk/sites/cms/content/view/article.asp?article=1132.
- ATTORNEY GENERAL'S OFFICES, 2004: «Infant Death Cases: Attorney-General's Review», *HL DEB* vol. 667, cols. 1567-66, 21 December 2004. *Review of Infant Death Cases*. BEECHER-MONAS, E., 2000: «The Heuristics of Intellectual Due Process: A Primer for Thiers of Sciences», *New York University Law Review*, 75.
- BLUME, J. H.; JOHNSON, S. L., y PAVOLA, 2007: «Every Juror Wants a Story: Narrative Relevance, Third Party Guilt and the Right to Present a Defense», *American Criminal Law Review*, 44.

¹⁴⁷ LAW COMMISSION, 2011: HC 829; véase ROBERTS, A., 2009: 551; ROBERTS, P., 2009.

¹⁴⁸ Cf. *Criminal Procedure Rules*, 2011, reglas 33.2 y 33.3. De manera general, véase SANDERS, 2007: 1559.

¹⁴⁹ Ahora en Inglaterra y Gales la responsabilidad es del Forensic Science Regulator y del Forensic Advisory Council; véase www.homeoffice.gov.uk/agencies-public-bodies/fsr. Para una discusión general, véase BROWN y WILLIS, 2009: 192.

¹⁵⁰ ATKIN, ROBERTS y JACKSON, 2010.

¹⁴³ Cf. EDMOND, 2002: 53.

¹⁴⁴ [2010] 2 Cr. App R. 24 [2].

¹⁴⁵ DENVER, 2008.

¹⁴⁶ [2010] 2 Cr. App R. 24 [203]. [2011]. [205].

- BROWN, S., y WILLIS, S., 2009: «Complexity in Forensic Sciences», *Forensic Science Policy and Management*, 1.
- COLE, S. A., 2004: «Grandfathering Evidence: Fingerprint Admissibility Rulings from *Jennings to Llera Plaza* and Back Again», *American Criminal Law Review*, 41.
- DAMAŠKA, M. R., 1997: *Evidence Law Adrift*, New Haven: Yale University Press.
- DENYER, R. (HHJ QC), 2008: *Case Management in the Crown Court*, London: Hart.
- DORAN, S.; JACKSON, J. D., y SEIGEL, M. L., 1995: «Rethinking Adversariness in Non-jury Criminal Trials», *American Journal of Criminal Law*, 23.
- DUHIG, C., 2003: «The Need for an Independent Institute: A View from One Small "Ology"», trabajo presentado al congreso del Solon Institute sobre *Independencia y objetividad en las ciencias forenses: Problemas y soluciones*, Cambridge, 9-12 de julio de 2003.
- EASTON, S. D., 2006: «That is Not all There Is: Enhancing *Daubert* Exclusion by Applying "Ordinary" Witness Principle to Experts», *Nebraska Law Review*, 84.
- EDMOND, G., 2002: «Constructing Miscarriages of Justice: Misunderstanding Evidence in High Profile Criminal Appeals», *Oxford Journal of Legal Studies*, 22.
- FERGUSON, P. R., 2006: «The Criminal Jury in England and Scotland: The Confidentiality Principle and the Investigation of Impropriety», *International Journal of Evidence and Proof*, 10.
- FRIEDMAN, R. D., 1998: «Anchors and Flotsam: Is Evidence Law "Adrift"?», *Yale Law Journal*, 107.
- GLENN, H. P., 2010: *Legal Traditions of the World*, 4.ª ed., Oxford: Oxford University Press.
- GOLDMAN, A. I., 1999: *Knowledge in A Social World*, Oxford: Oxford University Press.
- 2002: *Pathways to Knowledge: Private and Public*, Oxford: Oxford University Press.
- HACK, S., 2009: «Irreconcilable Differences? The Troubled Marriage of Science and Law», *Law and Contemporary Problems*, 72.
- HARRIS, G. C., 2000: «Testimony for Sale: The Law and Ethics of Snitches and Experts», *Pepperdine Law Review*, 29.
- HO, H. L., 2008: *A Philosophy of Evidence Law. Justice in the Search of Truth*, Oxford: Oxford University Press.
- HOWARD, M. N., 1991: «The Neutral Expert: A Plausible Threat to Justice», *Criminal Law Review*.
- JACKSON, J. D., 2002: «Making Juries Accountable», *American Journal of Comparative Law*, 60.
- JASANOFF, S. (ed.), 1997: *Science at the Bar: Law Science and Technology in America*, Cambridge: Harvard University Press.
- 2006: *States of Knowledge: The Co-production of Science and the Social Order*, New York: Routledge.
- JONES, C. A. G., 1994: *Expert Witnesses: Science, Medicine and the Practice of Law*, Oxford: Oxford University Press.
- JUDICIAL STUDIES BOARD, 2010: *Crown Bench Book. Directing the Jury*. Disponible en internet en <http://www.judiciary.gov.uk/publications-and-reports/judicial-college/crown-court-bench-book-directing-the-jury/>.
- KEE, R., 1986: *Trial & Error*, London: Hamish Hamilton.

- KIRBY, T., 2003: «Devoted mother's agonising wait to win freedom. After Three years, two appeals and a huge legal bill, prosecution file no. 89 vindicates solicitor jailed for killing her babies», *The Independent*, 30 de enero.
- LAUDAN, L., 2006: *Truth Error and Criminal Law: An Essay in Legal Epistemology*, Cambridge: Cambridge University Press.
- LAW COMMISSION, 2011: *Expert Evidence in Criminal Proceedings in England and Wales* (núm. 235, HC), London: TSO.
- LLOYD-BOSTOCK, S., y THOMAS, Ch., 2000: «The Continuing Decline of the English Jury», en VIDMAR, N. (ed.), *World Jury Systems*, Oxford: Oxford University Press.
- MNOOKIN, J., 2008: «Experts and Forensic Evidence», *Southwestern University Law Review*, 37.
- NANCY PENNINGTON, N., y HASTIE, R., 1993: «The Story Model for Juror Decision Making», en HASTIE, R. (ed.), *Inside the Juror: the Psychology of Juror Decision-Making*, Cambridge: Cambridge University Press.
- NELKEN, D. (ed.), 2000: *Contrasting Criminal Justice: Getting from Here to There*, London: Ashgate.
- NOBLES, R., y SCHIFF, D., 2002: «The Right to Appeal and Workable Systems of Justice», *Modern Law Review*, 65.
- ÖRÜCÜ, E., y NELKEN, D. (eds.), 2007: *Comparative Law: A Handbook*, London: Hart.
- PARDO, M. S., 2000: «Juridical Proof, Evidence and Pragmatic Meaning: Toward Evidentiary Holism», *Northwestern University Law Review*, 95.
- POSNER, R. A., 1999: «An Economic Approach to the Law of Evidence», *Stanford Law Review*, 51.
- QUIRK, H., 2007: «Identifying Miscarriages of Justice: Why Innocence in the UK is Not the Answer», *Modern Law Review*, 70.
- REDMAYNE, M., 1997: «Expert Evidence and Scientific Disagreement», *UC Davis Law Review*, 30.
- RISINGER, M.; DENBEAUX, M. P., y SAKS, M. J., 1989: «Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lessons of Handwriting Identification "Expertise"», *University of Pennsylvania Law Review*, 137.
- ROACH, K., 2009: «Forensic Science and Miscarriages of Justice: Some Lessons from Comparative Experience», *Jurimetrics*, 50.
- ROBERTS, A., 2009: «Rejecting General Acceptance, Confounding the Gate-keeper: the Law Commission and Expert Evidence», *Criminal Law Review*.
- ROBERTS, P., 2006: «Theorising Procedural Traditions: Subjects, Objects and Values in Criminal Adjudication», en DUFF, A.; FARMER, L.; MARSHALL, S., y TADROS, V. (eds.), *The Trial on Trial Volume Two: Judgement and Calling to Account*, London: Hart.
- 2008a: «Comparative Criminal Justice Goes Global», *Oxford Journal of Legal Studies*, 28.
- 2008b: «Faces of Justice Adrift? Damaska's Comparative Method and the Future of Common Law Evidence», en JACKSON, J.; LANGER, M., y TILLERS, P. (eds.), 2008: *Crime, Procedure and Evidence in a Comparative and International Context*, London: Hart.
- 2009: «The Science of Proof: Forensic Science Evidence in English Criminal Trials», en FRASER, J., y WILLIAMS, R. (eds.), *Handbook of Forensic Science*, Cullompton: Willan.

- 2010: «On Method: The Ascent of Comparative Criminal Justice», London: Sage.
- 2011: «Does Article 6 of the European Convention on Human Rights Require Reasoned Verdicts in Criminal Trials?», *Human Rights Law Review*, 11.
- 2012: «Renegotiating Forensic Cultures: Between Law, Science and Criminal Justice», *Studies in History and Philosophy of Science*, 43, en prensa.
- ROBERTS, P., y WILLMORE, Ch., 1993: *The Role of Forensic Science Evidence in Criminal Proceedings*. RCCJ Research Study No. 11, HMSO.
- ROBERTS, P., y ZUCKERMAN, A., 2010: *Criminal Evidence*, 2.ª ed., Oxford: Oxford University Press.
- ROZEMBERG, J., 2003: «Pathologist misled appeal judges in Sally Clark case», *Daily Telegraph*, 12 de abril.
- SANDERS, J., 2007: «Expert Witness Ethics», *Fordham Law Review*, 76.
- SHUMAN, D. W.; CHAMPAGNE, A., y WITAKER, E., 1996: «Juror Assessments of the Believability of Expert Witnesses: A Literature Review», *Jurimetrics Journal*, 36.
- SMITH, J. C., 1998: «Is Ignorance Bliss? Could Jury Trial Survive Investigation?», *Medicine, Science and the Law*, 38.
- STEIN, A., 2005: *Foundations of Evidence Law*. Oxford: Oxford University Press.
- TANTON, R. L., 1979: «Jury Preconceptions and their Effect on Expert Scientific Testimony», *Journal of Forensic Sciences*, 24.
- THOMPSON, W. C., 1993: «Evaluating the Admissibility of New Genetic Identification Tests: Lessons From the "DNA War"», *Journal of Criminal Law and Criminology*, 84.
- THORNTON, P., 2004: «Trial by Jury: 50 Years of Change», *Criminal Law Review*.
- TWINING, W., 1982: «Taking Facts Seriously», en GOLD, N. (ed.), *Essays on Legal Education*. Toronto: Butterworths, reimpresso como *Rethinking Evidence*, 2.ª ed., Cambridge: Cambridge University Press, 2006.
- 2007: «Taking Facts Seriously - Again», en ROBERTS, P., y REDMAYNE, M. (eds.), *Innovations in Evidence and Proof*. London: Hart.
- VAN KAMPEN, P., y NUBOER, H., 1997: «Daubert in the Lowlands», *UC Davis Law Review*, 30.
- VOGGER, R., 2005: *A World View of Criminal Justice*. London: Ashgate.
- VOGLER, R., y HUBER, B. (eds.), 2008: *Criminal Procedure in Europe*. Berlin: Max Planck Institut/Duncker and Humblot.
- WALKER, C., y STOCKDALE, R., 1999: «Forensic Evidence», en WALKER, C., y STAR-MER, K. (eds.), *Miscarriages of Justice: A Review of Justice in Error*. Oxford: Blackstone.
- WARD, T., 2004: «Experts, Juries and With-hunts: From Fitzjames Stephen to Angela Cannings», *Journal of Law and Society*, 31.
- WISTRICH, A. J.; GÜTHRIE, Ch., y RACHUNSKI, J. J., 2005: «Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding», *University of Pennsylvania Law Review*, 153.
- WONDER, A. K. Y., 1989: «Science and Law, A Marriage of Opposites», *Journal of the Forensic Science Society*, 29.
- ZUCKERMAN, A. S., 1986: «Law, Fact or Justice?», *Boston University Law Review*, 66.

PRUEBA CIENTÍFICA. UN MAPA DE RETOS

VII

Marina GASCON ABELLÁN

Universidad de Castilla-La Mancha.

1. PRUEBA CIENTÍFICA. SOBREVVALORACIONES Y PARADIGMAS

1.1. El mito

En los últimos años los constantes avances científicos y técnicos han tenido un profundo impacto en el ámbito de la prueba y juegan un papel cada vez más importante en todos los procesos. Los avances han sido particularmente espectaculares en el campo de la genética forense, que ha marcado un antes y un después en la resolución de numerosos problemas judiciales, como la investigación biológica de la paternidad, la identificación de personas o la investigación de indicios, es decir, el análisis de muestras biológicas de interés criminal, como manchas de sangre, saliva, espermatozoides¹. El potencial de la huella genética es de tal magnitud que su uso en los tribunales se ha convertido ya en moneda corriente. Pero la prueba de ADN no es la única protagonista de este *boom*. Hay también otras pruebas científicas que juegan un papel decisivo en muchas causas judiciales, como la dactiloscopia, la balística, el análisis de marcas o la prueba de locutores de voz. Esta última, por ejemplo, ha permitido fundar condenas en causas de terrorismo donde la prueba principal frente al

¹ Sobre éstas y otras aplicaciones de la prueba de polimorfismos ADN a la genética forense, véase MARTÍNEZ JARRERA, 1999: 129 y ss.

imputado es una llamada de teléfono avisando de la colocación de un artefacto explosivo. Hoy puede decirse, en suma, que las pruebas científicas se han convertido en la clave para probar muchos hechos que de otro modo resultarían difíciles de probar.

Pero la importancia de las pruebas científicas en la práctica procesal no ha ido acompañada de un proceso paralelo de cautelas y controles en relación con las mismas. Más bien ha sucedido lo contrario. Precisamente por el hecho de presentarse como «científicas» (y porque la mayoría de las veces —al menos en Europa— provienen de los laboratorios oficiales de la policía científica) estas pruebas han ido acompañadas de un aura de infalibilidad que ha frenado (cuando no claramente impedido) cualquier intento de revisión o reflexión crítica sobre las mismas, con el resultado de que su validez y valor probatorio se suelen asumir como dogmas de fe². Esta creencia ciega en la validez y valor de las pruebas científicas es algo que llama particularmente la atención si se considera que nada es menos «científico» que asumir como válido un conocimiento sin un previo control de sus postulados ajustado a una metodología científica. Pero sobre todo entraña un peligro, pues propicia que las decisiones probatorias apoyadas en este tipo de pruebas se asuman como incuestionables o irrefutables y, de paso, descarga al juez de hacer un especial esfuerzo por fundar racionalmente la decisión: basta con alegar que hubo prueba científica y que ésta apuntaba justamente en la dirección de la decisión probatoria final.

La beatificación de este universo probatorio se asienta en dos sobrevaloraciones. Por un lado sus resultados se aceptan como infalibles. Por otro, y sobre todo, se considera que esos resultados dicen cosas distintas de las que en realidad dicen³. La primera es una *sobrevaloración epistémica*. La segunda es, por así decirlo, una *sobrevaloración semántica*⁴.

² Por lo demás, esa desbordante confianza en la infalibilidad de la prueba científica se ha visto alimentada por el tremendo impacto que en el imaginario popular han tenido algunos *booms* televisivos como la norteamericana *CSI (Crime Scene Investigation)*, que han generado una especie de beatificación de estas pruebas. De modo que hoy estamos gozosamente dispuestos a creer en los informes provenientes de los laboratorios de la policía científica como si se tratase de una verdad revelada (advierte contra esta actitud IGARÚA, 2007).

³ La sobrevaloración de la prueba científica ha sido denunciada en el National Research Council Report, 2009, *Strengthening Forensic Science in the United States. A Path Forward*, donde se lee expresamente: «Durante décadas las disciplinas de la ciencia forense han producido pruebas valiosas que han contribuido a la imputación y condena de delinquentes y a la exoneración de personas inocentes [...] Sin embargo estos avances también han revelado que, en algunos casos, información sustantiva y testimonios basados en análisis defensores de la ciencia forense pueden haber contribuido a la condena errónea de personas inocentes. Este hecho ha demostrado el potencial peligro de atribuir un peso indebido a la información y testimonios derivados de pruebas y análisis defensores. Es más, los informes periciales imprecisos o exagerados han contribuido algunas veces a la admisión de pruebas erróneas o engañosas» (Summary, Introduction, Paragraph: *Challenges Facing the Forensic Science Community*, P. 4).

⁴ Sobre estas dos sobrevaloraciones, véase más extensamente GASCÓN, 2010: 81 y ss.

1.2. Sobrevaloración epistémica

Pero cabe preguntar, ¿por qué se asume con este fervor casi dogmático el mito de la infalibilidad de las pruebas científicas? ¿Por qué todos los esfuerzos que en los últimos tiempos se están realizando por introducir racionalidad en el ámbito de la prueba se dirigen fundamental y casi exclusivamente a las pruebas no científicas? ¿Por qué la prueba científica no se sitúa también bajo esa mirada crítica?

No voy a desarrollar aquí una respuesta detallada a estas cuestiones pero sí puede señalarse someramente su causa principal. No se ha reflexionado por sí que se asume que una prueba científica transcurre por derroteros epistemológicos distintos al resto de las pruebas; que la arquitectura inferencial de ambos tipos de prueba es distinta; que mientras una prueba no científica se estructura a través de un razonamiento *inductivo*, la prueba científica se articula mediante un razonamiento *deductivo*. Lo que se sostiene, en otras palabras, es que la prueba *no científica* está basada en leyes probabilísticas de débil fundamento epistémico, por lo común máximas de experiencia y leyes del actuar humano habitual, que además son aplicadas dentro de una metodología no científica, de modo que sus resultados son fallibles y han de medirse siempre en términos de simple probabilidad. En cambio la prueba *científica* —se argumenta— está basada en leyes universales (o en todo caso en leyes probabilísticas que gozan de un fuerte fundamento epistémico), leyes que además son aplicadas dentro de una rigurosa metodología científica, por lo que sus resultados pueden tenerse por incuestionables o fuera de toda duda. Se piensa, en definitiva, que el conocimiento que se obtiene en las salas de los tribunales es frágil, en cambio lo que sucede en los laboratorios de la policía científica es otra cosa.

Naturalmente esta convicción es *errónea*, pues las pruebas científicas no constituyen por lo general un razonamiento de tipo deductivo, sino que están basadas prevalentemente en leyes estadísticas y sus resultados han de ser aún interpretados a la luz de otros datos, y por consiguiente difícilmente puede hablarse de «objetividad»⁵ y mucho menos de infalibilidad en relación con las conclusiones obtenidas a raíz de las mismas. Por supuesto es muy probable que el mito de la infalibilidad de las pruebas científicas tenga mucho que ver

⁵ «Cualquier juicio sobre probabilidad en un caso particular, incluso aunque el juicio esté basado en una frecuencia relativa, tiene una componente basada en conocimiento personal». Esto es lo mismo que decir que cualquier juicio sobre probabilidad es esencialmente personal y, por tanto, subjetivo. Los que entienden los resultados estadísticos como resultados objetivos en el sentido de interpretarlos como incontrovertibles y universalmente alcanzables tienen una concepción irreal de lo que la ciencia puede llegar a hacer en la práctica. Puede hablarse de objetividad entendida como acuerdo intersubjetivo. En este sentido es más fácil que los científicos acepten unos resultados si están basados en frecuencias relativas que si lo están en valoraciones subjetivas de probabilidad, pero ese acuerdo no implica, en ningún caso, que los científicos crean que los resultados son incontrovertibles. Véase TARONI, AITKEN, GARBOLINO y BIEDERMANN, 2006: 21.

con la prueba del ADN, cuya metodología está ya tan perfeccionada y el grado de probabilidad que arroja es tan alto que en la práctica puede «actuarse como si» fuese infalible. No obstante, aún dando esto por sentado, conviene adoptar una actitud un poco más crítica, pues la calidad epistémica —y por tanto el valor probatorio— de los resultados de una prueba científica depende de varios factores que es necesario ponderar.⁶

En primer lugar depende de la *validez científica y/o metodológica* de la misma. Muchas de estas pruebas pueden realizarse por métodos científicos diferentes y no todos ellos gozan del mismo crédito en la comunidad científica correspondiente, de manera que la validez científica del método usado, y con ello la calidad de los resultados alcanzados, pudiera ser objeto de discusión. En segundo lugar, la fiabilidad atribuible a una prueba científica depende también de su *calidad técnica*. Cabe hablar a este respecto de *corrección técnico-procedimental*, en referencia a todo el proceso que conduce desde el descubrimiento o registro del vestigio o de la muestra hasta su análisis en el laboratorio (por ejemplo, a efectos de atribuir fiabilidad al resultado del análisis de una huella dactilar, el problema no es tanto, o no sólo, la validez científica de la prueba, sino, en primer lugar, saber quién tomó la huella, por orden de quién, en qué objeto estaba depositada, en qué punto concreto, cómo fue la cadena de custodia, etc.); y lo mismo con respecto al análisis de una mancha de sangre, orina, saliva)⁷. Pero cabe hablar también de *corrección técnico-científica*, en referencia a su correcta realización en laboratorio: por personal cualificado y siguiendo los protocolos apropiados. La regla aquí debería ser: «cuanto mayor es la expectativa de valor probatorio depositada en una prueba, más rigurosos deben ser los controles de realización de la misma». Y por último, en tercer lugar, no hay que olvidar que las pruebas las realizan personas de carne y hueso, y que por tanto pueden cometer errores. No son pocos los estudios que en los últimos tiempos llaman la atención sobre los *riesgos cognitivos* de algunas pruebas científicas⁸, sobre todo de aquellas tradicionales que, como la dactiloscopia y la grafística, tienen un fuerte componente comparativo que las deja enteramente bajo la supervisión del perito⁹.

⁶ En su informe, Champod y Vuille (2010) realizan un exhaustivo repaso por los factores que pueden influir negativamente en el valor probatorio atribuible a los resultados de las pruebas científicas. A continuación destacamos algunos, pero para una visión completa remitimos a ese excelente estudio (que puede consultarse en [www.coe.int/t20/Legal_affairs_pp_10_y_ss_1](http://www.coe.int/t/t20/Legal_affairs_pp_10_y_ss_1)).

⁷ Es evidente, por ejemplo, que, pese al potencial de la prueba de ADN, el estado en el que llegan los vestigios biológicos al laboratorio es crucial: si los vestigios no han sido bien recogidos o conservados (por ejemplo, porque han sido contaminados por un ADN extraño) la posibilidad y el rendimiento del análisis se reduce. Por eso la *regadía de indicios* ha de hacerse con sumo cuidado, y el *mantenimiento de la cadena de custodia* es fundamental para que los indicios no pierdan su valor probatorio. Llamo la atención a este respecto al entusiasmo con el que muchos países se han abierto a las pruebas científicas (particularmente a las de ADN) sin un marco normativo previo que regule los procedimientos de obtención y conservación de los datos y garanticen en consecuencia la fiabilidad de los resultados.

⁸ Con respecto a los riesgos cognitivos cfr. DROR, 2009: 93-110; DROR y COLE, 2010: 161-167.

⁹ CHAMPOD y VUILLE, 2010: 10-11.

En conclusión, la validez de la prueba científica (y por consiguiente la fiabilidad de sus resultados) no es algo que haya que dar por descontado, sino que depende de la *validez científica del método usado*, de que se haya utilizado la *tecnología apropiada* y de que se hayan seguido rigurosos *controles de calidad*. Por consiguiente, incluso en relación con la prueba de ADN, que con el tiempo ha conseguido un altísimo grado de solidez científica y se presenta como el modelo a seguir por el resto de los campos de la ciencia forense, no puede haber ninguna duda sobre la necesidad de prestar atención a estas dos últimas cuestiones a la hora de evaluar el crédito que dicha prueba merece. Ahora bien, cumplidas todas estas precauciones, el problema de las pruebas científicas no reside ni mucho menos en la prueba del ADN. El problema reside en que hay otras muchas áreas de la policía científica que están muy lejos de haber gozado del mismo grado de atención y desarrollo científico que la genética forense, pero sobre las que sin embargo se fundan diariamente muchas decisiones judiciales.

Del mito de la infalibilidad (o lo que es lo mismo, de la falta de reflexión seria sobre el estatuto epistémico de las pruebas científicas) derivan algunas consecuencias de importancia. Seguramente la más evidente y adversa sea la falta de control sobre su validez y fiabilidad, lo que sin duda permite la entrada en el proceso de auténtica *ciencia basura*, datos sin fundamento científico alguno con los que los laboratorios y los peritos hacen un auténtico negocio. Piénsese sin ir más lejos en la grafología o en el polígrafo. Por otra parte, y vinculado a lo anterior, el mito de la infalibilidad de la prueba científica nos enfrenta al peligro del desconocimiento de los errores judiciales que pueden cometerse con base en ella. El denominado *Innocent Project*, puesto en marcha por los abogados Barry SHECK y Peter NEUFELD en la Cardozo Law School para demostrar, mediante pruebas de ADN, la inocencia de un buen número de condenados, ha puesto de relieve no sólo la fragilidad de los medios de prueba tradicionales, como los testimonios y las confesiones, sino también de las pruebas científicas sobre las que se basaban algunas de estas condenas¹⁰. Por último, este mito entraña un riesgo adicional, un efecto adverso desde el punto de vista jurídico: en la medida en que el informe pericial se asume como infalible podría decirse que *es el propio perito quien indica al juez lo que debe creer sobre la hipótesis en consideración*, sin que el juez, por consiguiente, pueda separarse de tal juicio sin razones poderosas. Pero de este modo se termina convirtiendo a los peritos en decisores de la causa y, por consiguiente, instaurando un nuevo sistema de prueba fundado en la autoridad de los expertos.

1.3. Sobrevaloración semántica

Pero no es sólo el dogma de la infalibilidad de las pruebas científicas lo que debe ser revisado. También debe ser revisada otra cuestión: la que al prin-

¹⁰ Puede consultarse la página web del proyecto en www.innocenceproject.org.

cipio llamé la «sobreevaluación semántica»; más concretamente, el hecho de considerar que los resultados de la prueba (además de infalibles) dicen cosas distintas de las que en realidad dicen.

Si la prueba científica ha sido entronizada en el proceso es porque se da por descontado no sólo el altísimo valor probatorio o incluso infalibilidad de sus resultados, sino también que éstos hablan directamente de aquello que se pretende probar. Se piensa, en concreto, que el resultado de una prueba de ADN señala directamente la pertenencia o no del vestigio analizado a la persona de la que procede la otra muestra de ADN con el que aquél se contrasta (el acusado en una causa penal, por ejemplo); que el resultado de una prueba de balística dice directamente si el casquillo evaluado salió o no de la pistola del acusado; si la impresión de calzado analizada procede de la pisada de una cierta persona; si la escritura manuscrita examinada procede del demandado; si la voz analizada que realizó la llamada telefónica avisando de la colocación del artefacto explosivo pertenece o no al acusado, etc. Se piensa, en definitiva, que el resultado de una prueba científica habla en los términos en que el juez necesita pronunciarse.

Esta creencia expresa lo que en el ámbito de la ciencia forense se denomina el *paradigma de la individualización*, que se asienta sobre la supuesta capacidad de llegar a identificar plenamente a un individuo o a un objeto a partir de vestigios. Dicho paradigma se presenta como modelo a seguir para todas las técnicas identificativas en criminalística, y en términos generales consiste en sostener que el resultado de la prueba científica identifica (y además, categóricamente) un vestigio con una fuente, con exclusión de todas las demás¹¹.

Las cosas, sin embargo, no son así¹². Y de hecho en los últimos años el paradigma de la individualización ha recibido fuertes críticas por parte de la comunidad científica, que considera que, por más que se trate de un esquema

¹¹ Aludiendo al paradigma de la individualización, el NRC *Report, Strengthening Forensic Science in the United States: A Path Forward*, etc., dice que los analistas creen que algunas marcas son únicas: «(típicamente las impresiones de calzado y de neumáticos, las impresiones de crestas procedentes de la demás, marcas de herramientas y de armas de fuego, y exámenes de escritura manuscrita)». Y creen que «esa unicidad es transmitida fielmente desde la fuente a la evidencia que es objeto de examen (o en el caso de los exámenes de escritura manuscrita, que los individuos adquieren hábitos que dan lugar a una escritura individualizada). Cuando la evidencia y la fuente de donde pueda proceder se comparan, una conclusión de individualización implica que la evidencia se originó desde esa fuente, con exclusión de todas las demás fuentes posibles» (cap. I, Introduction, Section *Pressures on the Forensic Science System*, Paragraph *Questionable or Questioned Science*: 43-44).

¹² Uno de los ensayos más estimulantes de los últimos tiempos sobre el uso de la estadística en la ciencia forense, es la obra *Statistical Evidence* (ROYALL, 1997: xi) que denuncia explícitamente que el uso de métodos estadísticos estándar conduce muchas veces a tergiversar los resultados de las pruebas: unas veces dándoles un peso mayor o menor del que realmente tienen; otras, considerando que los *datos estadísticos apoyan un resultado cuando en realidad sucede justo lo contrario*. Y esto último sucede, no porque los expertos usen equivocadamente la estadística, sino justamente porque domina el paradigma de la individualización.

de análisis simple e intuitivo, no resulta aceptable: la pretensión no comprobada de que es posible vincular un vestigio desconocido a una única fuente representa una equivocada intuición probabilista que iguala infrecuencia con unicidad¹³. O dicho muy brevemente, se considera imposible obtener prueba concluyente de la unicidad¹⁴.

Como resultado de lo anterior los científicos consideran que hay que interpretar los resultados de las pruebas, no en términos de identificación (de un vestigio con una fuente) sino en términos de *razón de verosimilitud* (*likelihood ratio*). Lo que se sostiene, más concretamente, es que no hay fundamento científico alguno para que un perito sostenga *categoricamente* que ha sido capaz de identificar a una persona o un objeto a partir de los análisis llevados a cabo en el laboratorio. Pero la crítica va más allá: no hay tampoco fundamento científico alguno para que en el informe pericial se realice siquiera una valoración *probabilística* sobre la posibilidad de atribuir a una persona o a un objeto el vestigio analizado¹⁵. El resultado de un análisis de voz, por ejemplo, o el de una comparación de perfiles de ADN, no dice (ni conclusiva ni probablemente) que la voz, o el ADN del vestigio analizado, pertenece a tal o a cual persona, sino que sólo aporta *datos* que, una vez *interpretados* con las adecuadas herramientas estadísticas, dicen cosas del siguiente tipo: «es X veces más probable que se observe tal rasgo en la voz analizada si ésta pertenece al acusado que si no pertenece a él»; o «es X veces más probable que coincidan los perfiles genéticos si el vestigio analizado procede del acusado que si procede de una fuente distinta». En otras palabras, las pruebas científicas hablan de la *probabilidad de los datos analíticos y técnicos* resultantes tras el análisis en el laboratorio a la luz de las hipótesis judiciales examinadas, y no al revés; es decir no hablan de la probabilidad de las hipótesis judiciales consideradas a la luz de esos datos. Por eso la extendida costumbre de expresar los resultados de las pruebas científicas en términos de identificación categórica o incluso probabilística (particularmente extendida en las técnicas forenses tradicionales) debe ser abandonada. Por lo demás, como ha subrayado CHAMPOD en su dura crítica al paradigma de la individualización, la ciencia forense, para ser útil, no necesita manejar ese paradigma, por lo que debe ser erradicado de la comunidad forense¹⁶.

¹³ SAKS y KOEHLER, 2008: 199-219; KAYE, 2009: n. 39, véase en línea: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1261970. Como KAYE comenta (n. 39), existe un riesgo nunca igual a cero al aceptar cualquier inferencia sobre un parámetro poblacional. «La distancia entre la muestra y la población requerirá siempre un salto de fe. Lo único que vale la pena debatir es la longitud del salto». Cfr. también STONEY, 1991.

¹⁴ Por eso, «un científico o un juez objetivo que revisara el estado actual de las ciencias forenses tradicionales probablemente calificaría las afirmaciones hechas por éstas como plausibles, carentes de suficiente investigación y sobreevaluadas». SAKS, KOEHLER, 2008: 892.

¹⁵ Para un análisis detallado de las críticas recibidas por el paradigma de la individualización, remito al trabajo conjunto GASCÓN, LUCENA y GOZÁLEZ, 2010.

¹⁶ CHAMPOD, 2009.

2. EL PARADIGMA DE LA VEROSIMILITUD

Pero si el paradigma de la individualización no resulta válido, entonces cabe preguntar, ¿cómo deben interpretarse los datos aportados por una prueba científica?, ¿cuál es el valor que hay que dar a los resultados de este tipo de pruebas? El llamado por ROYALL *paradigma de la verosimilitud*¹⁷ ofrece una respuesta a esta cuestión. Dicho paradigma se erige sobre la distinción entre las tres preguntas básicas que cabe formular una vez que el perito ha realizado ya los análisis pertinentes en una prueba científica: ¿qué nos dicen los datos u observaciones resultantes de esos análisis sobre la hipótesis A en relación con la hipótesis B?, ¿qué debemos creer a partir de esos datos? y ¿qué debemos hacer?¹⁸ Como enseña se verá, esta distinción se presenta como una importante herramienta a la hora de reconstruir cómo interacciona con el proceso judicial una prueba que se desarrolla fuera del mismo. De hecho permite realizar un deslinde neto entre las tareas del perito y las del juez.

2.1. Lo que dicen los datos

Responder a la primera pregunta (*qué dicen los datos*) es la principal tarea del perito en el proceso, quien (por ejemplo en una prueba de comparación de voces) deberá *interpretar el resultado* de la prueba en el laboratorio (supongamos que positivo: las voces comparadas coinciden en el rasgo-R) dándole un alcance en relación con las hipótesis enfrentadas (supongamos que hipótesis-A: la voz pertenece al acusado; o hipótesis-B: la voz no pertenece al acusado). Muy simplemente, se trata de interpretar qué nos dicen los datos resultantes de la prueba sobre una hipótesis en relación con la otra (por ejemplo en estos términos: «es X veces más probable que se observe el rasgo-R en la voz analizada si ésta pertenece al acusado que si no pertenece a él»). Insistamos en que la interpretación de esos resultados ha de hacerse en términos de *razón de verosimilitud*, y es importante que se haga bien, porque es lo que se refleja en el informe de conclusiones y por consiguiente lo que se va a comunicar al juez.

2.2. Lo que debe creerse

Una vez que los datos han sido interpretados y expresados adecuadamente en el informe pericial hay que responder a la segunda pregunta (*qué debe creerse a partir de esos datos*), es decir hay que *evaluar* la veracidad de las

¹⁷ ROYALL, 1997.

¹⁸ Para un desarrollo más extenso del alcance de este paradigma véase GASCON, LUCENA y GOZALEZ, 2010.

hipótesis enfrentadas (A y B), y ello ha de hacerse a partir de lo que dicen esos datos pero teniendo en cuenta también lo que dice el resto de las pruebas e informaciones disponibles.

Es evidente que esta tarea de valoración de las hipótesis en juego a partir de los datos aportados por la prueba corresponde al juez, al menos por dos razones. Primero porque es el juez quien tiene *institucionalmente* atribuida la función de proveer una solución jurídica para un caso controvertido o litigioso, y por tanto la de determinar previamente cuáles han sido los hechos que han dado origen al conflicto. Y segundo porque también el principio de *libre convicción*, en la medida en que consagra que el juez no debe dar por probado un hecho entretanto no se haya convencido de su verdad, exige atribuir a este la tarea de determinar lo que hay que creer a partir de la prueba, incluida la científica¹⁹. Resulta claro, por consiguiente, que el paradigma de la individualización no es comparable con el hecho de que sea el juez quien tiene asignada en exclusiva la tarea de valorar la prueba, pues si el informe pericial hablara en términos de «el vestigio y proviene (o no proviene) de la fuente f» habría sido el perito (y no el juez) quien habría determinado lo que hay que creer sobre las hipótesis en conflicto.

En suma, la neta distinción entre la tarea del perito (interpretar y comunicar lo que dicen los datos) y la del juez (valorarlos a la luz de los demás datos y pruebas disponibles) es el núcleo del paradigma de la verosimilitud. El paradigma de la individualización, por el contrario, no distingue entre la tarea del perito y la del juez, pues propicia que el perito declare en términos de «el vestigio y procede de la fuente f (por ejemplo, el acusado)», que son los términos en que debe expresarse el juez en ejercicio de la tarea de valoración de la prueba que sólo a él le corresponde²⁰. Pero de este modo importantes principios jurídicos, como el de la libre convicción, resultan conculcados.

2.3. Lo que debe hacerse

Por último, el paradigma de la verosimilitud también traza la distinción entre *lo que debe creerse* y *lo que debe hacerse* a partir de la realización de una prueba científica. Dicha distinción resulta sumamente apropiada para reconstruir otra diferenciación esencial dentro del universo de la prueba: la

¹⁹ Por eso la «cientificidad» de la prueba, por sí misma, no cierra la cuestión de su valor probatorio, que debe ser resuelto por el tribunal en cada caso en virtud del principio de libre valoración. En el mismo sentido WROBLEWSKI (1969: 185): «La referencia a la ciencia especializada no excluye la valoración de las pruebas».

²⁰ Por lo demás, este paradigma comporta un riesgo importante de comisión de falacias. Como la conocida falacia del fiscal (*prosecutor fallacy*). Sobre la *falacia del fiscal* y la *falacia de la defensa*, cfr. THOMPSON y SCHUMANN, 1987: 167. Sobre la falacia del fiscal puede consultarse más recientemente, LEUNG, 2002: 44-50.

existente entre la valoración de la prueba y la decisión sobre los hechos probados. Mientras que lo que debe *creerse* fundadamente sobre una hipótesis se expresa como la probabilidad de esa hipótesis y puede identificarse con la *valoración* de la prueba, lo que debe *hacerse* (o sea, la toma de decisiones en presencia de esa probabilidad) expresa la cuestión de los *estándares* de prueba; es decir, la determinación del grado de probabilidad que debe haber alcanzado una determinada hipótesis para que podamos considerarla probada y actuar en consecuencia. Y así como la valoración de la prueba es (o debería ser) una cuestión completamente librada a la *racionalidad epistémica*, la fijación de estándares de prueba es también una cuestión de *policy*, es decir política o valorativa, dependiente de lo tolerante que el sistema esté dispuesto a ser con cada uno de los dos errores que pueden cometerse al adoptar una decisión: declarar probado lo falso y declarar no probado lo verdadero. Es evidente, por ejemplo, que el estándar de prueba que suele exigirse en el proceso civil es mucho más bajo que el que se exige en el penal. De manera que mientras que en el primero podríamos aceptar que la prueba de un hecho descansa sobre una probabilidad preponderante (digamos, del 55 por 100), en el segundo sólo estamos dispuestos a condenar con una probabilidad muy cualificada; y ello por la simple razón de que somos menos tolerantes con el error consistente en condenar a un inocente que con el consistente en absolver a un culpable.

En suma, *lo que hay que creer* sobre una cierta hipótesis depende de lo que dicen los datos sobre la misma y del resto de las informaciones disponibles. *Lo que debe hacerse* depende obviamente de lo que hay que creer, pero también del contexto normativo en el que se encuadra la decisión a adoptar. Por eso, ante una misma creencia, no es lo mismo tomar una decisión en el marco de un proceso civil que en el marco de un proceso penal. Y aún en este último supuesto, no es lo mismo que la decisión se adopte en el marco de la instrucción que en el del juicio oral.

3. LOS RETOS

De lo dicho hasta aquí derivan algunas consecuencias de importancia que plantean importantes retos de cara a la consecución del objetivo final de racionalización de la función de juzgar cuando se utiliza prueba científica. La primera tiene que ver con el control de la validez y fiabilidad de la prueba y deriva de la conciencia de falibilidad de las pruebas. La segunda está relacionada con el contenido y sentido del informe pericial, esto es, con el modo en cómo han de formularse las conclusiones en el informe, y deriva del paradigma de la verosimilitud. La tercera, por último, hace referencia al modo en que ha de valorarse la prueba científica y deriva también del paradigma de la verosimilitud.

3.1. Sobre la admisibilidad de la prueba científica. El juez como *gatekeeper*

La admisibilidad es el control de entrada de la prueba en el proceso. Controlar que los informes periciales admitidos al proceso tengan un sólido fundamento científico es de un interés epistémico máximo, pues dicho control aspira a dejar a la pseudociencia fuera del ámbito de las decisiones judiciales. De hecho todos los esquemas de control de validez y fiabilidad (en definitiva, de calidad) de las pruebas científicas comparten ese objetivo. Cómo conseguir esto es lo que marca la diferencia entre ellos.

En línea de principio cabe adoptar dos grandes posturas sobre la admisibilidad de la prueba científica. Cabe sostener, por un lado, que *no es necesario un control específico de admisibilidad científica de la prueba* (validez científica) *junto con el control de admisibilidad procesal* (relevancia y legalidad) y que, por consiguiente, todas las pruebas científicas relevantes y legalmente obtenidas deben ser admitidas, dejando la cuestión de su validez científica para el momento de la valoración²¹. Pero también cabe sostener que *junto al control de admisibilidad procesal debe haber otro de admisibilidad científica* diferente de la fase de valoración de la prueba, es decir de atribución a la misma de valor probatorio²².

En Europa en general se ha seguido la primera postura. Las pruebas científicas relevantes y legalmente válidas entran en el proceso sin traba alguna y son valoradas por el juez, que les atribuye mayor o menor fuerza probatoria dependiendo, entre otros factores, de la validez científica y fiabilidad de la metodología usada para obtenerlas. En otras palabras, la validez y fiabilidad de una prueba científica no se evalúa en el trámite de admisibilidad de la prueba sino que se pondera en el momento de la valoración y se refleja en el valor probatorio que se le atribuye a dicha prueba. Pero esta opción presenta algunos inconvenientes.

En primer lugar una importante cuestión conceptual que, no obstante, tiene trascendencia práctica. Este modelo no distingue entre la fase de admisibilidad científica de la prueba y la fase de valoración de la misma, de modo que la *fiabilidad científica* de la prueba se confunde a menudo con su *valor probatorio*. Debe observarse, sin embargo, que admitir una prueba científica no implica pronunciarse sobre su valor probatorio sino meramente sobre la posibilidad de tomarla en consideración para la decisión del caso. Ambos conceptos (validez científica y valor probatorio) deben ser claramente distinguidos.

²¹ Lo que se sostiene, en definitiva, es que, si la prueba no es válida, ya se encargará el juez en el momento de la valoración de privarla de valor probatorio.

²² Puede consultarse in extenso una confrontación de los modelos de control aludidos en HANAJ-NEE y AL-RAWASHDEH, 2010: 187 y ss.

En segundo término está también el problema de que los jueces, por lo general, carecen de los conocimientos científicos necesarios para evaluar la validez y fiabilidad de estas pruebas. Por eso dejar enteramente en sus manos la libre apreciación de este dato, sin establecer un estándar objetivo para guiarlos, puede conducir a rechazar pruebas científicamente válidas que podrían haber ayudado al esclarecimiento de la verdad en el proceso, o —lo que es más probable— puede conducir a dar crédito a pruebas que son pura charlatanería pseudocientífica y que pueden convertirse en fundamento de decisiones judiciales erróneas. No puede desconocerse la gran capacidad de persuasión e influencia que una prueba científica puede tener sobre la convicción judicial, con lo que si se admite su entrada en el proceso sin ningún control se corre el riesgo de que, al final, la decisión venga determinada por pura y simple *juris science*.²³ Por eso parece más apropiada la opción alternativa, esto es, la que aboga por establecer un estándar objetivo de admisibilidad científica de las pruebas junto al examen de su admisibilidad procesal.

Ésta es la opción que se ha seguido en Estados Unidos. Allí, el que se conoce como test Daubert²⁴ ha establecido como estándar de admisibilidad, junto con la *aceptación general* de la teoría y técnica usadas por la comunidad científica concernida, la *efectiva validez científica* de la técnica aplicada, que ha de acreditarse evaluando los siguientes factores: a) si la teoría o la técnica se puede probar y si ha sido efectivamente *puesta a prueba* (verificabilidad y refutabilidad como criterios mínimos de científicidad de cualquier teoría); b) si se han realizado publicaciones especializadas en el tema sometidas a una *revisión de pares*, y c) cuáles son los *márgenes de error* potencial asociados al método. No es que el test Daubert exija que cualquier técnica científica usada como fuente de prueba cumpla todos los criterios mencionados, pues esto es algo que queda al criterio del juez. Pero sí exige a los jueces —y esto es lo más significativo— escribir de manera más estricta muchas pruebas científicas que hasta ahora no habían sido cuestionadas y conduce así a que éstos recuperen su papel de *gatekeeper*.²⁵ Por decirlo muy sumariamente, la sentencia *Daubert*

²³ La admisión al proceso de prueba científica no fáble «confunde a los jueces, socavando la correcta función de las decisiones en materia penal.» (ROBERTS Y ZUCKERMAN, 2004: 69).

²⁴ El estándar ha sido establecido en el caso *Daubert v. Merrill Dow Pharmaceuticals Inc.*, 113 S. Ct. 2786 (1995). El caso que allí se abordaba era la admisibilidad de nueva o inusual prueba científica. Y la cuestión legal que se debatía era si el tradicional *Frye test* vigente desde 1923, que atribía la admisibilidad de la prueba en la «aceptación general» del método usado por parte de la comunidad científica, había sido sustituido por la *Federal Rule of Evidence 702*, de 1975, que establecía que el testimonio de los expertos resultará admisible cuando, aparte de relevante, esté basado en datos suficientes que sean el resultado de principios y métodos científicos válidos que hayan sido aplicados correctamente por el experto. La sentencia *Daubert* respondió afirmativamente a esa cuestión y estableció además —y éste es el corazón del test Daubert— que son los tribunales los que deben escrutar si el testimonio de los expertos cumple esos parámetros de relevancia y de efectiva «validez científica» de método usado (agradecemos a Susan HAACK las atinadas precisiones que me hizo sobre el tema).

²⁵ Prácticamente una réplica del test Daubert cabe encontrar en los criterios sugeridos para el Reino Unido por la LAW COMMISSION, 2009: 53 y ss. (LAW COMMISSION, *The Admissibility of Expert Evidence in Criminal Proceedings in England and Wales. A New Approach to the Determination of*

supone un llamamiento a los jueces para que miren más críticamente las pruebas científicas.²⁶ De hecho, después de esta sentencia, ámbitos tradicionales de la policía científica, como las huellas papilares, la grafística y la balística empiezan a ser puestos seriamente en cuestión.

Pero el test Daubert tampoco ha zanjado la polémica porque su aplicación presupone que los jueces son capaces de comprender y aplicar los conceptos científicos necesarios para evaluar la validez de estas pruebas; lo que acaso no siempre sea cierto. Por supuesto cuando la prueba ha sido ya objeto de un intenso debate y aceptación por parte de la comunidad científica, como sucede con la prueba del ADN, no se plantea ninguna cuestión sobre su admisibilidad. Pero siempre habrá otras especialidades de la ciencia forense con mayor fragilidad científica que la prueba de ADN. No debe extrañar por ello que para muchos cada vez esté más clara la necesidad de formar a los jueces en los distintos ámbitos de especialidad forense, y sobre todo la necesidad de proporcionarles una formación específica en las nuevas áreas que vayan surgiendo. En todo caso muchas cuestiones permanecen abiertas: ¿es realista instar a los jueces a adquirir conocimientos científicos para poder evaluar la fiabilidad de la prueba científica?²⁷ Puesto que no todos los jueces tendrían los mismos conocimientos, ¿no acentuaría esto la disparidad de las actuaciones en práctica? ¿Pero no es aún más peligroso dejar esta tarea enteramente en manos de los

Evidentiary Reliability: A Consultation Paper, www.law.com.gov.uk/expert_evidence.htm, Paper 2009, 190, pp. 53 y ss.).

²⁶ Para determinar si la prueba pericial científica (o supuestamente científica) es lo suficientemente fáble como para ser admitida, el tribunal debe considerar los siguientes factores y cualesquiera otros que se consideren relevantes:

a) Si los principios, técnicas y presupuestos en los que se basa han sido adecuadamente sometidos a prueba y, de ser así, el grado en que los resultados de tales pruebas demuestran su solidez.

b) El margen de error asociado a la aplicación de esos principios, técnicas y presupuestos, así como a las conclusiones obtenidas a partir de los mismos.

c) Si hay un cuerpo de literatura especializada en ese campo de conocimiento.

d) En qué medida los principios, técnicas y presupuestos han sido examinados por otros científicos —por ejemplo en publicaciones sometidas a revisión por pares— y, de ser así, en qué medida se consideran sólidos en la comunidad científica.

e) La cualificación, experiencia y publicaciones relevantes del testigo experto y su prestigio en la comunidad científica.

f) La validez científica de los puntos de vista opuestos (si es que los hay) y la cualificación, experiencia y prestigio profesional en la comunidad científica de los científicos que los mantienen.

g) Si hay pruebas que sugieran que el testigo experto no ha actuado de conformidad con su primordial deber de imparcialidad.

²⁷ Con elocuentes palabras CHAMPROD y VUURLE (2010: 26), la sentencia *Daubert* «insiste de manera implícita sobre el escrutinio que el juez debe mantener hacia el perito, el cual deja de ser considerado como el miembro de una élite con autoridad y pasa a ser un agente social comparable a cualquier otro, eventualmente sometido a presiones de orden político y económico que pueden alterar su dicterio».

²⁸ Ésta es la razón por la que, en su informe de 2009, la Law Commission propone para el Reino Unido la posibilidad de que excepcionalmente, en los casos muy delicados, el juez pueda solicitar la ayuda de los expertos para evaluar la validez científica de la técnica propuesta (me refiero al informe *The Admissibility of Expert Evidence in Criminal Proceedings in England and Wales. A New Approach to the Determination of Evidentiary Reliability*, 2009). Esto mismo esta sugerencia sumamente interesante.

expertos, que podrían tener intereses ocultos en la promoción de ciertas técnicas? ¿No sería posible un sistema donde los jueces decidieran con la ayuda de un órgano independiente y científicamente competente?²⁸

3.2. Sobre el modo de transmitir los resultados de la prueba científica

Aparte del central asunto de la fiabilidad de la prueba científica hay otra importante cuestión relativa al informe pericial y a la que desgraciadamente no se le ha prestado demasiada atención²⁹. La cuestión concierne, más precisamente, al modo en que los peritos deberían interpretar y expresar los resultados de los test en su informe para evitar que sean malentendidos o sobervalorados por los jueces. Muy brevemente, cómo deben formularse las conclusiones del informe pericial para comunicar correctamente los resultados de las pruebas a los tribunales.

De acuerdo con el paradigma de la verosimilitud al perito no le corresponde emitir una opinión sobre la hipótesis en consideración sino sólo *dar cuenta de los datos de un modo científicamente riguroso*, permitiendo así que el juez entienda exactamente su significado y pueda valorarlos junto con el resto de las pruebas disponibles. En relación con esto conviene insistir en una cuestión central: las conclusiones de la prueba que el perito formula en su informe se orientan a determinar el significado (o el grado de probabilidad) de los datos analíticos o técnicos resultantes a la luz de todas las hipótesis en consideración, y no el significado (o el grado de probabilidad) de las hipótesis en consideración a la luz de los datos analíticos o técnicos resultantes. Entender que las conclusiones de la prueba hablan en estos últimos términos, esto es, que las pruebas le dicen al juez justamente lo que él debe determinar para fundar su decisión, es propio del paradigma de la individualización, que no distingue entre lo que dicen los datos (una tarea pericial) y lo que debe creerse a partir de los datos (una tarea judicial), y da lugar a múltiples malinterpretaciones y, en ocasiones, a serios errores judiciales. Esto, por ejemplo, es lo que explica que se produzcan las conocidas como falacia del fiscal y falacia de la defensa³⁰.

²⁸ «Una posibilidad podría ser establecer un órgano europeo, al estilo del British Forensic Science Advisory Council and Forensic Science Regulator, que actuara como el principal consultor de las autoridades políticas y jurídicas en relación con la fiabilidad de las técnicas científicas que se usen [...]. Ese órgano no tendría poderes vinculantes, pero podría emitir recomendaciones para los tribunales nacionales que auxiliarían a los jueces cuando tuviesen que pronunciarse sobre la admisibilidad de una nueva técnica forense o sobre la fiabilidad de una nueva forma de prueba, o cuando pareciera adecuado abandonar una forma de prueba que hubiera quedado obsoleta», CHAMPOD y VUILLE, 2010: 109.

²⁹ De ello se lamentaba no hace mucho, entre otros, WILLIS, 2009: 523-545. El *NRC Report* también presta atención a este asunto, al establecer que es absolutamente necesario sentar las bases «para la correcta valoración y comunicación de los hallazgos forenses» (*Reporting Results Section*, p. 186).

³⁰ Sobre la *falacia del fiscal* y la *falacia de la defensa*, cfr. THOMPSON y SCHUMANN, 1987: 167 y ss. Y, entre nosotros, CARRACEDO, 1999: 303. Sobre la falacia del fiscal puede consultarse más recientemente LEUNG, 2002: 44-50.

Probablemente no hay fórmulas mágicas para expresar las conclusiones en el informe, y probablemente no hay una sola fórmula. Pero lo que sí está claro es que debe huirse de las expresiones (categóricas o incluso sólo probabílicas) sobre las hipótesis planteadas por las partes con respecto a un caso criminalístico, y debe en cambio optarse por pronunciamientos del tipo «es *R* veces más probable que coincidan los perfiles genéticos si el vestigio *x* procede de la fuente *f* que si no procede de *f*»; o «es *R* veces más probable que coincidan las huellas dactilares si la huella es del acusado que si no lo es»³¹. Por eso las llamadas escalas verbales de probabilidad, ampliamente usadas en el análisis de huellas digitales, en el examen de escritura, en la balística, en la comparación de caras, en la identificación de voz mediante análisis lingüístico o en la comparación de huellas de calzado, de neumáticos y de herramientas son erróneas. De hecho, estas escalas reproducen el paradigma de la identificación estableciendo, no lo que dicen los datos, sino lo que debe creerse sobre la hipótesis en consideración a partir de los mismos. En este sentido un importante reto que la ciencia forense tiene ante sí es el establecimiento de unos estándares de conclusiones que guíen el trabajo de los laboratorios y que al propio tiempo aseguren el uso de criterios uniformes en este punto³².

Y una última observación al respecto. La estadística juega un papel crucial en la interpretación de los datos resultantes de las pruebas científicas. De hecho, los datos plasmados en las conclusiones del informe pericial que después han de ser valorados por los jueces son, en la mayoría de los casos, datos estadísticos. Ahora bien, está comprobado el peso desproporcionado (*overweighing thesis*) que el público en general y los jueces en particular atribuyen a la estadística³³. Precisamente por eso no parece descabellado plantear la necesidad de que los jueces reciban alguna formación en este sentido. Obviamente no se trata de convertir a los jueces en matemáticos, o en *amateur scientists*, pero sí de proporcionarles los conocimientos necesarios para conjurar el riesgo de malinterpretación o de sobrevaloración de los datos estadísticos en los que se expresa el resultado de una prueba científica.

De todos modos, y aunque no cabe dudar de los beneficios que reportaría un adiestramiento como el comentado, ello no anularía el importante papel auxiliar de los peritos en la tarea de interpretación de los datos del informe. Muy simplemente, no siempre resultará fácil para el juez, como para cual-

³¹ Me permito remitir aquí, de nuevo, al trabajo de GASCÓN, LUCENA y GOZÁLEZ, 2010.

³² En relación con este punto cabe mencionar que el Laboratorio Sueco de Ciencia Forense (SKL) anunció en 2009, tras años de debate interno, su decisión de unificar el modo de redactar las conclusiones para así proporcionar la misma información a los tribunales independientemente del tipo de prueba de que se trate.

³³ Por ceñirnos sólo a Europa, son conocidos, entre otros, el caso de Sally Clark, en Inglaterra, o el de Lucia de Berk, en los Países Bajos, absuelta en abril de 2010. Ambos, y algunos más, son citados como ejemplo de lo dicho por CHAMPOD y VUILLE, 2010: 10.

quiera que no sea un experto, aislar en dicho informe *qué cosa* dicen exactamente los datos y *qué valor* hay que atribuir a eso que dicen los datos. Por más claro y riguroso que sea el informe, puede ser inevitablemente complejo y objeto de malinterpretaciones. Por eso es necesario que, como regla general, se requiera la presencia del experto en el acto de juicio para ayudar a interpretar los datos del informe y evitar así que éstos sean malinterpretados por el juez. Esta presencia representa, en este sentido, una *garantía epistemológica*³⁴.

3.3. Sobre la valoración de la prueba

La valoración de la prueba es el núcleo de la decisión probatoria y consiste en determinar lo que hay que creer sobre la hipótesis en consideración. Valorar la prueba científica, por consiguiente, consiste en atribuirle valor probatorio en relación con la hipótesis que describe los hechos del caso. Son dos las cuestiones relevantes que se plantean en relación con la valoración de la prueba: *quién debe valorar y cómo* llevar a cabo la valoración.

¿Quién debe valorar? La respuesta a esta cuestión puede parecer evidente, porque así es.

Si es el juez quien tiene la tarea de proveer una solución normativa para el caso planteado, entonces es también a él a quien corresponde la tarea previa de determinar autoritadamente cuáles han sido los hechos que han dado origen al conflicto. Sin embargo no es ocioso insistir en ello, pues, precisamente en consonancia con el paradigma de la identificación, muchos tribunales acúan «como si» fuese el perito, y no el juez, quien ha de valorar la prueba. Y es que, en efecto, al expresar las conclusiones de la prueba conforme al paradigma de la identificación (es decir, al declarar en términos de, por ejemplo, «a la vista de los datos resultantes de los test, la hipótesis es probable en un grado x»),

³⁴ Si esta insistencia tiene sentido es porque en los últimos tiempos algunas veces sostienen que no es necesario que los peritos de las instituciones oficiales —esto es, aquellos cuya independencia debemos presumir— acudan al proceso, puesto que lo que tenían que decir ya lo han dicho en su informe, y en consecuencia abogan por una reforma legislativa que confiera a los informes periciales de los laboratorios de la policía científica carácter de prueba documental. Se trataría, por consiguiente, de extender también a las pruebas científicas provenientes de los laboratorios oficiales el mismo tratamiento que ya la reforma del art. 788 de la Ley de Enjuiciamiento Criminal española confirió a los informes forenses emitidos por laboratorios oficiales sobre la naturaleza, cantidad y pureza de las drogas. «Este tipo de propuestas me parecen profundamente desacertadas. Podrán servir —esto no se cuestiona— para agilizar el acto de juicio (puesto que no será necesario que declaren los peritos sino que el juez, por su cuenta, examinará el documento del informe pericial), así como para optimizar los recursos de los laboratorios oficiales y facilitar el trabajo de sus expertos, que no se verán obligados a acudir al tribunal. Pero, al menos en el proceso penal, lesionan valores fundamentales vinculados al derecho de prueba, como la publicidad, la contradicción, la oralidad y la inmediación, como lesionan también el que constituye el principio medular de la decisión probatoria: el de *libre valoración*» (GASCÓN Y LUENNA, 2010: 105-106).

«los datos resultantes de los test hacen la hipótesis probable en un grado x» o «son compatibles con la hipótesis en un grado x», o aún más directamente, «el vestigio y procede de la fuente f»), es el experto (y no el juez) quien habrá establecido lo que hay que creer sobre la hipótesis. En virtud del paradigma de la verosimilitud, sin embargo, es el juez quien ha de valorar la prueba, o sea, quien ha de decir lo que ha de creerse sobre la hipótesis que se pretende probar.

Y una última consideración relacionada con este aspecto. Suele decirse muchas veces que, en virtud del principio de libre convicción, «el juez es *peritus peritorum*», expresión con la que se pretende destacar que es él quien tiene la última palabra sobre el valor que debe atribuirse a los datos expresados en el informe pericial y que, por consiguiente, *no está estrictamente vinculado* a dicho informe; lo cual —se dice también— resulta un tanto paradójico si se tiene en cuenta que él no es un experto y que por tanto carece de los conocimientos necesarios para formarse una opinión propia sobre la prueba realizada. La cuestión, sin embargo, está mal planteada, pues en realidad carece de sentido afirmar que «el juez es *peritus peritorum*». Conforme al paradigma de la verosimilitud lo que el juez debe hacer no es, o bien *vincularse* al informe pericial incluso si está convencido de que debe apartarse del mismo (infringiendo entonces el principio de libre convicción) o bien, por el contrario, *apartarse* del mismo en su condición de *peritus peritorum* (asumiendo entonces que posee conocimientos especializados que muy probablemente no posee). Lo que el juez debe hacer es otra cosa: *determinar lo que hay que creer sobre la hipótesis en consideración a la luz del informe pericial y (también) del resto de las pruebas e informaciones que obran en la causa*. El juez valora la prueba, no es perito de peritos.

¿Cómo valorar? La segunda y central cuestión es la de cómo realizar la valoración (o sea, cómo determinar lo que hay que creer sobre la hipótesis a probar) cuando se ha aportado prueba científica.

Cabe distinguir dos supuestos. Cuando únicamente existe prueba científica en apoyo de la hipótesis, lo que hay que creer sobre la misma viene enteramente determinado por (o es coincidente con) el valor probatorio atribuido a la prueba científica. Cuando, por el contrario, además de la prueba científica existe también prueba no científica, que es la situación más frecuente, determinar lo que hay que creer sobre la hipótesis a probar requiere ponderar el valor probatorio de los datos estadísticos aportados en el informe pericial con el valor probatorio atribuido al resto de las pruebas. Naturalmente el juez realiza normalmente esta valoración sin auxilio alguno de las matemáticas; «subjetivamente», por así decirlo. Sin embargo, dado que el valor probatorio de la prueba científica puede ser expresado en términos estadísticos, se han propuesto también instrumentos matemáticos para llevar a cabo la valoración.

El principal intento se ha realizado mediante la aplicación del teorema de Bayes³⁵ a las inferencias jurídicas basadas en probabilidades subjetivas³⁶. En su formulación más simple la fórmula bayesiana afirma que la probabilidad de un evento H, dado un evento E, puede determinarse en función de la frecuencia estadística con la que dado H se verifica E y de la probabilidad atribuida precedentemente al evento H.

$$P(H/E) = \frac{P(E/H)}{P(E/\text{no-H})} \times P(H)$$

P(H/E): probabilidad de H, dado un evento E (*posterior probability*).

P(E/H): grado de confianza de que ocurra E si H es verdadera.

P(E/no-H): grado de confianza de que ocurra E independientemente de que H sea verdadera.

P(E/H)/P(E/no-H) (*Likelihood ratio* de E respecto de H, que mide la mayor o menor probabilidad con la que ocurre E siendo H verdadera que sin serlo).

P(H): grado de creencia de que H es verdadera antes de saber si E ocurría (*prior probability*).

La utilidad de la fórmula bayesiana en el proceso reside en que permite combinar la información estadística aportada por una prueba científica (la procedente del informe pericial y expresada mediante una *likelihood ratio*) con el valor atribuido anteriormente por las otras pruebas a la hipótesis que se pretende probar (la *prior probability*), y el resultado de esa combinación es la *posterior probability*, que expresa el resultado de la valoración de la prueba. En realidad, como puede verse, la fórmula bayesiana mide el impacto que, sobre el hecho que se pretende probar, provoca la inclusión de nuevos datos de naturaleza estadística, como son los resultantes de las pruebas científicas³⁷.

Sin embargo las ventajas de utilizar el teorema de Bayes para mejorar la precisión y controlabilidad (en definitiva, la fiabilidad) de la valoración judicial sobre hechos se ven atenuadas por los problemas que presenta su uso en el proceso. De hecho el bayesianismo, como modelo general de valoración,

³⁵ El teorema de Bayes no es el único instrumento matemático que ha sido propuesto como instrumento de valoración. Las *redes bayesianas* también constituyen una poderosa herramienta al servicio de las inferencias probabilísticas en ciencia forense cuando concurren múltiples pruebas, científicas o no, bajo las hipótesis defendidas por las partes que intervienen en el proceso. Cf. TARONI, ARKEN, GAROLINO y BIEDERMANN, 2006.

³⁶ La probabilidad subjetiva indica el grado de creencia personal de un *factfinder* racional sobre la ocurrencia de un suceso incierto.

³⁷ FINKELSTEIN y FAURLEY, 1970. Sobre el bayesianismo pueden verse además los trabajos del simposio *Decision and Inference in Litigation*, *Cardozo Law Review*, 13, 1991.

ha sido objeto de numerosas críticas. Algunas de estas críticas están animadas por consideraciones *prácticas*. Por ejemplo, llaman la atención sobre lo difícil que resulta para el juzgador cuantificar su valoración subjetiva previa (o sea la *prior probability*); o sobre el riesgo que entraña poner en manos de los jueces instrumentos estadísticos incomprensibles para ellos. Además, la fórmula bayesiana supone un cálculo matemático relativamente simple en el caso-base de tener que valorar *una sola prueba* que versa *directamente* sobre la hipótesis a probar, que es además una *hipótesis simple*. Pero la complicación del cálculo aumenta de manera impresionante cuando se usa para resolver situaciones más complejas, como la *plurialidad* de elementos de prueba relativos a una hipótesis, la *cascaed inferencia* (o prueba mediata) o la prueba de una hipótesis referente a un *hecho complejo*. Otras críticas, en cambio, *ponen en cuestión la validez epistemológica del modelo en el campo judicial*, en la medida en que conduciría a resultados contraintuitivos y resultaría dudosamente compatible con principios fundamentales del ordenamiento, como la presunción de inocencia³⁸.

No cabe duda de que las críticas aludidas son serias y deben ser tomadas en consideración. Pero en cualquier caso los problemas del bayesianismo como modelo general de valoración de la prueba no anulan la utilidad de las herramientas estadísticas como modo de cuantificar (mediante una *likelihood ratio*) el valor probatorio de una prueba científica. Esto significa que en la valoración de la prueba acaso puedan coexistir dos distintos conceptos de probabilidad. De un lado la *probabilidad matemática o estadística*, que es usada para cuantificar el *valor probatorio* de la prueba científica y que se expresa mediante una *likelihood ratio*. De otro lado la *probabilidad lógica o inductiva*, que es usada para expresar la *convicción sobre la hipótesis* a probar a la luz de todas las pruebas e informaciones disponibles (científicas o no).

BIBLIOGRAFÍA

CARRACEDO, A., 1999: «Valoración de la prueba del ADN», en MARTÍNEZ JARRETA, B. (dir.), *La prueba del ADN en la medicina forense. La genética al servicio de la ley en el análisis de indicios criminales y en la investigación biológica de la paternidad*, Barcelona: Masson.

³⁸ El blanco de las críticas lo constituye aquí el papel desempeñado por la *probabilidad previa* del hecho que se quiere probar. Si el valor atribuido a la misma fuera bajo, también sería bajo el resultado de la valoración final, con independencia de que la probabilidad indicada por el nuevo material probatorio fuese alta. Pero además, y conectado con ello, si la formulación de la *probabilidad previa* estuviera viada de un error inicial, podría atentar contra principios básicos del ordenamiento, como la presunción de inocencia. En suma, el uso del teorema de Bayes produce una sobrevaloración de la probabilidad inicial y una minusvaloración del peso que tienen las nuevas pruebas en el cálculo de la probabilidad final. Para las deficiencias del bayesianismo, véanse TRIBE, 1971; TARUFFO, 2004; COHEN, 1977. Me remito también, en todo caso, a la atinada crítica de LAUDAN, 2005, a la aplicación del cálculo probabilístico en la construcción de un estándar de prueba.

- CHAMPOD, C., 2009: «Interpretation of evidence and reporting in the light the 2009 NRC reports». Keynote Speech of the Interpretation and Evaluation Session of the V Conference of the European Academy of Sciences held in Glasgow (Scotland), University of Strathclyde.
- CHAMPOD, C., y VUILLE, J., 2010: *Scientific Evidence in Europe. Admissibility, Appraisal and Equality of Arms*, European Committee on Crime Problems, 10.
- COHEN, H. I., 1977: *The Probable and the Provable*, Oxford: Clarendon Press.
- DROR, I., 2009: «How can Francis Bacon help forensic science? The four idols of human biases», *Jurimetrics Journal*, 50.
- DROR, I. E., y COLE, S. A., 2010: «The vision in "blind" justice: Expert perception, judgment, and visual cognition in forensic pattern recognition», *Psychonomic Bulletin & Review*, 17.
- FINKESTEIN, M. O., y FARLEY, W. B., 1970: «A Bayesian Approach to Identification Evidence», *Harvard Law Review*, 83.
- GASCON, M., 2010: «Prueba científica: mitos y paradigmas», en *Anales de la Cátedra de Francisco Suárez número conmemorativo: «Un panorama de filosofía jurídica y política (50 años de Anales de la Cátedra Francisco Suárez)»*, Granada, 44.
- GASCON, M., y LUCENA, J. J., 2010: «Pruebas científicas. La necesidad de un cambio de paradigma», *Jueces para la Democracia*, 71 (11).
- GASCON, M.; LUCENA, J. J., y GÓZÁLEZ, J., 2010: «Razones científico-jurídicas para valorar la prueba científica: una argumentación multidisciplinar», *Revista La Ley*, 7481 (4/10).
- HANNINER, A. M., y AL-RAWASHDEH, S. H., 2010: «Theoretical Approaches to Admitting Scientific Evidence in the Adversarial Legal System», *European Journal of Scientific Research*, 2 (41).
- IGARÚA, J., 2007: «Prueba científica y decisión judicial: unas anotaciones propedeuticas», *Diario La Ley*, 6812 (2/11).
- KAYE, D. H., 2009: *Probability, Individualization and Uniqueness in Forensic Science: Listening to the Academics*, 30th June, Social Science Research Network (SSRN). Disponible en línea: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1261970.
- LAUDAN, L., 2005: «Por qué un estándar de prueba subjetivo y ambiguo no es un estándar», *Doxa*, 28.
- LAW COMMISSION, 2009: *The Admissibility of Expert Evidence in Criminal Proceedings in England and Wales: A New Approach to the Determination of Evidentiary Reliability. A Consultation Paper*, www.lawcom.gov.uk/expert_evidence.htm, Part 190.
- LEUNG, W. C., 2002: «The prosecutor's fallacy - a pitfall in interpreting probabilities in forensic evidence», *Medicine Science and the Law*, 42.
- MARTÍNEZ JARRETA, B. (dir.) 1999: *La prueba del ADN en Medicina Forense. La Genética al servicio de la ley en el análisis de indicios criminales y en la investigación biológica de la paternidad*, Barcelona: Masson.
- NATIONAL RESEARCH COUNCIL OF THE NATIONAL ACADEMIES (NRC), 2009: *Report, Strengthening Scientific Evidence in the United States: A Path Forward*, Washington: The National Academies Press.
- ROBERTS, P., y ZUCKERMAN, A., 2004: *Criminal Evidence*, Oxford: Oxford University Press.

- ROYALL, R., 1997: *Statistical Evidence: A Likelihood Paradigm*, *Monographs on Statistics and Applied Probability*, London: Chapman-Hall/CRC.
- SAKS, M. J., y KOEHLER, J. J., 2005: «The Coming Paradigm Shift in Forensic Identification Science», *Science*, 5736 (309).
- 2008: «The Individualization Fallacy in Forensic Science Evidence», *Vanderbilt Law Review*, 1 (61).
- STONE, D. A., 1991: «What Made Us Ever Think We Could Individualize Using Statistics?», *Forensic Science Society*, 31 (2).
- TARONI, F.; ATTEN, C.; GARBOLINO, P., y BIEDERMANN, A., 2006: *Bayesian Networks and Probabilistic Inference in Forensic Science*, Chichester: John Wiley and Sons Ltd.
- TARUFFO, M., 2004: *La prueba de los hechos*, Madrid: Trotta.
- THOMPSON, W., y SCHUMANN, E., 1987: «Interpretation of statistical evidence in criminal trials. The prosecutor's fallacy and the defense attorney fallacy», *Law Human Behaviour Journal*, 11 (3).
- TRIBE, L. H., 1971: «Trial by Mathematics», *Harvard Law Review*, 6 (84).
- WILLIS, S. M., 2009: «Forensic science, ethics and criminal justice», en FRASER, J., y WILLIAMS, R. (eds.), *Handbook of forensic science*, Cullompton, UK: William Publishing.
- WRÓBLEWSKI, J., 1969: citado por la traducción de EZQUIAGA GANZUAS, F. J. (1989): «La referencia a la ciencia especializada no excluye la valoración de las pruebas», en *La prueba jurídica: axiología, lógica y argumentación*, en *id.*, «Sentido» y «hecho» en el derecho. San Sebastián: UPV-EHU.

VIII

LA APLICACIÓN DE ESTÁNDARES CIENTÍFICOS
A LAS CIENCIAS SOCIALES Y FORENSES *

Michele TARUFFO

Universidad de Pavia y Cátedra de Cultura Jurídica
de la Universidad de Girona

1.

El surgimiento (podría decirse: la explosión) del problema del uso de la ciencia como método para probar los hechos debatidos en cualquier procedimiento judicial es uno de los fenómenos más importantes de las últimas décadas y su importancia teórica y práctica crece día a día. En realidad, los métodos científicos han sido usados por la administración de justicia desde hace mucho tiempo, pero el problema solía concebirse sólo respecto al uso práctico de la prueba pericial o con los peritos como asesores del juez, sin prestar especial atención a la calidad y fiabilidad de la prueba obtenida por medio de métodos científicos o técnicos.

Quien quiera determinar un punto de partida en el desarrollo moderno y actual de lo relativo a la «prueba científica» en el significado moderno de la expresión, debe probablemente remitirse a 1993, fecha en la que la Corte Suprema de Estados Unidos adoptó la decisión *Daubert vs. Merrell Dow Pharm., Inc.* En su voto para la mayoría, el magistrado Blackmun intentó determinar

* Traducción de Maximiliano ARAMBURU.

Los criterios que el juez, como *portero (gatekeeper)*, en inglés) que controla la admisión de la prueba científica, debe aplicar para seleccionar la prueba que es «científicamente válida». Declarando inválido el llamado test Frye (establecido en 1923) de acuerdo con el cual un elemento probatorio era «científico» si el método aplicado tenía «la aceptación general» de la comunidad científica relevante, Blackmun escribió que un elemento de prueba es científicamente válido si: a) una teoría o técnica puede ser (y ha sido) sometida a prueba; b) la teoría o técnica ha sido sometida a revisión entre pares y publicada; c) se conoce la tasa de error actual o potencial; d) hay estándares que controlan la operación de la ciencia o técnica, y e) la ciencia o técnica tiene un grado significativo de aceptación dentro de la comunidad científica relevante.

Los argumentos ofrecidos por Blackmun han sido criticados desde el punto de vista epistemológico, dado que la epistemología popperiana en la que fueron inspirados parece ser dudosa por muchas razones. Además, la doctrina *Daubert* no ha sido aceptada por todas las jurisdicciones estatales de Estados Unidos, aunque es aplicada en los tribunales federales. Pero esa doctrina se amplió en el caso *Kumho* también a pruebas técnicas (y no únicamente a las científicas). Adicionalmente, *Daubert* produjo en el año 2000 un cambio importante en la regla 702 de las *Federal Rules of Evidence*, de acuerdo con la cual ahora un perito debe ser «calificado» como experto por su conocimiento, habilidad, experiencia, entrenamiento o educación, y puede testificar si «1) el testimonio se basa en hechos o datos suficientes; 2) el testimonio es el producto de principios y métodos fiables; y 3) el testigo ha aplicado de manera fiable los principios y los métodos a los hechos en cuestión». Es claro, por tanto, que la regla 702 se concentra en la *fabilidad* de las pruebas presentadas por el perito, enfatizando la necesidad, ya subrayada en *Daubert*, de que la «prueba científica» sea realmente «científica».

Algunos comentaristas dicen, de todos modos, que *Daubert* ha sido y sigue siendo sumamente importante porque se concentra —aunque desde una perspectiva epistemica discutible— en los problemas básicos de la validez científica de las pruebas que se emplean para decidir sobre los hechos del caso. Aun más, algunos académicos dicen que *Daubert* tomó el camino correcto al enfatizar el papel del juez en el control y la selección de la prueba que es verdaderamente «científica», pese a que no avanzó lo suficiente en esa dirección. Un papel tan activo del juez es algo relativamente nuevo en la estructura adversarial del proceso judicial estadounidense, pero es menos nuevo en los sistemas de *civil law*, donde usualmente el juez ya juega un papel activo en la presentación de pruebas. Sin embargo, aun en esos sistemas el mensaje de *Daubert* sobre la necesidad de una efectiva validez científica de la prueba tiene un impacto relevante, fundamentalmente en contextos en los que ese control no se hacía en el pasado, o solía hacerse de manera superficial o ineficiente.

Por supuesto, no debemos estar indebidamente influenciados por el «mito de la ciencia» o por un cientificismo *naïf*, dado que los epistemólogos nos enseñan a no creer que la Ciencia (con C mayúscula) siempre descubre Verdades (con V mayúscula), y nos dicen que el uso de la ciencia no es la panacea para todos los complejos problemas que el juzgador de los hechos debe resolver. Además, dicen que la ciencia se diversifica y evoluciona con el tiempo, de manera que la mera referencia a la ciencia no dice nada relevante. Y añaden que muchas veces no es fácil —o que incluso es imposible— trazar una distinción clara y precisa entre aquellos métodos científicamente válidos y los que no lo son, en la medida en que la diferencia entre «ciencia» y «no ciencia» es a menudo borrosa o incierta.

Sin embargo, estas críticas a un enfoque científico de los problemas de la prueba judicial, no deberían llevarse muy lejos, aun cuando indudablemente tienen algo de razón.

Por un lado, en efecto, aunque se reconozca que los límites de la ciencia (y de la no-ciencia) no están exacta y precisamente definidos, también debe admitirse que no todos los métodos que se dicen fiables son realmente capaces de suministrar un «buen» conocimiento de los hechos de la causa. En otras palabras, es necesario —tanto en el contexto judicial como en otros contextos— tener estándares o criterios para establecer si se puede obtener información y datos fiables, en lugar de información falsa y no fiable. Cualquiera que sea el estándar o paradigma que se use para definir lo que entendemos por «ciencia», debe racionalizarse y justificarse su aplicación, sólo para poder descartar el conocimiento que pretenda ser tal pero que realmente carezca de aceptabilidad racional.

Por otro lado, si no nos remitimos a la ciencia (porque en un caso en concreto no haya ciencia disponible, o porque por alguna razón no la usamos), estamos abocados al «sentido común», a la «cultura promedio» del «hombre razonable», es decir, al «acervo de conocimientos» personal que cualquier miembro de la sociedad tiene en su mente y su memoria. Y es que, el juzgador de los hechos (juez o jurado) debe decidir de todas maneras acerca de los hechos de la causa: si no puede confiar en pruebas científicamente válidas, se remitirá a su propio acervo de conocimientos. Pero William TWAINING tiene toda la razón cuando subraya que ese acervo de conocimientos es una «aglomeración no bien definida de creencias que consiste, típicamente, en una compleja sopa de información más o menos bien fundada, modelos sofisticados, recuerdos anecdóticos, impresiones, historias, mitos, proverbios, deseos, este-reotipos, especulaciones y prejuicios» y añade que en esta «sopa» la distinción entre hechos y valores no es muy clara, ni es clara tampoco la distinción entre hechos, fantasía y ficción.

Podría ser cierto, por tanto, que la ciencia puede no ser ontológicamente diferente de otros tipos de conocimiento más comunes, pero parece claro que

los distintos tipos de conocimiento tienen muy diversos grados de fiabilidad, y que ese nivel de fiabilidad es mucho mayor en la ciencia que en cualquier otra forma de conocimiento, real o pretendido.

1.1.

Un argumento crítico más común acerca de la conexión entre ciencia y derecho, y particularmente acerca del uso de la ciencia en el contexto judicial, es que los métodos de la ciencia y los métodos judiciales para la decisión sobre los hechos del caso son diferentes. Prima facie el argumento parece correcto, pero una vez más, no debe llevarse muy lejos.

Por una parte, es cierto que las investigaciones científicas están orientadas al descubrimiento de teorías generales o leyes, mientras que el proceso judicial está orientado a buscar la verdad acerca de eventos específicos y particulares; pero esto no es un argumento en contra del uso de la ciencia en contextos judiciales, pues en muchos casos se usan conocimientos científicos para decidir acerca de hechos concretos (tal como sucede, por ejemplo, con el diagnóstico médico de un paciente en particular). Por otra parte, las reglas procesales usualmente permiten la utilización de métodos científicos como algo común, y generalmente admiten que un experto le suministre al juzgador de los hechos el conocimiento científico o técnico que éste no tiene.

El hecho de que haya reglas que rigen la admisión y presentación (y, en algunos casos, también la valoración) de la prueba en los procesos judiciales puede ser un problema, pero ello no implica que esos procesos sean completamente diferentes de cualquier otro contexto de la experiencia en el que deba tomarse una decisión acerca de si un hecho realmente ocurrió o no. En otras palabras, no hay verdad «procesal» o «formal» que pueda alcanzarse en un proceso judicial, como cosa diferente de una verdad «real» o «material» que pueda ser alcanzada sólo en contextos no judiciales. Si un proceso está bien organizado como empresa dirigida a la búsqueda de la verdad de los hechos del caso, la verdad que puede alcanzarse en ese proceso no es ontológicamente diferente de la que puede alcanzarse en otro tipo de situaciones.

Aquí el problema no tiene que ver con la naturaleza de la verdad, sino con el hecho de que, por diversas razones, muchos sistemas procesales tienen reglas que son claramente antiepiestémicas, como las que excluyen pruebas relevantes (por ejemplo, con el propósito de proteger diversos secretos acerca de cuestiones privilegiadas). Sin embargo, debe tenerse en cuenta que esas reglas varían de un sistema procesal a otro y que son el resultado de la historia, las tradiciones y ciertas elecciones políticas y, por tanto, no son lógicas o epistémicamente necesarias y muchas veces carecen de justificación racional. Algunas veces, además, esas reglas son la expresión de diversos valores (como

la eficacia procesal, la economía de los recursos, la protección de derechos individuales) que están en conflicto con el propósito de la búsqueda de la verdad en el contexto judicial. No obstante, debe considerarse que la corrección de la decisión de los hechos de la causa es un valor básico en la administración de justicia, pues —como dijo Jerome FRANK alguna vez— una decisión justa no puede basarse en hechos equivocados. Entonces la cuestión no es abandonar los valores de corrección y verdad de la decisión fáctica sino —más bien— establecer cuándo y bajo qué condiciones tales valores podrían verse superados por otros valores y necesidades. Por supuesto, esa sería una decisión tremendamente difícil, si se tiene en cuenta que una reconstrucción verdadera de los hechos es una condición necesaria (aunque no suficiente) de la justicia y la legitimidad de la decisión final. En pocas palabras: las reglas procesales antiepiestémicas pueden aceptarse y estar justificadas, pero sólo cuando se considere que esas reglas están dirigidas a implementar un valor que se estima más importante que la decisión correcta y verdadera de los hechos del caso, esto es, un valor que se considera más importante que la correcta y legítima administración de justicia.

Sobre este punto, probablemente estaba en lo cierto Jeremy BENTHAM cuando escribió que un derecho probatorio ideal consistiría únicamente en el principio de relevancia inclusiva, de acuerdo con el cual debe admitirse todo elemento probatorio relevante (hoy, probablemente BENTHAM aceptaría la exclusión de la prueba ilícitamente obtenida). De otro lado, debe reconocerse que algunas reglas procesales son epistémicas porque se orientan a la búsqueda de la verdad, como por ejemplo las dirigidas a proteger al jurado de sus propios errores, o las dirigidas a suministrar un método eficiente para la presentación de la prueba oral (como en el interrogatorio cruzado, si se acepta el conocido *dictum* de WIGMORE).

Debe quedar claro, por tanto, que la existencia de reglas procesales de la prueba judicial no contradice de ninguna manera, per se, la utilidad y la admisibilidad de la prueba científica. Por el contrario, si hay reglas procesales antiepiestémicas, la consecuencia debe ser que se prescindirá de esas reglas; mientras que las reglas procesales epistémicas pueden facilitar la presentación de pruebas científicamente válidas y, por tanto, pueden hacer más probable el establecimiento de la verdad de los hechos del caso.

2.

Al menos desde finales del siglo XIX, algunos filósofos (como DILTHEY y WINDELBAND) comenzaron a introducir importantes distinciones dentro del ámbito general de la «ciencia». De acuerdo con la más reconocida de esas distinciones, hay una diferencia entre las llamadas «ciencias de la explicación» (es decir, las ciencias «naturales» o «duras»), como por ejemplo la fi-

sica, la química, la ingeniería, la genética, etc., y las llamadas «ciencias de la comprensión» (es decir, las ciencias «sociales», «humanas» o «blandas»), como por ejemplo la psicología, la sociología, la historia, la antropología, la etnología, etc. Este no es un asunto cultural y filosófico abstracto: también tiene efectos significativos, en la medida en que tiene que ver con el uso de las ciencias en los contextos judiciales.

Realmente, el uso de expertos de las ciencias sociales para establecer o evaluar e interpretar hechos específicos en casos concretos, ha sido reconocido desde hace algunas décadas, y se está haciendo cada vez más y más frecuente, no sólo en casos tan famosos como el conocido proceso en el que un juez inglés debía decidir acerca del valor literario de la novela *El amante de Lady Charterley* de LAWRENCE. Sólo por plantear unos cuantos ejemplos, podría considerarse que:

— Se consulta frecuentemente a expertos en *sociología* en casos sobre discriminación laboral.

— Se consulta a *historiadores* acerca de la reconstrucción de eventos históricos.

— Se escucha a los *historiadores del arte* cuando se debe apreciar la importancia histórica de un edificio.

— Se necesitan *antropólogos* para establecer si un delito tuvo «motivación cultural».

— Se consulta a los *psicólogos* en un amplio espectro de casos relativos al estado mental de las personas.

En términos generales, se admite la referencia a las ciencias sociales y humanas en los procedimientos judiciales, pero hay algunos problemas relevantes relacionados con esa práctica.

Por una parte, debe tenerse en cuenta que dichas ciencias no pueden ser evaluadas de acuerdo con los estándares *Daubert*. *Daubert* se deriva de una cuestión relacionada con las ciencias duras, dado que el problema era establecer si un fármaco llamado Bendectin podría haber provocado daños a un recién nacido. Por tanto, cuando el magistrado Blackmun se aventuró a definir los estándares antes citados, tenía en mente el problema de la validez científica de las ciencias naturales y nada más: por tanto, esos estándares fueron concebidos concibiendo un concepto general de ciencias duras. Ni siquiera cuando se extendió la aplicación de los estándares *Daubert* a ámbitos «técnicos» —como en *Khuno*— se pensó en las ciencias sociales.

Entonces la cuestión es que algunos de esos estándares podrían ser aplicados también a las ciencias sociales (principalmente el estándar de la «aceptación general», que ya fue aplicado en *Frye* y ha sido usado comúnmente desde 1923, y tal vez también el de la publicación en revistas arbitradas), pero ello no es posible para los demás —y más importantes— estándares, como el

haber sido sometido a prueba y el conocimiento de la tasa de error. El núcleo de la cuestión es que las «ciencias de la comprensión» no se corresponden con las «ciencias de la explicación» a las que se refería el juez Blackmun. Para decirlo en términos de KUHIN, hay una clara diferencia entre los *paradigmas* de ambos grupos de ciencias: las ciencias humanas, en particular, no usan el paradigma nomológico hempeliano, que se dice propio de las ciencias naturales. Dado que los estándares *Daubert*, como cualquier otro estándar aplicable a las ciencias duras, no puede ser aplicado a las ciencias con paradigmas fundamentalmente diferentes, estamos frente al problema de establecer qué estándares de validez científica, si es que hay alguno, pueden aplicarse a las ciencias sociales.

Un importante problema adicional es que esas ciencias no forman un conjunto homogéneo: cada una de ellas tiene su propio paradigma (o paradigmas). Dentro de semejante variedad, algunos criterios de validez científica deben definirse simplemente tomando en consideración las características específicas de cada ciencia. Ciertamente, una referencia de las estadísticas económicas o sociales no sigue el mismo método que debe emplearse para establecer el valor histórico de un edificio, y a su vez éste no es el mismo método que se aplica para analizar la capacidad mental de una persona, o de un menor, lo cual —una vez más— no se corresponde con cómo un experto evalúa la motivación cultural de un delito o establece si y cómo ocurrió un hecho histórico. Entonces el problema es determinar no uno, sino varios estándares de fiabilidad científica para cada tipo de ciencia que se tome en consideración.

Al afrontar este complejo problema, existe el riesgo omnipresente de admitir el uso judicial de métodos que podrían ser adecuadamente considerados como «ciencia basura» porque carecen de toda validez científica, y porque no hay ningún análisis epistémico acerca de su naturaleza y de la credibilidad de sus resultados. Aun admitiendo, como dicen algunos epistemólogos, que las fronteras entre la buena ciencia y la mala ciencia no son claras, no deberíamos intentar adscribirle valor científico alguno a la lectura de hojas de té o a la astrología, ni debería admitirse como perito a quien se presente como médium. Por otro lado, un buen método para probar la fiabilidad de una práctica nunca puede consistir en preguntárselo a quienes ganan dinero con ella. Incluso la referencia al test *Frye* de la aceptación general es discutible, simplemente porque la gente que practica la interpretación de las hojas de té estaría de acuerdo en que es un método de conocimiento ampliamente aceptado en la comunidad relevante. Lo que se necesita, por el contrario, es un análisis independiente de la construcción de estándares que puedan ser apropiados para cada paradigma de formas válidas de conocimiento, y debería ser capaz de detectar qué prácticas o métodos son «basura» por carecer de todo control epistémico.

3.

En un artículo reciente, David KAYE escribió que «estos son días oscuros para las ciencias forenses». Y la principal razón es que un informe de la National Academy of Sciences (NAS) de Estados Unidos, publicado en 2009 con el título *Strengthening Forensic Science in the United States. A Path Forward*, planteó en más de trescientas páginas los problemas de la fundamentación científica de las llamadas «ciencias forenses». El problema no es nuevo, dado que en las últimas dos décadas la literatura ha cuestionado ampliamente la fiabilidad de diversas técnicas forenses. Pero el informe de la NAS es especialmente importante porque examina no menos de doce tipos de pruebas forenses y, además de una discusión general de los problemas relativos a esas técnicas, desarrolla un amplio y detallado análisis para cada uno de ellos. Este informe es un punto de inflexión en el estudio de las pruebas forenses, y desde su aparición hasta hoy, se han publicado docenas de artículos discutiendo su contenido, tanto en términos generales como en cuanto a señalamientos concretos acerca de las prácticas forenses más comunes.

Por supuesto, el contenido del informe de la NAS no puede ser resumido aquí, pero vale la pena por lo menos resaltar los problemas más importantes que aborda. El núcleo fundamental del análisis desarrollado en el informe es que *todas* las ciencias forenses, con la única excepción de los exámenes de ADN, carecen de validez científica. Se dice lo mismo de las técnicas usadas con mayor frecuencia, como la comparación de huellas dactilares (usada habitualmente en Estados Unidos desde 1911) y el análisis grafológico, así como de otras técnicas usadas en la investigación forense. Además, el informe observa que los jueces estadounidenses han sido (y siguen siendo) «totalmente ineficientes» en el control de la fiabilidad de las pruebas forenses, pues admiten regularmente, y usan en sus decisiones, información o datos que no tienen ninguna validez científica. También dice que esta situación se deriva de una carencia sistemática de investigación científica acerca de la posibilidad (o imposibilidad) de validación de las técnicas forenses desde un punto de vista realmente científico, y de la amplia ignorancia de los jueces acerca de la no fiabilidad de casi todas esas técnicas. El informe NAS se refiere a la situación estadounidense, pero lo que dice en términos generales probablemente sea cierto también en muchos otros países.

Los exámenes de ADN son la única excepción relevante y en las últimas décadas se ha convertido en el «estándar de oro» de la fiabilidad científica, simplemente porque se apoyan en un cuerpo de investigaciones científicas que existen como base de los análisis genéticos (aunque el desarrollo de este tipo de pruebas no ha sido de ninguna manera lineal y sencillo, como lo muestra el reciente libro de David KAYE sobre el tema). Además, los exámenes de ADN han determinado la crisis de muchas otras técnicas, ahora es claro que cientos

o incluso miles de casos de falsas condenas (cuya falsedad se descubrió mediante exámenes de ADN) estaban basadas en diversas técnicas forenses que llevaron a los juzgadores de los hechos a decisiones equivocadas.

Merecen mencionarse al menos dos de los diversos problemas relativos a la fiabilidad de buena parte de las pruebas forenses. El primero se deriva del hecho de que en muchos casos (por ejemplo, el análisis grafológico), el resultado se basa esencialmente en la *comparación subjetiva* de dos elementos de prueba. Si la comparación es necesariamente *subjetiva*, entonces existe la posibilidad de que dos o más expertos lleven a cabo diferentes comparaciones subjetivas, con resultados sustancialmente diferentes: pero en esos casos no hay forma de establecer cuál de las comparaciones es la correcta y cuál no lo es, simplemente porque ambas son subjetivas y no se apoyan en ninguna base científica objetiva. Cada experto dirá que su propio análisis es el correcto y que el otro no lo es, pero esto seguirá básicamente en la incertidumbre.

El segundo problema surge porque el resultado de muchas técnicas forenses se expresa usualmente en términos de *unicidad* (por ejemplo, sólo una persona en todo el mundo puede tener esas huellas dactilares), pues los expertos forenses suelen hablar en términos de «tasa de error cero», o en términos de probabilidad de 1 en X millones (lo que en la práctica significa una tasa de error de casi cero). Sin embargo, en muchos casos la unicidad es imposible de probar, porque no hay base científica que justifique esa conclusión. Si, de un lado, se admite (aunque usualmente se desconoce) una probabilidad de error teóricamente baja (o extremadamente baja o de no unicidad), entonces la cuestión que se plantea es determinar qué tasa (desconocida) de error es aceptable como base para concluir a favor de una condena o de una absolución. Una vez más, la falta de información científicamente fiable priva a las técnicas forenses de cualquier credibilidad objetiva y controlable.

Al tener esos problemas en cuenta, el informe NAS hace una larga lista de recomendaciones, entre las cuales la Recomendación 3 es particularmente clara al enfatizar la necesidad de estudios «que establezcan las bases científicas que demuestren la validez científica de los métodos forenses» y sugerir el desarrollo y establecimiento de «medidas cuantificables de fiabilidad y precisión de los análisis forenses», determinando además los límites de la confiabilidad y precisión de esos análisis. Además, la Recomendación 5 resalta la conveniencia de investigaciones sobre «los prejuicios del observador humano y las fuentes del error humano en el examen forense», mientras que la Recomendación 6 esboza la necesidad de «desarrollar herramientas para las mediciones avanzadas, la validación, la fiabilidad, el intercambio de información y los exámenes de competencia», así como la necesidad de «establecer protocolos de análisis forense, métodos y prácticas». No es necesario señalar que dichas recomendaciones son particularmente importantes cuando se describen y proponen nuevos métodos de prueba forense: si no cumplen con los estándares de

validez científica, no deberían admitirse como prueba en los procedimientos judiciales.

4.

Estas breves anotaciones muestran que la aplicación de estándares de validez científica a los diversos tipos de prueba que se emplean en el contexto judicial es una empresa sumamente compleja y confusa. Sin embargo, y a pesar de su dificultad, es una empresa necesaria: si creemos que una decisión judicial correcta requiere de una reconstrucción verdadera y precisa de los hechos del caso, y si consideramos que la ciencia está convirtiéndose en un medio común para establecer esos hechos, entonces la determinación de qué método es realmente «científico» y cuál no es científicamente fiable, es una condición crucial para una correcta administración de justicia. Como he tratado de explicar, hay muchas diferencias importantes dentro del amplio y heterogéneo universo de la prueba científica, pero el problema de la validez científica es sumamente serio para las ciencias sociales, para las ciencias naturales y para las ciencias forenses.

Realmente, en la administración de justicia moderna, que se basa en la correcta aplicación del derecho, no se necesitan decisiones arbitrarias y no fiables de los hechos del caso. Se necesitan decisiones racionales y verdaderas, esto significa que cuando las decisiones estén basadas en la ciencia, debe tratarse de ciencia válida y buena.

BIBLIOGRAFÍA

- ALLEN, R. J., y NAFISI, E., 2010: «*Daubert* and its Discontents», *Brooklyn Law Review*, 50.
- BORUTTI, S., 1999: *Filosofia delle scienze umane. Le categorie dell'Antropologia e della Sociologia*. Milano: Bruno Mondadori.
- FAGNAND, D. L.; SACKS, M. J.; SANDERS, J., y CHENG, E. K., 2008: *Modern Scientific Evidence. Standards, Statistics and Research Methods*, ThomsonWest.
- GRANNELLI, P. C., 2011: «*Daubert* and Forensic Science: The Pitfalls of Law Enforcement Control of Scientific Research», *University of Illinois Law Review*, 2011.
- HACK, S., 2009: «Irreconcilable Differences? The Troubled Marriage of Science and Law», *Law and Contemporary Problems*, 72.
- 2010a: «Federal Philosophy of Science: A Deconstruction - And a Reconstruction», *NYU Journal of Law and Liberty*, 5.
- 2010b: «Seis signos de cientismo», *Discusiones filosóficas*, 2010.
- KAYE, D. H., 2010a: «Probability, Individualization and Uniqueness in Forensic Science Evidence: Listening to the Academics», *Brooklyn Law Review*, 75.
- 2010b: *The Double Helix and the Law of Evidence*, Cambridge (Mass.)-London: Harvard University Press.

- LUKE, A., 2010: «Employment Discrimination Litigation: Social Science Evidence and a Solution for the Problem of Presumptions», *Temple Journal of Science, Technology and Law*, 29.
- MNOOKIN, J. L.: «The Courts, the NAS and the Future of Forensic Science», *Brooklyn Law Review*, 75.
- MORGAN, H., 2009: «Painting the Past and Paying for It: the Demise of *Daubert* in the Context of Historical Expert Witnesses», *Wake Forest Law Review*, 44.
- Reference Manual on Scientific Evidence*, 2000: 2.ª ed., Federal Judicial Center.
- SHELTON, D. E., 2011: *Forensic Science in Court. Challenges in the Twenty-First Century*, Lanham: Rowman & Littlefield.
- SLOBODIN, Ch., 2007: *Proving the Unprovable. The Role of Law, Science and Speculation in Adjudicating Culpability and Dangerousness*, Oxford: Oxford University Press.
- Strengthening Forensic Sciences in the United States. A Path Forward*, 2009. Washington D. C.: National Academy Press.
- TARUFFO, M., 2008: «La prueba científica», en *id.*, *La prueba*, Madrid-Barcelona-Buenos Aires: Marcial Pons.
- TWINING, W., 2006: *Rethinking Evidence. Exploratory Essays*, 2.ª ed., Cambridge: Cambridge University Press.

IX

UNA APROXIMACIÓN PRAGMATISTA AL TESTIMONIO COMO EVIDENCIA

Andrés PÁEZ *
Universidad de los Andes

1. INTRODUCCIÓN

El testimonio es nuestra mayor fuente de creencias. La gran mayoría de las cosas que creemos han sido adquiridas a partir de las palabras de los demás, no de evidencia recolectada directamente en el mundo a través de la observación. Hechos acerca del pasado remoto o acerca de lugares lejanos son claros ejemplos de creencias testimoniales. Pero también lo son nuestras creencias acerca de hechos muy personales, tales como el lugar y la fecha de nuestro nacimiento, o los nombres de nuestros tatarabuelos. Las teorías y los descubrimientos científicos son aceptados en la gran mayoría de los casos a partir de lo que otros nos informan. Que el cigarrillo causa cáncer, o que las anomalías en las observaciones de las rotaciones de las galaxias pueden ser explicadas por la teoría de la materia oscura, son afirmaciones que ninguna persona puede verificar individualmente. En muchos casos ni siquiera puedo imaginar o entender los medios a través de los cuales mis

* Profesor asociado y director del Departamento de Filosofía de la Universidad de los Andes, Bogotá, Colombia. El autor quiere agradecer a todos los participantes del *Workshop on Standards of Proof and Scientific Evidence*, organizado por la Facultad de Dret de la Universitat de Girona, por sus valiosos aportes y comentarios. Esta publicación es uno de los resultados de la investigación: «Umbral de Evidencia Científica», financiada por el Departamento Administrativo de Ciencia, Tecnología e Innovación de la República de Colombia (COLCIENCIAS).

creencias testimoniales podrían ser verificadas, y sin embargo tengo plena confianza en ellas.

La mayoría de los testimonios son aceptados irreflexivamente, sin que medie un proceso de deliberación consciente. La credulidad humana es un tema filosófico con una larga tradición. Por ejemplo, REID afirmaba que todos los seres humanos poseen «la disposición a confiar en la veracidad de los demás, y a creer lo que nos dicen» (2000: 194). Nuestra credulidad probablemente está apuntalada por mecanismos subconscientes muy estables que sobrepasan la plausibilidad de las afirmaciones de los demás (FRICKER, 1995), su consistencia con las creencias del oyente y su credibilidad intrínseca y contextual (THAGARD, 2005). Estos mecanismos psicológicos de «vigilancia epistémica» (SPERBER, 2010) parecen necesarios desde un punto de vista evolutivo. Nuestra confianza en las palabras de los demás son un requisito para la interacción social en general, y es esencial con el fin de alcanzar el nivel de cooperación necesario para la supervivencia de la especie humana (STREBLIN, 2012).

Sin embargo, nuestra tendencia natural a aceptar testimonios algunas veces nos mete en problemas, especialmente cuando la aceptación desprevenida de un falso testimonio tiene consecuencias negativas significativas. En los negocios, el derecho, la ciencia y en la vida diaria debemos tomar decisiones importantes e implementar políticas cruciales a partir, casi exclusivamente, de las palabras de los demás. En estos contextos el testimonio ha de ser aceptado como el resultado de un proceso de decisión racional. La aceptación del testimonio requiere una evaluación epistémica para determinar su relevancia, su credibilidad y su valor explicativo y evidencial en un contexto dado.

Es bien conocido el intento de HUME de justificar nuestra aceptación del testimonio reduciéndolo a una simple inducción por enumeración. En «De los milagros», que corresponde al capítulo X de la *Investigación sobre el entendimiento humano* (1743), HUME afirma que podemos confiar en el testimonio humano porque la experiencia ha demostrado que generalmente es confiable. La aproximación reduccionista de HUME no ha encontrado muchos defensores y ha sido criticada severamente por COADY (1973, 1992), como veremos en el apartado 2.

La gran mayoría de los autores ha optado por aproximaciones no reductivas, en las cuales no se hace ningún intento por verificar o justificar directamente nuestras creencias testimoniales. SHAPIN (1984), por ejemplo, intenta transformar la pregunta de tal modo que lo que esté en juego no sea la justificación de una creencia sino la determinación del grado de confianza que nos produce la persona que presenta el testimonio. Como nuestro juicio acerca de si alguien es digno de confianza es un juicio moral, el problema del testimonio en el fondo sería un problema moral. El problema es que las fuentes del falso testimonio incluyen no sólo las intenciones mendaces, sino también la

incompetencia honesta. La aproximación de SHAPIN no logra incluir este último aspecto del testimonio.

HARDWIG (1985: 1991), por su parte, también arguye que la confianza es la base de la autoridad epistémica que sirve de base para el avance de la ciencia. Su argumento es que un científico no puede ser independiente y autónomo, ni siquiera en su propia área de especialización, si quiere tener acceso a la mejor evidencia disponible en su campo. En su opinión, un investigador puede tener buenas razones para creer una afirmación testimonial si tiene buenas razones para creer que la fuente tiene buenas razones para creerlo. En otras palabras, puede aceptar un testimonio, no porque sepa cómo verificarlo o porque tenga la posibilidad de hacerlo, sino porque cree que la fuente es un experto en su campo. A mi modo de ver, sin embargo, la propuesta de HARDWIG no es muy satisfactoria, pues consiste en reemplazar un problema por otro: el problema de las creencias testimoniales por el problema del conocimiento experto, el cual es tanto o más difícil de resolver que el problema original.

En este ensayo presento una aproximación al problema del testimonio desde una nueva perspectiva reduccionista. Mi intención es mostrar que las creencias testimoniales fundamentadas pueden estar basadas en las mismas reglas de inferencia y en los mismos mecanismos de adquisición de creencias que las creencias no testimoniales. Al mismo tiempo, mi aproximación preserva los rasgos característicos de la evidencia testimonial. La idea fundamental es que la aceptación deliberada de un testimonio es un proceso que debe hacerse en dos etapas. En la primera, se busca una justificación de la afirmación testimonial a través de la determinación de aquellos hechos que aumentan la probabilidad epistémica de que la afirmación sea verdadera, esto es, de que la fuente esté reportando verazmente un hecho. El conjunto de hechos relevantes puede ser obtenido a partir del contenido del testimonio y del contexto en el que es presentado, incluyendo la identidad de la fuente o del testigo. La aceptación de estos hechos y relaciones probabilísticas busca un equilibrio entre su credibilidad y su valor informativo. En la segunda etapa, se juzga si la justificación obtenida es suficiente para aceptar el testimonio como evidencia en un contexto dado, de nuevo intentando equilibrar su credibilidad y su valor informativo a la luz de los intereses y metas cognitivas del investigador.

La estructura del ensayo es la siguiente. En los siguientes dos apartados presentaré las objeciones principales en contra de una teoría reduccionista del testimonio, prestando particular atención a la acusación de que cualquier intento de justificar un testimonio es irremediablemente circular. En los apartados restantes presentaré mi propia aproximación al problema del testimonio. Desde el comienzo debe quedar claro que el siguiente análisis no pretende dar cuenta del gran número de complicaciones que surgen al estudiar las pruebas testimoniales en el contexto jurídico. Mi propósito es ofrecer más bien un análisis epistemológico de la evidencia testimonial que

quizá pueda iluminar indirectamente el problema del testimonio en contextos jurídicos.

2. LOS PROBLEMAS DE LAS APROXIMACIONES REDUCTIVAS AL TESTIMONIO

En las páginas iniciales de «De los milagros», HUME anota que «no hay un tipo de razonamiento más común, más útil o incluso más necesario para la vida humana que el derivado de los testimonios de los hombres y los informes de los testigos presenciales y de los espectadores» (1748: 135). Al indagar sobre la fuente de la confianza que tenemos en los testimonios de los demás, afirma que «nuestra seguridad en cualquier argumento de esta clase, no deriva de ningún otro principio que la observación de la veracidad del testimonio humano y de la habitual conformidad de los hechos con los informes de los testigos» (135-136). HUME parece estar diciendo que si alguien hace una afirmación del tipo X, a través del cual reporta un estado de cosas, creemos la afirmación porque hemos verificado en el pasado que las afirmaciones del tipo X generalmente corresponden al estado de cosas que describen. La evidencia testimonial se reduce entonces a una simple inducción por enumeración.

COADY encuentra dos defectos en esta propuesta: «HUME nos dice que confiamos en el testimonio sólo porque la experiencia nos ha mostrado que éste es confiable, pero si por "experiencia" quiere decir la observación individual y las expectativas que ésta genera, la afirmación es claramente falsa; por otra parte, si se refiere a la experiencia común (*i. e.*, a la que depende de las observaciones de los demás), entonces con toda seguridad es una petición de principio» (1973: 150).

Como lo señala COADY, la evidencia textual parece indicar que HUME tenía en mente la experiencia común de la humanidad, no las observaciones de un individuo, como base para la inferencia testimonial. Ahora bien, como sólo podemos saber lo que los demás han observado a través de su testimonio, la propuesta de HUME es claramente circular. Si circunscribimos la teoría de HUME a la esfera de la observación individual, podremos encontrar algunos ejemplos en los que parece funcionar bien, pero éstos se limitarían a los testimonios de nuestra familia y amigos, y a situaciones cotidianas en las que los eventos reportados pueden ser verificados directamente. En la gran mayoría de los casos, sin embargo, no existe la oportunidad de verificar si un testimonio es confiable y por eso considero que COADY tiene razón al concluir que nuestras inferencias testimoniales carecen del tipo de justificación que HUME les asigna.

COADY también señala que HUME no ofrece ninguna pista acerca de cómo debemos clasificar los tipos de testimonio para los cuales buscamos una justi-

ficación. «Hume pretende que estas conjunciones sean similares a las clases de conjunciones que él cree son necesarias para establecer leyes causales e incluso leyes de la naturaleza. En tales casos, las conjunciones constantes decisivas son entre una clase de objeto y otra clase de objeto. Pero independientemente de lo que pensemos sobre la idea de clases de objetos, la noción de clase de reporte claramente requiere alguna explicación en este contexto» (51).

Hay dos formas obvias en que se puede clasificar un reporte: por el tipo de declarante o testigo, y por el tipo de contenido. En el primer caso, los tipos relevantes estarían determinados por el área de expertise: meteorólogo, historiador del arte, bioquímico, etc. Con el fin de determinar si un declarante pertenece a alguna de estas áreas, debemos hacer uso de información adicional, la cual generalmente no es directamente observable y debe ser obtenida del testimonio de otros. Terminamos así una vez más en el problema de la circularidad. Por otra parte, si intentamos determinar el tipo de testimonio por el tipo de contenido, nos enfrentamos al siguiente problema: un mismo contenido puede ser clasificado de muchas maneras diferentes, dificultando la tarea de establecer una única inferencia inductiva. «Presumiblemente habría que decidir si el reporte "Hay un león enfermo en el Parque Zoológico de Taronga" es un reporte de tipo médico o geográfico o empírico o existencial» (52). Además, dicha decisión también dependería del testimonio de otros¹.

3. APROXIMACIONES REGLA-REDUCTIVAS AL TESTIMONIO

Como lo señala Peter LIPTON, la descripción que hace HUME de nuestras prácticas inductivas es demasiado rudimentaria y no representa adecuadamente los diferentes aspectos de nuestras inferencias no demostrativas. «Esta debilidad de la teoría general de Hume acerca de nuestras prácticas inductivas también podría generar una debilidad en la argumentación de Coady, porque me parece que Coady se apresura al pasar del fracaso del intento de Hume de reducir el testimonio a una primitiva inducción por enumeración a la conclusión de que no existe ninguna teoría de la justificación inductiva que pueda incluir inferencias sobre la confiabilidad del testimonio» (1998: 17).

LIPTON arguye que puede haber dos aproximaciones reduccionistas posibles al testimonio: una aproximación premisa-reductiva y una regla-reductiva. «La primera es un intento de mostrar cómo cada creencia testimonial que el agente acepta de manera fundamentada puede ser justificada en términos que no apelen ellas mismas a creencias basadas en testimonios» (23). Éste es el tipo de aproximación que COADY rechaza y es claro que es un proyecto sin

¹ COADY ofrece una tercera objeción que consiste en una reducción al absurdo del argumento de HUME. Esta objeción no es muy relevante para el propósito de este ensayo y la dejaré de lado. Más bien me interesa profundizar en el problema de la circularidad en la justificación del testimonio.

futuro dada la precaria base empírica sobre la cual descansaría cualquier inferencia testimonial inductiva. Y no sólo sería un proyecto sin perspectivas de éxito sino también uno indeseable, puesto que no le haría justicia a nuestras prácticas científicas y cotidianas, en las cuales muchos testimonios que no han sido verificados directamente por la observación son utilizados en la justificación de todo tipo de afirmaciones.

LIPTRON cree que una aproximación regla-reductiva tiene mejores perspectivas. Dicha aproximación mostraría «que las creencias testimoniales fundamentadas están basadas en reglas de inferencia o mecanismos de adquisición de creencias que son válidas para creencias de diversas fuentes, no solamente para las provenientes del testimonio» (24). La plausibilidad de este tipo de aproximación se deriva de la idea de que existe no sólo una dimensión vertical sino también una horizontal en la justificación de las creencias en general, una dimensión en la que la coherencia, la cohesión y la integración explicativa juegan un papel fundamental, y en la cual las creencias testimoniales se combinan libremente con creencias de otras fuentes como la percepción o la memoria. LIPTRON concluye: «Este mecanismo holista es una fuerte indicación de que, en la medida en que la adquisición de creencias a partir del testimonio es inferencial, las reglas de inferencia no son particulares al testimonio» (24).

El mecanismo inferencial adoptado por LIPTRON es la Inferencia a la Mejor Explicación (IME). Una exposición detallada de la IME puede encontrarse en su libro del mismo nombre (2004). De acuerdo con la teoría de LIPTRON, siempre debemos inferir aquella hipótesis que proporcione la mejor explicación de la evidencia disponible. «En el caso del testimonio, la aplicación más sencilla de la Inferencia a la Mejor Explicación sería decir que el agente infiere que lo dicho por el informante es verdadero sólo si la verdad de lo dicho es (parte de) la mejor explicación de (entre otras cosas) el hecho de que el informante lo haya dicho» (1998: 27). El tipo de explicación que LIPTRON tiene en mente en el caso del testimonio es la explicación psicológica de una clase de comportamientos verbales en los cuales los juicios acerca de la sinceridad y la competencia del hablante juegan un papel central. «Cuando explicamos por qué una persona dice algo, la explicación no tiene por qué depender de la determinación previa de la verdad de lo dicho. Así, podemos juzgar que la mejor explicación de por qué alguien dijo lo que dijo es que esa persona probablemente no está siendo engañada ni está engañándose en esa clase de asunto, sin tener que establecer independientemente si lo que está diciendo es verdad» (29).

LIPTRON no nos da más detalles acerca de la aplicación de la IME al caso del testimonio. Simplemente nos asegura que hay dos razones por las que la IME

es adecuada para proveernos con una epistemología del testimonio. Por una parte, las características de las inferencias testimoniales son comunes a las inferencias inductivas en general. En particular, la IME aprueba las inferencias verticales y enfatiza el papel de la coherencia. Por otra parte, la IME da cuenta de dos características que sólo le pertenecen a las inferencias testimoniales: la mendacidad y la incompetencia honesta.

Un análisis completo de la IME en general, y de cómo funcionaría en el caso del testimonio, nos desviaría de nuestro objetivo principal en este ensayo. Sólo quiero señalar por qué considero que la IME no es adecuada para dar cuenta de las inferencias testimoniales. El problema principal es que lo que se requiere en el caso de las creencias testimoniales no es una explicación sino una justificación. La búsqueda de una explicación nos lleva en la dirección equivocada. Consideremos por un momento el *explandium* de una IME y la pregunta por el porqué que pone en marcha la inferencia. El hecho que ha de ser explicado es que un individuo *a* dijo que *p*, y la pregunta que nos debemos formular es «¿Por qué dijo *a* que *p*?». La mejor explicación de por qué *a* dijo que *p* naturalmente incluirá información muy dispar. En su análisis del testimonio, GOLDMAN (1999: 104) considera la variedad de elementos involucrados:

Primero, el comunicador debe seleccionar cuál de los hechos observados comunicará. Si acaba de observar diez verdades pero no es factible comunicarlas todas, debe decidir qué subconjunto reportará. Segundo, para cada una de las verdades observadas, existe la opción de reportarla sinceramente *versus* la opción de distorsionarla o falsificarla. La tercera dimensión es «cómo» comunicarla. Puede haber más de un medio de comunicación disponible y los medios pueden diferir, por ejemplo, en términos de su probabilidad de llegar a la audiencia deseada. Así que hay que tomar decisiones. Finalmente, en la dimensión del «a quién», el observador debe elegir la audiencia a la cual se dirigen los mensajes seleccionados. ¿A qué individuos se deben dirigir los diversos mensajes? ¿Hay potenciales receptores que el reportero quiera impedir que reciban los mensajes? Las decisiones en las cuatro dimensiones pueden ser interdependientes. El medio o la manera de transmisión escogido pueden constreñir el número de mensajes que se pueden enviar. La audiencia escogida para los mensajes puede influenciar su contenido, y viceversa.

Todos estos aspectos forman parte de la explicación psicológica completa de por qué *a* dijo que *p*. Por muy interesantes que sean, la mayoría de estas explicaciones parciales son de poca utilidad para alguien que esté contemplando la aceptación del testimonio de *a* y que busque algún tipo de justificación para hacerlo. Por una parte, muchas de estas explicaciones no pueden ser verificadas porque hacen referencia a decisiones y procesos psicológicos inaccesibles al receptor del testimonio. Por otra, lo que el receptor quiere saber es, por ejemplo, si el testigo estaba ubicado en un lugar que le permitiera observar los hechos reportados, o si tiene el conocimiento o el entrenamiento necesario

² El fundarcentismo de HAACK (1993: 2003) y la teoría de la coherencia de THAGARD (2000) son aproximaciones que han explorado a fondo la naturaleza multidimensional de la justificación de las creencias.

para hacer las afirmaciones que hace, o si tiene buenos antecedentes reportando los hechos, etc. Ninguno de estos elementos pertenece a la explicación psicológica de por qué *a* dijo lo que dijo; pertenecen, más bien, a la justificación que el receptor puede ofrecer para aceptar el testimonio de *a*. En breve, la propuesta de utilizar la IME como base para la epistemología del testimonio fracasa porque las explicaciones generalmente son insuficientes para justificar una creencia testimonial³.

En este ensayo quiero ofrecer una aproximación completamente diferente al problema del testimonio. Conuerdo con LIPSON en que tal aproximación debe ser regla-reductiva, es decir, debe estar «basada en reglas de inferencia o mecanismos de adquisición de creencias que son válidas para creencias de diversas fuentes, no solamente para las provenientes del testimonio» (1998: 24), y también es fundamental que la aproximación haga justicia a las características específicas de las inferencias testimoniales. Sin embargo, en lugar de la IME, mi aproximación estará basada en la teoría de la revisión de creencias desarrollada por LEVI (1967; 1980; 1991)⁴, y en particular en su aproximación a la investigación⁵ como un problema de decisiones.

4. LA INVESTIGACIÓN COMO UN PROBLEMA DE DECISIONES

En *The Fixation of Belief and Its Undoing*, LEVI (1991) ofrece una teoría de la investigación como un problema de decisiones, inspirada en el modelo duda-creencia propuesto originalmente por PERCE (1877). De acuerdo con el modelo duda-creencia, un agente investigador presupone que todo aquello que se ha comprometido a creer plenamente es verdadero. Esto no significa que la verdad o la falsedad sean relativas a lo que el agente cree. Pero los juicios del agente acerca de la verdad y la falsedad son relativos a lo que actualmente cree.

Una investigación parte de un trasfondo de creencias firmemente mantenidas: el estado de creencia plena del agente, considerado por éste como libre de errores. Dos desideratas guían las decisiones de un investigador al modificar su actual estado de creencia: i) la adquisición de nueva información valiosa, y ii) la evitación del error. Dadas las metas cognitivas del agente, un cambio de creencia óptimo es aquel que provee información nueva y libre de errores.

³ Esto no quiere decir que la integración explicativa no juegue un papel en la aceptación de un testimonio. Como veremos en el apartado 5, el rol explicativo de un testimonio en una investigación debe ser tomado en cuenta al determinar su valor epistémico.

⁴ La teoría de revisión de creencias fue desarrollada de manera independiente, y con un enfoque distinto, por ALCHOURRÓN, GARDENFORS y MAKINSON (1985).

⁵ El término «investigación» debe ser entendido en sentido amplio, como cualquier actividad guiada por el intento de resolver una pregunta o problema práctico o teórico, no necesariamente científico o jurídico.

Sin embargo, la evitación del error y la adquisición de nueva información son metas que apuntan en direcciones opuestas. Por un lado, la mejor forma de evitar el error es permanecer en el estado original de creencia plena. Por el otro, la forma más efectiva de adquirir información es migrando hacia estados de creencia que pueden contener errores e inconsistencias. «La moraleja de la historia es que aquellos cambios que aumentan el valor informacional generan un riesgo de error» (LEVI, 1991: 12).

Con el fin de esclarecer la manera en que estos dos fines cognitivos opuestos pueden ser equilibrados, LEVI comienza explicando en qué consiste un estado de creencia plena. Cuando un agente investigador acepta conscientemente un conjunto de proposiciones se compromete con las consecuencias lógicas de esas proposiciones. El agente quizás no pueda reconocer todas las consecuencias de sus creencias en ese momento, pero no puede negarlas racionalmente si alguien se las señalara. Al aceptar conscientemente todas las consecuencias de sus creencias, el agente está cumpliendo con su *compromiso* doxástico. Pero debido a sus limitadas habilidades de razonamiento, su *desempeño* doxástico siempre se quedará corto con respecto a su compromiso doxástico. El conjunto de enunciados con los cuales se ha comprometido el agente en un momento dado es su estado de creencia plena, denotado por *K*.

De acuerdo con el modelo duda-creencia, una investigación comienza ante la presencia de la duda. «La duda es un estado de inquietud e insatisfacción del que luchamos por liberarnos para pasar a un estado de creencia; este último es un estado calmado y satisfactorio que no queremos evitar o cambiar por una creencia diferente [...] llamaré a esta lucha, *investigación*» (PERCE, 1877: 114). Por supuesto, no todas las dudas nos fuerzan a comenzar una investigación. Algunas dudas son consideradas más importantes que otras, y no existe ningún principio que gué nuestras investigaciones distinto a nuestros propios intereses y metas cognitivas.

Como mencioné al comienzo, la gran mayoría de nuestros testimonios son aceptados y creídos irreflexivamente, de manera automática. Los únicos requisitos parecen ser que el testimonio sea consistente con el conjunto de creencias del agente y que exista un compromiso previo con la confiabilidad del testigo o de la fuente de la información. Esta aceptación automática del testimonio corresponde a lo que LEVI llama una «expansión rutinaria» de *K* (1991: cap. 3). La búsqueda de una justificación racional para un testimonio, es decir, un expansión no rutinaria de *K*, sólo surge cuando el asunto es muy importante o interesante para el investigador a la luz de sus fines prácticos o cognitivos; o cuando su aceptación tiene consecuencias significativas para el agente o para otras personas; o simplemente cuando la curiosidad intelectual del agente no quede satisfecha hasta encontrar suficiente evidencia para una afirmación. Si una afirmación es tautológica, o si confirma algo que el agente ya cree, o si es irrelevante o inconsecuente, será incapaz de generar duda o

curiosidad alguna en el agente y éste no emprenderá la tarea de encontrar una justificación para ella.

La aproximación al testimonio que presento aquí toma como punto de partida un contexto epistémico en el cual un agente investigador cree que un reporte testimonial no está libre de duda y donde el agente considera que la tarea de alcanzar una creencia firme sobre dicho reporte es digna de ser emprendida a la luz de sus fines e intereses. Si un reporte testimonial es dudoso pero ninguna decisión o investigación depende de él, o si va a ser utilizado como evidencia en algún asunto de importancia pero no hay una duda razonable acerca de su verdad, el agente no comprometerá recursos cognitivos importantes buscando una justificación para él.

5. LA BÚSQUEDA DE UNA JUSTIFICACIÓN PARA EL TESTIMONIO

Una vez que se ha puesto en marcha una investigación, motivada por un testimonio dudoso, en mi opinión el agente se verá enfrentado a dos tareas diferentes. La primera será buscar una justificación para la afirmación testimonial, intentando encontrar todos aquellos hechos que aumenten la probabilidad de que la afirmación sea verdadera, esto es, de que el testigo o la fuente esté reportando verdicilmente un hecho. El conjunto de hechos relevantes puede ser extraído del contenido de la afirmación testimonial y del contexto en el que aparece, incluyendo la identidad del testigo o fuente. La evidencia no puede estar restringida a hechos observables. De otro modo, la aproximación se convertiría en premisa-reductiva. Es posible recopilar información factual a partir de lo dicho por otros y la valoración de la credibilidad de su testimonio formará parte del proceso de justificación. Naturalmente la afirmación testimonial que sirve como justificación no puede ser más dudosa que el testimonio que estamos evaluando. La aceptación de estos hechos y relaciones probabilísticas estará basada en el mejor equilibrio entre su credibilidad y su valor informativo relativos al conjunto K de creencias del agente.

La segunda tarea a la que se verá enfrentado el agente es la determinación de si la justificación obtenida en el primer paso es suficiente para la aceptación del testimonio en ese contexto. Esta segunda tarea será analizada en el apartado 6.

Para ilustrar el primer paso, voy a presentar una serie de ejemplos de testimonios en diferentes contextos y de algunos de los hechos que hacen más probables que sean verdaderos. Posteriormente presentaré un tratamiento formal de estas relaciones.

Ejemplo 1: Un testigo reporta que el auto utilizado por los ladrones de un banco en su huida era rojo. La siguiente es una lista parcial de los datos que aumentarían la probabilidad de que el testigo esté diciendo la verdad:

- El testigo estaba muy cerca del banco en el momento del robo.
- Las condiciones de iluminación permitían ver claramente el color del auto.
- El testigo es un maestro de escuela muy respetado.
- El testigo no es daltónico⁶.

Ejemplo 2: El Doctor G. ha publicado una serie de resultados experimentales en una revista muy respetada en su disciplina. Los siguientes hechos aumentan la probabilidad de que el reporte sea verdadero:

- El Dr. G. tiene un título doctoral en esa disciplina.
- El Dr. G. es profesor titular en una prestigiosa universidad.
- La revista somete todos los artículos a una revisión por pares.
- Otros investigadores independientes han replicado en el pasado los resultados experimentales obtenidos por el Dr. G.

En ambos ejemplos las creencias de fondo que se requieren para aceptar que la información mencionada aumenta la probabilidad de que el testigo esté diciendo la verdad no son muy sofisticadas. Si la información proviene de fuentes confiables, el agente puede aceptarla y utilizarla posteriormente para decidir si acepta el testimonio original, como veremos en el próximo apartado.

Ejemplo 3: Un desconocido se te acerca en la calle y te dice que el nadador Michael Phelps, el mayor medallista olímpico de todos los tiempos, se acaba de ahogar. Tu primera reacción es de incredulidad y decides investigar el asunto. Logras verificar que Michael Phelps tenía planeado utilizar la piscina local hoy, pero no logras averiguar nada más. Aunque este hecho aumenta ligeramente la probabilidad de que el reporte del desconocido sea verdadero, decides no creerlo.

Ejemplo 4: En un foro de discusión en internet acerca del valor de las acciones en bolsa de la compañía X aparecen contribuciones de inversionistas y especuladores. Los participantes intercambian información acerca de eventos que tienen el potencial de afectar el precio de la acción de la compañía. Como no hay información alguna acerca de la identidad, el contexto o la motivación de los participantes, es muy difícil darle credibilidad a los posibles factores que puedan aumentar la probabilidad de que sus testimonios sean verdaderos. Asumiendo que éstos no son intrínsecamente increíbles, la única estrategia disponible para aceptar cualquier afirmación es razonar inductivamente a la manera sugerida por HUMBE. Uno sólo debe creer las afirmaciones de aquellos participantes que regularmente reportan eventos que son posteriormente

⁶ Este sencillo ejemplo puede hacerse más complejo si incluimos la clase de hechos que son generalmente mencionados en el análisis psicológico de la confiabilidad de un testigo en el contexto jurídico. Véase, por ejemplo, WELLS y OLSON, 2003.

confirmados. Pero si no existe un patrón en la información proporcionada por alguien, sería irracional tomar cualquiera de estos testimonios como evidencia pues carecen de toda justificación⁷.

En todos estos casos es muy importante tener en cuenta que los hechos que aumentan la probabilidad de la verdad de cada testimonio no son necesariamente los mismos hechos que aumentan la probabilidad del estado de cosas reportado a través del testimonio. En el primer ejemplo, el hecho de que el testigo sea un maestro de escuela muy respetado ni aumenta ni disminuye la probabilidad de que el auto de los ladrones haya sido rojo, pero sí aumenta la probabilidad de que la persona haya sido honesta. Lo mismo ocurre con el hecho de que la revista del segundo ejemplo tenga una política de revisión por pares, un hecho que no tiene ninguna conexión causal en absoluto con los resultados experimentales obtenidos por el Dr. G. en su laboratorio.

THAGARD (2005: 298) identifica cuatro elementos que pueden generar dudas acerca de un testimonio: la falta de credibilidad de la fuente, el comportamiento sospechoso de la fuente, la incompatibilidad del testimonio con los fines prácticos o epistémicos, y la inconsistencia del testimonio con las creencias del agente. Ahora bien, de acuerdo con el modelo duda-creencia, si un testimonio es inconsistente con el estado de creencia K del agente, debe ser rechazado inmediatamente. Sin embargo, hay casos en los que el agente no quiere prejulgar negativamente el testimonio, y decide contraer hipotéticamente su estado de creencia, esto es, rechazar momentáneamente aquellas creencias que generan una contradicción con el testimonio⁸. En lo que sigue asumiré que el testimonio no es inconsistente con K , ya sea porque K no implica la negación del testimonio, o porque K ha sido contraído para eliminar hipotéticamente la inconsistencia. La actitud epistémica inicial del agente será la de suspender el juicio acerca de la verdad del testimonio.

Supongamos que K es el estado de creencia de una persona que está considerando aceptar un testimonio, ϕ es una oración que afirma que el testimonio es verdadero, y $p(\phi)$ es la probabilidad epistémica⁹ en K de que el testigo o la fuente están diciendo la verdad. Un conjunto de oraciones Ψ es una justificación potencial de ϕ con respecto a K si y sólo si:

- i) $K \cup \Psi$ es consistente.
- ii) $\Psi \not\subset K$.

⁷ Es interesante anotar que si las afirmaciones hechas en el foro de discusión en internet son acerca del precio de la acción, no podríamos utilizar la estrategia inductiva de HUME porque no hay una forma directa de verificar el precio. La única verificación posible es a través de un reporte oficial de la bolsa de valores.

⁸ Para un tratamiento formal de la contracción de un estado de creencia, véase LEVI, 2004, y GÄRBERG, 1998. Volveré a este asunto en el próximo apartado.

⁹ La probabilidad epistémica de una oración es el grado de confirmación o apoyo provisto por la evidencia total disponible, en este caso, por las oraciones en el conjunto de creencias K .

- iii) Hay una oración $\psi \in \Psi$.
- iv) Hay una oración $p(\phi|\psi) > p(\phi) \in \Psi$.
- v) No hay ninguna $\gamma \in K$ tal que $p(\phi|\psi \& \gamma) = p(\phi|\sim\psi \& \gamma)$.
- vi) ϕ y ψ son lógicamente independientes.
- vii) Nada más es un elemento de Ψ .

Una justificación potencial es, pues, un conjunto que contiene una oración singular ψ y una oración de probabilidad que afirma la relevancia probabilística del hecho descrito por ψ para el hecho de que el testimonio sea verdadero. A continuación examinaré cada una de estas condiciones.

La primera condición afirma que una justificación potencial debe ser consistente con el corpus de creencias con respecto al cual está siendo evaluado el testimonio. Es cierto que hay muchos testimonios que nos obligan a realizar cambios radicales en nuestras creencias. Sin embargo, el efecto de la aceptación de un testimonio sobre el estado de creencia es independiente de la aceptación de las creencias que aumentan la probabilidad de que ese testimonio sea verdadero. Supongamos que yo creo que Juan, quien trabaja para la compañía X , es el culpable de un desfalco. La compañía ha contratado a un detective que, trabajando en conjunto con la contadora, ha rastreado las cuentas bancarias de todos los empleados. Según el testimonio del detective, Juan es inocente y la verdadera culpable es María. Supongamos que la evidencia documental presentada por el detective aumenta en gran medida la probabilidad de la verdad de la versión del detective. Su testimonio me obligará entonces a revisar mi conjunto de creencias, pero la aceptación del testimonio mismo estará basada en información consistente con mi conjunto de creencias inicial¹⁰.

La segunda condición afirma que la justificación potencial no puede estar previamente aceptada en K . Podría parecer que en algunos casos la justificación potencial ya forma parte del conjunto de creencias del agente. Sin embargo, aunque puede suceder que el hecho relevante ψ ya se encuentre en K , la oración de probabilidad no puede estarlo, puesto que el testimonio es nuevo para el agente. Como Ψ es una conjunción que incluye la oración de probabilidad, la violación de esta condición es tan sólo aparente.

La tercera y la cuarta condición pueden ser analizadas conjuntamente. La tercera afirma que la justificación potencial debe incluir una oración ψ que describa el factor relevante. La cuarta, que esa oración ψ es positiva o negativamente relevante para la verdad del testimonio. La relevancia de ψ es relativa a K . En la mayoría de los casos sólo se requiere sentido común para determinar los hechos que aumentan la probabilidad de que un testimonio sea

¹⁰ Puede ocurrir, por supuesto, que un testimonio que fue rechazado tomando como base a K pueda ser aceptado posteriormente a raíz de una revisión de K consistente con nuevas justificaciones potenciales.

verdadero, pero en otros casos el contexto puede ser muy especializado y se requerirá el conocimiento de expertos. En estos casos el no experto terminará con un conjunto empobrecido de justificaciones potenciales, lo cual lo puede llevar a rechazar testimonios que un experto habría aceptado. La ignorancia puede ser un obstáculo en la búsqueda de evidencia justificada.

Es obvio que un testimonio que afirme que β siempre será más probable a la luz de una oración que describa el hecho β . Siempre y cuando esta última oración sea consistente con K , mi propuesta permite que haya justificaciones potenciales que contengan una oración que describa el hecho β y una oración de probabilidad que afirme que β aumenta la probabilidad de que el testimonio sea verdadero. Como el agente no quiere prejuzgar el asunto a favor del testimonio, a esta justificación potencial se le deberá asignar el valor informacional más bajo posible. Más aún, si el testimonio es intrínsecamente poco creíble, su probabilidad doxástica será extremadamente baja. En ambos casos, la justificación potencial que incluye a la oración que describe a β no calificará como una justificación legítima, como veremos más adelante.

La cuarta condición no impone ninguna restricción sobre el valor de $p(\phi)$ en K . Esto permite que haya casos en los que la probabilidad epistémica de ϕ sea bastante alta. Sería un error, sin embargo, establecer un límite superior para la probabilidad epistémica previa de ϕ en K . Incluso cuando un testimonio parece plausible a los ojos del agente, éste puede preguntarse si puede utilizarlo como evidencia al tomar una decisión muy importante, y simplemente quiere estar seguro. Dado que una de las desideratas al expandir un estado de creencia es evitar el error, la confianza que el agente tenga en la verdad del testimonio debe ser compensada por el valor de la información que contiene. Si el riesgo de error es bajo, pero la información es prácticamente inútil, quizás no valga la pena correr el riesgo.

La quinta condición garantiza que ϕ y ψ no estén correlacionadas de manera espuria. De nuevo, esto sólo puede ser garantizado relativo a K y en algunos casos el conocimiento no experto puede no ser suficiente para descartar justificaciones potenciales en las cuales ϕ y ψ están de hecho espuriamente correlacionadas.

La sexta condición previene la inclusión de casos triviales en los que $p(\phi|\psi) = 1$ porque $\psi \vdash \phi$.

La última condición establece que cada justificación potencial contiene sólo un factor relevante. Normalmente habrá varios factores relevantes para la verdad de cualquier testimonio, y por ende varias justificaciones potenciales.

Utilizando nuestra definición de justificación potencial, podemos caracterizar ahora la noción de un *espacio de justificación*. Un espacio de justificación puede ser entendido como el conjunto de todas las posibles justifica-

nes potenciales de ϕ , independientemente de si el agente es consciente de ellas o no. Más formalmente:

(E1 ϕ) Para cada oración ϕ en K , hay un conjunto $\{Y_1, Y_2, \dots, Y_n\}$ tal que Y_i es un elemento del conjunto si y sólo si es una justificación potencial de ϕ . El conjunto, denotado E1 ϕ , es el *espacio de justificación* de ϕ .

El espacio de justificación contendrá justificaciones potenciales que son lógicas y/o empíricamente equivalentes. Por una parte, si $Y_1 = \{\psi, p(\phi|\psi) > p(\phi)\}$ y $Y_2 = \{\eta, p(\phi|\eta) > p(\phi)\}$, donde ψ y η son lógicamente equivalentes, entonces Y_1 y Y_2 son lógicamente equivalentes. Si un agente acepta Y_1 , también se compromete con Y_2 . Por otra parte, si ψ y η contienen términos singulares o predicados coextensivos, que ocupan los mismos lugares en ψ y η , Y_1 y Y_2 serán justificaciones potenciales empíricamente equivalentes. Sin embargo, el valor informacional y la credibilidad de Y_1 y Y_2 no serán evaluadas del mismo modo a menos que el agente sea consciente de que los términos singulares y predicados son coextensivos.

El espacio de justificación provee todas las justificaciones posibles de un testimonio relativas a un trasfondo de conocimiento dado, desde las más recónditas hasta las más obvias. Generalmente, sin embargo, no logramos identificar la totalidad de sus elementos. Nuestras opciones están restringidas a aquellas justificaciones potenciales que pueden ser obtenidas del contexto y del contenido del testimonio. En consecuencia debemos restringir las opciones disponibles para la expansión de K a un subconjunto del espacio de justificación. Este subconjunto será la base para la evaluación del riesgo de error incurrido y del valor informacional obtenido cuando aceptamos una justificación potencial.

Un conjunto de opciones justificativas con respecto a E1 ϕ , denotado $O\phi$, es el subconjunto del espacio de justificación de ϕ que contiene todas las justificaciones potenciales de ϕ que el agente ha logrado identificar. Para simplificar el análisis, cada justificación potencial Y_i en el espacio de justificación será representada en $O\phi$ por la conjunción de sus elementos, esto es, por la conjunción de una oración singular y de una oración de probabilidad, y denotado como J_i .

Como $O\phi$ contendrá justificaciones potenciales bastante dispersas, el agente debe dividirlas entre aquellas que parecen creíbles, aquellas que parecen sospechosas y aquellas que parecen extravagantes. En otras palabras, el agente debe asignar una distribución de probabilidad doxástica a las oraciones del conjunto $O\phi$. La función de probabilidad doxástica resultante, denotada C , determina el riesgo de error incurrido al aceptar una justificación potencial en $O\phi$. Para cada justificación potencial J_i , el riesgo de error es: $1 - C(J_i)$.

El riesgo de error en el que se incurre al aceptar una justificación del conjunto de opciones justificativas debe ser compensado por el valor informacio-

nal obtenido. Las justificaciones potenciales en $O\phi$ serán más o menos valiosas para el agente dependiendo de cuánta información contengan y de qué tan útil sea esa información para sus fines cognitivos. ¿Cómo debemos evaluar el contenido informacional de una justificación potencial? El reto principal es encontrar una base objetiva para la evaluación. La idea que adoptaré aquí fue propuesta inicialmente por POPPER (1935) y ha sido defendida de diversas maneras por CARNAP, BAR HILLEL, LEVI y otros. POPPER afirma que el contenido de una hipótesis, su «grado de falsación», debe ser medido a partir del número de hipótesis rivales que excluye. Del mismo modo, sostengo que el primer paso en la evaluación del contenido de una justificación potencial en $O\phi$ debe ser la determinación del número de justificaciones potenciales excluidas. El contenido por sí solo no será suficiente para obtener un ordenamiento completo de todos los elementos de $O\phi$ con respecto a su valor informacional, pero por lo menos proveerá una base objetiva para el ordenamiento.

La idea de POPPER se puede precisar de la siguiente manera. LEVI (1984: cap. 5) ha propuesto un método para medir el contenido informacional de las expansiones potenciales de un estado de creencia, el cual captura la idea de que el contenido de una hipótesis debe ser medido de acuerdo al número de hipótesis rivales que excluye. La propuesta de LEVI puede ser adaptada fácilmente para medir el contenido informacional de una justificación potencial. La idea es introducir una medida M que le asigne valores no negativos a los elementos de $O\phi$ tal que su suma sea 1 y tal que el valor M de cualquier disyunción de elementos de $O\phi$ sea igual a la suma de sus valores M . El incremento en valor informacional al expandir K añadiendo J_i es la suma de todos los valores M de los elementos de $O\phi$ que son excluidos. En otras palabras, el contenido informacional de J_i , denotado $Cont(J_i)$, es $1 - M(J_i)$. Como LEVI lo señala, la función M tiene las propiedades formales de una función de probabilidad.

La siguiente pregunta que debemos considerar se refiere a la manera en que se deben asignar los valores M . CARNAP, BAR HILLEL y otros utilizaron nociones probabilísticas de contenido, pero la función M siempre fue interpretada como una función de probabilidad doxástica o como una medida del grado de confirmación de una hipótesis. Lo novedoso de la propuesta de LEVI es que él rechaza estas interpretaciones e introduce una probabilidad de contenidos, la cual está basada en el orden parcial de las expansiones potenciales de K , introducida por una relación de consecuencia clásica. El conjunto de expansiones potenciales de K es un álgebra booleana en la cual el máximo es K y el mínimo es O , el estado inconsistente. Si la función M se define sobre este conjunto, generará un orden parcial de sus elementos; y si el único elemento que tiene probabilidad cero es O , las expansiones potenciales aumentarán estrictamente en probabilidad a medida que decrece su fuerza lógica. Cuando la función M se define sobre el conjunto de potenciales expansiones

de interés para el investigador, esto es, para una «partición última» en la jerga de LEVI, obtenemos una medida del contenido informacional de las potenciales expansiones relevantes. En el caso de las justificaciones potenciales de un testimonio, yo he transformado la función M en una medida del contenido informacional de las justificaciones potenciales en $O\phi$.

LEVI establece una distinción entre contenido informacional y valor informacional, una distinción que es capturada por el requisito de monotonicidad débil (RMD). Aquí adoptaré un requisito análogo para el caso de las justificaciones potenciales:

(RMD) Si una justificación potencial J_1 en $O\phi$ contiene al menos la misma cantidad de nueva información que otra justificación potencial J_2 en $O\phi$, J_1 contiene al menos la misma cantidad de valor informacional nuevo que J_2 .

El requisito de monotonicidad débil genera un cuasiorden de las justificaciones potenciales en $O\phi$, el cual está basado en el orden parcial generado por la función M , pero que no es necesariamente idéntico a éste. Así, un agente puede considerar que la justificación potencial J_1 , que es consecuencia de otra justificación potencial J_2 de acuerdo con el orden parcial, contiene el mismo valor informacional que la justificación potencial más fuerte porque la información adicional en J_2 no tiene ningún valor para él. Pero la justificación potencial más fuerte no puede tener *menos* valor informacional que la más débil. Aunque alguna información es inútil, nunca carece completamente de valor.

Aquí encontramos el corazón del problema del valor informacional. Por una parte, a pesar de que el cuasiorden generado por el requisito de monotonicidad débil está basado en el orden parcial generado por la función M , diferentes agentes pueden evaluar el valor informacional de maneras diferentes, todas ellas compatibles con la función M y con el requisito de monotonicidad débil. Por otra parte, típicamente habrá elementos de $O\phi$ que no son comparables en términos de fuerza lógica. En consecuencia, no podrán ser parcialmente ordenados por la función M . Son casos de justificaciones potenciales residuales.

Como el agente quiere adoptar las mejores justificaciones disponibles, puede utilizar criterios adicionales para evaluar el valor informacional de todos los elementos de $O\phi$. La pregunta es hasta qué punto puede ser completado el cuasiorden y qué consideraciones son relevantes en el caso de la justificación del testimonio. Hay muchos valores epistémicos que son mencionados en el contexto de la selección de teorías científicas: poder explicativo, poder predictivo, simplicidad, precisión, etc., pero no todos ellos pueden ser utilizados al juzgar el valor informacional de una justificación potencial. Sin embargo, no hace falta establecer un conjunto único de criterios. Si los criterios son tales que el agente tiene una idea clara acerca de su importancia con respecto a sus

intereses y fines epistémicos, y acerca de cómo deben ser utilizados en casos particulares, puede utilizarlos al evaluar las justificaciones potenciales. De esa manera podrá completar, en mayor o menor grado, el cuasiorden generado por el requisito de monotonicidad débil con base en la función M .

Sea $O^*\phi$ un conjunto de opciones justificativas tal que $O^*\phi \subseteq O\phi$ y tal que el valor M de cada elemento del conjunto haya sido determinado. Combinando la función de probabilidad doxástica C que determina el riesgo de error incurrido al aceptar una justificación potencial en $O^*\phi$ con la función M definida sobre los elementos de $O^*\phi$, obtenemos un valor que el agente puede usar para seleccionar las justificaciones aceptables de ϕ . Llamaré a este resultado el valor epistémico de una justificación potencial:

$$(VE) \quad VE(J) = \alpha C(J) + (1 - \alpha) \text{Cor}(J).$$

Si dividimos VE por α y restamos del resultado $q = (1 - \alpha)/\alpha$, obtenemos la siguiente transformación afín positiva de VE :

$$(VE) \quad VE(J) = C(J) - qM(J),$$

donde q es el índice de audacia. Como las justificaciones que el agente quiere aceptar no deben ser falsas sin importar qué tanto valor informativo tengan, debemos exigir que $\alpha \geq 0.5$, y por tanto que $0 \leq q \leq 1$. «Entre más alta sea q , más alta será la importancia relativa dada a la adquisición de nueva información en comparación con la evitación del error. En este sentido la audacia del agente investigador aumenta con el aumento de q » (LEVI, 1996: 172). Una vez se haya adoptado un índice q , el agente debe rechazar una justificación potencial en $O^*\phi$ si $VE(J)$ es negativo, negarse a comprometerse con ella si es 0, y aceptarla si es positiva. Cualquier justificación potencial en $O^*\phi$ con un valor epistémico positivo es una justificación de ϕ en K . La disyunción de todas estas justificaciones es la justificación de ϕ en K :

($J\phi$) La justificación de ϕ en K , denotada $J\phi$, es la disyunción de todas las justificaciones potenciales en $O^*\phi$ con un valor epistémico positivo.

Si no se encuentra ninguna justificación para el testimonio en cuestión, la duda original que motivó al agente a emprender la investigación no ha sido eliminada. El agente debe continuar con la suspensión del juicio hasta encontrar nueva información, y el testimonio no puede ser usado como evidencia por el momento¹¹. Por otra parte, la existencia de una justificación para ϕ no es suficiente para aceptarla. Como veremos en el próximo apartado, la pregunta de si un testimonio debe ser aceptado no depende sólo de la existencia de una justificación para él.

¹¹ Esto no se aplica a aquellos casos en que el agente ha contraído hipotéticamente su estado de creencia con el fin de no prejuzgar el testimonio. Discutiré tales casos al final del ensayo.

6. LA ACEPTACIÓN DEL TESTIMONIO COMO EVIDENCIA

Buscar y aceptar una justificación para ϕ en K es tan sólo la primera etapa del proceso de decisión. La aceptación del testimonio mismo es otra historia. No todo testimonio justificado debe ser aceptado como evidencia por el agente. Podría suceder que incluso la mejor justificación en $J\phi$ ofrezca un apoyo muy débil al testimonio. En tales casos, una justificación débil puede dejar intacta nuestra actitud epistémica hacia el testimonio.

La decisión de aceptar un testimonio debe estar basada en la misma estrategia que fue usada en la aceptación de su justificación. Es decir, la aceptación de ϕ debe estar basada en su valor informativo y su probabilidad doxástica, pero no con respecto a K , el conjunto de creencias que fue utilizado como base para evaluar las justificaciones de ϕ , sino con respecto al conjunto de creencias $K + J\phi$, el cual ya incluye las justificaciones aceptadas, y que será denotado como K^* .

Comencemos con la probabilidad doxástica de ϕ . El papel de $J\phi$ es proporcionar buenas razones para creer que el testimonio es verdadero. Si lo logra efectivamente, entonces no debería proporcionar buenas razones para creer que es falso. Más precisamente, $J\phi$ proporciona buenas razones para creer ϕ sólo si $C(\phi) > C(\sim\phi)$ en K^* . Ahora bien, como no hay restricción alguna sobre el valor de $C(\phi)$ en K , el conjunto de creencias original desde el cual fue emprendida la búsqueda de una justificación, podría suceder que $C(\phi)$ ya sea mayor que $C(\sim\phi)$ en K . En tales casos, el papel de $J\phi$ es reducir el riesgo de error al aceptar ϕ al incrementar su probabilidad doxástica. Sin embargo, la decisión final de aceptar ϕ deberá tener en cuenta su valor informativo, como veremos más adelante.

Si $C(\phi) = C(\sim\phi)$ en K^* , el agente debe suspender el juicio acerca de ϕ , y si $C(\phi) < C(\sim\phi)$ en K^* , debe rechazarla. El primer caso ocurriría si $\sim\phi$ era inicialmente más creíble que ϕ en K , y la justificación hizo a ϕ igualmente creíble. El segundo caso ocurre en el ejemplo de Michael Phelps, en el que el agente acepta una oración que aumenta la probabilidad de que el testimonio sea verdadero, a saber, «Michael Phelps tenía planeado utilizar la piscina local hoy», pero dicha información no logra que cambie su actitud epistémica acerca del testimonio que oyó de boca de un extraño.

El hecho de que $C(\phi) > C(\sim\phi)$ en K^* es sólo una condición necesaria pero no suficiente para la aceptación de ϕ . Hay dos razones por las que tener buenas razones para creer no garantiza la aceptación de un testimonio. La primera es que la justificación encontrada para ϕ puede no ser suficiente para remover completamente las dudas que motivaron la investigación. Esto puede pasar si no hay suficiente información disponible en el contexto o si el testimonio no

es enteramente claro o si el investigador no es muy competente en la búsqueda de justificaciones. Por cualquiera de estas razones, un testimonio justificado aún puede generar dudas razonables en el agente. La segunda razón es que la decisión de aceptar ϕ también debe tomar en consideración el valor informativo del testimonio. Incluso si un testimonio es más creíble que increíble, el agente puede decidir no aceptarlo si su valor informacional no es muy alto.

Pasemos entonces a considerar el valor informacional del testimonio. Si la pregunta acerca de la verdad de un testimonio surge en el curso de una investigación, el valor informacional de ϕ dependerá de lo útil que sea para los fines cognitivos y prácticos del agente. Así formulado, el valor informacional es una medida comparativa. El testimonio puede ser visto como (parte de) una respuesta a una pregunta por el qué, el cómo o el porqué, y como tal, puede ser comparada con otras respuestas potenciales a la misma pregunta. La forma de medir este contenido es la misma que en el apartado anterior, a través de la función M , que mide el contenido informativo de las potenciales respuestas a la pregunta de investigación.

Así, una vez el contenido informacional y la credibilidad de ϕ han sido evaluadas en K^* , procedemos como antes para encontrar el valor epistémico del testimonio:

$$(VE) \quad VE(\phi) = C(\phi) - qM(\phi)$$

El índice de audacia q debe ser el mismo que se utilizó en la selección de las justificaciones potenciales de ϕ porque el contexto epistémico no ha cambiado. Como antes, el agente debe rechazar ϕ si $VE(\phi)$ es negativo, negarse a comprometerse con ella si es 0, y aceptarla si es positiva.

En el caso de la aceptación del testimonio existe una complicación adicional que no contemplamos en el apartado anterior. Como se mencionó antes, hay casos en los que el agente contrae hipotéticamente su estado de creencia con el fin de no prejuzgar negativamente un testimonio. Cuando un agente contrae su estado de creencia hay un principio que guía el proceso, el principio de economía informacional: «La información, en general, no es gratuita, y por tanto las pérdidas innecesarias de información deben ser evitadas» (GÄRDENFORS, 1988: 49). Si un agente quiere contraer su estado de creencia eliminando $\sim\phi$, el estado $K - \sim\phi$ debe preservar tanto como sea posible de K , pero al mismo tiempo debe eliminar cualquier creencia o combinación de creencias que impliquen lógicamente a $\sim\phi$. El problema es que típicamente habrá varias formas posibles de formar una contracción de K que no implique a $\sim\phi$. ¿Cómo se debe determinar exactamente que creencias deben eliminarse?

Existen varias respuestas a esta pregunta en la literatura sobre revisión de creencias, entre las cuales se encuentran los subconjuntos maximales de K que no implican $\sim\phi$ (ALCHOURRÓN y MAKINSON, 1982) o subconjuntos míni-

mos de K que contienen $\sim\phi$ (ALCHOURRÓN y MAKINSON, 1985). La solución que quiero examinar brevemente aquí fue sugerida por GÄRDENFORS (1988) y GÄRDENFORS y MAKINSON (1988). Está basada en la idea intuitiva de que no todas las creencias en el estado de creencia de un agente están epistémicamente atrincheradas en el mismo grado¹². En los casos en los que se ve obligado a escoger entre dos creencias, un agente prescindirá de aquella que esté menos atrincherada.

El atrincheramiento epistémico de una oración no está conectado con su probabilidad doxástica. Si una oración ha sido aceptada en K , es juzgada como máximamente probable. Sin embargo, un individuo no considera que todas sus creencias estén igualmente atrincheradas. El criterio informal para determinar el grado de atrincheramiento epistémico de una oración es «cuán útil es para la investigación y deliberación» (GÄRDENFORS, 1988: 87). Algunas oraciones verdaderas son centrales para nuestras empresas cognitivas y otras son epistémicamente inertes. El atrincheramiento epistémico está conectado con el poder explicativo y predictivo, y en general con el valor informativo de una oración. De acuerdo con la definición ofrecida por GÄRDENFORS, es posible determinar el atrincheramiento epistémico relativo de todas las oraciones en un estado de creencia K , pero la definición es puramente cualitativa y no proporciona un procedimiento para asignar valores numéricos a las oraciones individuales. A pesar de esto, GÄRDENFORS presenta una serie de postulados para el atrincheramiento epistémico que sirven como base para la construcción de una función de contracción¹³.

Ahora bien, supongamos que un agente ha contraído su estado de creencia pasando al estado $K - \sim\phi$, utilizando el atrincheramiento epistémico para minimizar la pérdida de información. Sea $K^* = K - \sim\phi$ y $K^{**} = (K - \sim\phi) + (J\phi + \phi)$. Hay tres posibles resultados al pasar de K a K^* :

- 1) El agente puede decidir que $\sim\phi$ es verdadera después de todo porque no se encontró ninguna justificación para ϕ . El agente regresa a K sin tener que pasar a la segunda etapa del proceso de decisión. No hay ni pérdida ni ganancia en valor informacional durante el ejercicio hipotético de rechazar $\sim\phi$.
- 2) El agente encuentra justificaciones para ϕ y decide que ϕ es verdadera porque su VE es positivo con respecto a $K^* + J\phi$. El valor informacional de K^{**} puede ser mayor, menor o igual al de K^* , puesto que, por una parte añade ϕ y $J\phi$, pero pierde $\sim\phi$ y cualquier oración en K que la implique lógicamente.
- 3) El agente no puede decidirse porque el $VE(\phi) = 0$ en $K^* + J\phi$, y se niega a añadir ϕ o $\sim\phi$ a su corpus de creencias. El agente sufre una pérdida de valor informacional.

¹² Levi (1991: 142) usa el concepto contrario: los «grados de correteabilidad». El grado de atrincheramiento aumenta con una disminución en el grado de correteabilidad. Ambos conceptos deben ser entendidos en sentido epistemológico, no psicológico.

¹³ Los detalles técnicos pueden ser consultados en GÄRDENFORS (1988).

Desde el punto de vista de la teoría de las decisiones, la mejor opción es negarse a contraer K a K^* , a menos que K^{**} tenga mayor valor informacional que K . «No tiene sentido correr el riesgo de terminar con K^* a menos que haya posibilidad de beneficiarse al terminar con K^{**} » (LEVI, 1991: 156). Informalmente, uno no debe poner en riesgo su estado de creencia a menos que exista la posibilidad de acceder a otro estado de creencia que tiene el potencial de ser más valioso en términos informacionales.

Ahora bien, el análisis presentado en este apartado asume que un testimonio puede ser comparado con otras respuestas potenciales a una pregunta de investigación con el fin de determinar su valor informacional. El problema es que no todos los testimonios son respuestas a preguntas de investigación; muchos ni siquiera son respuestas en absoluto. Si estoy leyendo una revista y me encuentro con una afirmación dudosa, no hay una pregunta previa involucrada. Quizás emprenda la tarea de averiguar si la afirmación puede ser justificada, simplemente para satisfacer mi curiosidad, no porque sea parte de una investigación o porque tenga alguna utilidad práctica. ¿Cómo debemos evaluar el valor informacional de un testimonio en tales casos?

Mi sospecha es que cuando una duda acerca de un testimonio surge de manera aislada, dissociada de cualquier contexto cognitivo o práctico, y cuando el testimonio no puede ser usado como evidencia para investigaciones o deliberaciones ulteriores, su valor informacional es extremadamente bajo. La duda en tales casos puede ser un ejemplo de lo que THAGARD llama una «duda recalcitrante», una duda que «implica una actitud hacia una proposición que es tanto emocional como cognitiva» (2004: 391). Si el valor informacional del testimonio es extremadamente bajo, la decisión de aceptarlo tendrá que basarse en el hecho de que es más creíble que increíble a la luz de las posibles justificaciones que el agente pueda encontrar para él¹⁴.

7. CONCLUSIÓN

En este ensayo he presentado una aproximación regla-reductiva al testimonio, esto es, una aproximación basada en reglas de adquisición de creencias que se aplican a creencias de todo tipo y no sólo a las creencias testimoniales. La aceptación o el rechazo de un testimonio obedece los mismos principios de la teoría de las decisiones que guían la aceptación o el rechazo de cualquier otra creencia. Al mismo tiempo, la propuesta toma en consideración las características específicas de las creencias testimoniales. En primer lugar, busca

detectar las fuentes principales de falsedad en el testimonio: las intenciones maliciosas y la simple incompetencia. Ambas son detectadas en la etapa justificativa al identificar los factores que promueven la verdad del testimonio. Al examinar si hay razones para confiar en un testigo o fuente, o si su localización espacio-temporal hace más probable que haya observado el evento reportado, mi aproximación minimiza las posibilidades de error al aceptar un testimonio. En segundo lugar, al enfocarse en la duda como el estímulo para emprender cualquier investigación o indagación, mi aproximación también explica la diferencia entre el testimonio aceptado automáticamente y el testimonio que es aceptado a través de un proceso de deliberación. Finalmente, mi aproximación provee un criterio para determinar cuándo un testimonio puede ser usado como evidencia.

A algunos la propuesta les parecerá excesivamente subjetiva. Por una parte, la evaluación del riesgo de error y del valor informacional tanto del testimonio como de sus posibles justificaciones se lleva a cabo a partir del estado de creencia de un agente individual. Por otra parte, es siempre la duda de un agente individual lo que pone en marcha el proceso de búsqueda de una justificación para el testimonio. Las preocupaciones sobre el primer aspecto pueden ser mitigadas señalando que es posible extender la aproximación presentada aquí de tal modo que la evaluación del riesgo de error y del valor informacional se haga a partir de un corpus de creencias que represente los acuerdos compartidos de la comunidad científica. Yo mismo he presentado una aproximación semejante para el caso del concepto de explicación (PÁEZ, 2006).

Con respecto al problema de la duda individual, considero que PERCE tiene razón al afirmar que el punto de partida de los investigadores individuales es irrelevante para el establecimiento de una opinión común: «Mentes diferentes pueden comenzar con puntos de vista completamente opuestos, pero el proceso de investigación los conduce por una fuerza externa a ellos mismos hacia la misma conclusión. [...] A la opinión que está destinada a ser aceptada por todos aquellos que investigan, es a lo que llamo la verdad, y el objeto representado en esta opinión es lo real» (1878: 138-139).

BIBLIOGRAFÍA

- ALCHOURRÓN, C. E.; GÄRDENFORS, P., y MAKINSON, D., 1985: «On the Logic of Theory Change: Partial Meet Functions for Contraction and Revision», *Journal of Symbolic Logic*, 50.
- ALCHOURRÓN, C. E., y MAKINSON, D., 1982: «On the Logic of Theory Change: Contraction Functions and their Associated Revision Functions», *Theoria*, 48: 14-37.
- 1985: «On the Logic of Theory Change: Safe Contractions», *Studia Logica*, 44.
- COADY, C. A. J., 1973: «Testimony and Observation», *American Philosophical Quarterly*, 10.

¹⁴ Formalmente, con respecto a K^* , si $\text{Cont}(\phi)$ es casi cierto, i. e., si $M(\phi) \approx 0,99$, entonces para un índice promedio de audacia q de 0,5, $\text{VE}(\phi)$ sólo será positivo si $C(\phi) \geq 0,5$, puesto que $\text{VE}(\phi) = C(\phi) - qM(\phi)$.

- 1992: *Testimony. A Philosophical Study*, Oxford: Clarendon Press.
- FRICKER, E., 1995: «Telling and Trusting: Reductionism and Anti-Reductionism in the Epistemology of Testimony», *Mind*, 104.
- GÄRDENFORS, P., 1988: *Knowledge in Flux. Modeling the Dynamics of Epistemic States*, Cambridge: MIT Press.
- GÄRDENFORS, P., y MAKINSON, D., 1988: «Revisions of Knowledge Systems Using Epistemic Entrenchment», *Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning about Knowledge*, San Francisco: Morgan Kaufmann.
- GOLDMAN, A., 1999: *Knowledge in a Social World*, New York: Oxford University Press.
- HAAK, S., 1993: *Evidence and Inquiry*, Oxford: Blackwell.
- 2003: *Defending Science within Reason*, Amherst: Prometheus Books.
- HARDWIG, J., 1985: «Epistemic Dependence», *The Journal of Philosophy*, 82.
- 1991: «The Role of Trust in Knowledge», *The Journal of Philosophy*, 88.
- HUME, D., 1748: «Of Miracles», Citado por la traducción al castellano de DE SALAS, J., 1980, «De los milagros», *Investigación sobre el entendimiento humano*, Madrid: Alianza Editorial.
- LEVI, I., 1967: *Gambling with Truth*, New York: Knopf.
- 1980: *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability and Chance*, Cambridge: MIT Press.
- 1984: *Decisions and Revisions*, Cambridge: Cambridge University Press.
- 1991: *The Fixation of Belief and Its Undoing*, New York: Cambridge University Press.
- 2004: *Mild Contraction. Evaluating Loss of Information Due to Loss of Belief*, Oxford: Clarendon Press.
- LIPFON, P., 1998: «The Epistemology of Testimony», *Studies in the History and Philosophy of Science*, 29.
- 2004: *Inference to the Best Explanation*, 2.ª ed., London: Routledge.
- PÁEZ, A., 2006: *Explanations in K. An Analysis of Explanation as a Belief Revisory Operation*, Oberhausen: Athena Verlag.
- PERCE, C. S., 1877: «The Fixation of Belief». Citado por la edición de HOUSER, N., y KLOESEL, C., 1992: en *The Essential Peirce. Selected Philosophical Writings* (1), Bloomington: Indiana University Press.
- POPPER, K., 1935: *Logik der Forschung*, Vienna: Springer.
- REID, T., 1764: *Inquiry into the Human Mind on the Principles of Common Sense*. Citado por la edición de BROOKES, D. R., 2000, Edinburgh: Edinburgh University Press.
- SHAPIN, S., 1994: *A Social History of Truth*, Chicago: University of Chicago Press.
- SPEER, D.; CLÉMENS, F.; HEINTZ, C.; MASCARO, O.; MERCIER, H.; ORIGGI, G., y WILSON, D., 2010: «Epistemic Vigilance», *Mind and Language*, 25.
- STERELNY, K., 2012: *The Evolved Apprentice*, Cambridge: MIT Press.
- THAGARD, P., 2000: *Coherence in Thought and Action*, Cambridge: MIT Press.
- 2004: «What is Doubt and When Is It Reasonable?», en EZCURDIA, M.; STAIN-TON, R., y VIGER, C. (eds.): *New Essays in the Philosophy of Language and Mind. Canadian Journal of Philosophy. Supplementary Volume*, Calgary: University of Calgary Press.
- 2005: «Testimony, Credibility, and Explanatory Coherence», *Erkenntnis*, 63.
- WELLS, G., y OLSON, E., 2003: «Eyewitness Testimony», *Annual Review of Psychology*, 54.

COLECCIÓN «FILOSOFÍA Y DERECHO»

ÚLTIMOS TÍTULOS PUBLICADOS

(Véase la lista completa en www.filosofiyderecho.es/titulospublicados.php)

Teoría de los principios
Humberto Ávila

Esta obra, escrita originalmente en portugués, obtuvo un enorme éxito en Brasil, donde se publicaron once ediciones en tan solo cinco años. La notoriedad alcanzada motivó su traducción al alemán y al inglés, y la obra experimentó el mismo éxito tanto en Alemania como en los países de lengua inglesa. En esta edición, totalmente revisada y ampliada en relación con el trabajo original, el autor presenta, con originalidad e independencia, un serio estudio sistemático sobre el significado y el fundamento de los principios jurídicos, proponiendo criterios distintos a los que la doctrina comúnmente emplea para distinguir entre principios y reglas, así como para su adecuada interpretación y aplicación.

Las pretensiones normativas del derecho
Un análisis de las concepciones de Robert Alexy y Joseph Raz
Paula Gaido

La presente obra parte de una pregunta que puede ser formulada de diferentes maneras: ¿en qué idea se está pensando cuando se afirma que el derecho es normativo? ¿por qué es importante entender al derecho como fuente de normas que generan deberes, y no meramente como un hecho, como un conjunto de contenidos semánticos o como un conjunto de actos de imposición de poder? Hay quienes sostienen que cuando se piensa en el derecho está implícita la idea de que es normativo, y que esta es una cuestión conceptual. Sin la idea de normatividad no se puede comprender al derecho. El análisis de las respuestas dadas por los filósofos del derecho Robert Alexy y Joseph Raz, se toma como brújula en la exploración del problema. Las nociones de pretensión de corrección, en el caso de Alexy, y la de pretensión de autoridad legítima, en el caso de Raz son, de este modo, objeto de principal análisis en el libro y, a través de ellas, el sentido en que el derecho, se dice, es fuente de razones justificatorias.

Cómo deciden los jueces
Richard A. Posner

Cómo deciden los jueces es un estudio acerca del complejo y diverso entramado de factores que llevan a los jueces a tomar las decisiones que toman. Posner pretende «descorrier el velo» que cubre una de las actividades fundamentales del derecho: la actividad de juzgar. Su objetivo es someter a escrutinio público los condicionantes de naturaleza sociológica, psicológica, económica, política, filosófica y también jurídica que de hecho influyen en la actividad decisoria de los tribunales. Guiado por este objetivo, aborda toda una serie de aspectos que van desde las condiciones laborales de la profesión de juez (sueldo, promoción, estabilidad en el puesto) hasta el papel jugado por sus convicciones ideológicas y políticas, sus filias y fobias partidistas y su función ante la opinión pública, pasando por la psicología y la personalidad de los jueces y los problemas que tienen que ver con la manera en que aceptan y conciben las pautas del método jurídico.

Pero el libro de Posner no sólo se mueve en este nivel descriptivo o explicativo, sino que contiene también un modelo normativo de juez. El escenario de contraste de este modelo es asimismo la jurisprudencia estadounidense, centrilmente los tribunales de apelación y el Tribunal Supremo. El autor reivindica, por ejemplo, que una comprensión cabal del papel del Tribunal Supremo pasa por entenderlo como tribunal político, y somete a revisión alguna de sus tendencias actuales, en concreto la de tomar la jurisprudencia constitucional de otros ordenamientos jurídicos como fuente autoritativa. Posner apuesta, en definitiva, por un modelo pragmatista, pero, en sus palabras, se trata de «un pragmatismo sensible y no uno de cortos vuelos».

Causalidad y responsabilidad
Un ensayo sobre derecho, moral y metafísica
Michael S. Moore

El concepto de causalidad es fundamental para asignar responsabilidad, moral y jurídica, por eventos. Pero la relación entre la causalidad y la responsabilidad permanece poco clara. ¿Cuál es, exactamente,

la conexión entre el concepto de causalidad usado al atribuir responsabilidad y las explicaciones de las relaciones causales ofrecidas por la filosofía de la ciencia y la metafísica? ¿Cuánto de lo que llamamos responsabilidad causal es definido, en verdad, por factores no causales? Este libro sostiene que una gran parte de la teoría jurídica sobre estas cuestiones es confusa e incoherente, y representa el primer intento exhaustivo, desde el trabajo de Hart y Honore, de aclarar el trasfondo filosófico de los debates jurídicos y morales.

En primer lugar, el libro ubica a la causalidad en el derecho penal y el derecho de daños, y define la metafísica presupuesta por la teoría jurídica. Luego analiza las mejores explicaciones teóricas de la causalidad ofrecidas por la filosofía de la ciencia y la metafísica y, valiéndose de ellas, critica a muchos de los conceptos jurídicos centrales que circundan a la causalidad, como los de la causalidad sobrevenida, la predictibilidad del daño y la participación. Considera y rechaza las propuestas radicales que tienden a erradicar a la causalidad del derecho, usando, en cambio, cálculos de riesgos para atribuir responsabilidad.

El análisis resulta ser un argumento poderoso para que revisemos nuestro entendimiento del papel que juega la causalidad en la atribución de responsabilidad, tanto jurídica como moral.

Jerarquías normativas y dinámica de los sistemas jurídicos

Jordi Ferrer Beltrán y Jorge L. Rodríguez

El libro aborda dos problemas centrales de la teoría del derecho: la jerarquía normativa y la dinámica de los sistemas jurídicos. Ambos han tenido hasta el momento menos atención de la que merecen, pero lo realmente novedoso de este volumen es su tratamiento conjunto, con un análisis muy cuidadoso de sus implicaciones mutuas. A partir de aquí, se revisan nociones clave como las de sistema jurídico, las concepciones de las normas, la pertenencia al sistema de las normas derivadas, o la validez y aplicabilidad de las normas, así como los efectos de la preferencia jerárquica. En resumen, el lector encontrará en este libro, un estudio profundo de algunas nociones centrales de la teoría del derecho, revisadas a la luz de la intersección entre la estructura jerárquica y la dinámica de los sistemas jurídicos.

El realismo jurídico genovés

Jordi Ferrer Beltrán y Giovanni B. Ratti

A partir de la obra de Giovanni Taruffo, y bajo su maestría, se ha conformado en Génova un grupo de investigadores, que contó enseguida con el amparo de Silvana Castiglione y Riccardo Guastini, y en el que se formaron filósofos del derecho como Mauro Barberis, Paolo Comanducci y Pierluigi Chiassoni, entre otros. En el ámbito de la teoría del derecho, ese grupo, conocido por muchos como *Escuela genovesa*, se ha caracterizado por la defensa de tesis propias del realismo jurídico y ha tenido una gran influencia en el debate filosófico italiano, francés e iberoamericano de los últimos 20 años. El realismo jurídico *à la genovese* hundió sus raíces en las dos grandes tradiciones lusitánicas, americana y escandinava. Sin embargo, ha desarrollado progresivamente sus propias tesis y refinado con la metodología analítica algunas de las asunciones básicas del realismo jurídico clásico, en especial del americano. En este libro el lector encontrará la presentación de esas tesis por parte de sus más destacados defensores y su discusión a cargo de algunos de los más importantes filósofos iberoamericanos actuales.

La teoría principalista de los derechos fundamentales

Estudios sobre la teoría de los derechos fundamentales de Robert Alexy
Jan-R. Sieckmann (ed.)

La teoría de los derechos fundamentales es un tema tan crucial como controvertido. Debido al aumento de Estados que se configuran como «constitucionales» y «democráticos», y también al desarrollo de sistemas europeos e internacionales de protección de derechos humanos, la búsqueda de modelos generales de protección justiciables cobra una enorme importancia. Entre estos modelos, la *Teoría de los derechos fundamentales* de Robert Alexy ha sido uno de los que más aceptación ha merecido tanto en el ámbito académico como en la práctica del derecho. Pese haber sido blanco de numerosas críticas, constituye, por su solidez teórica y reconocimiento internacional, uno de los referentes más adecuados sobre el tema. Los trabajos reunidos en este volumen analizan y discuten si la teoría principalista desarrollada por Alexy ofrece el marco idóneo para elaborar una teoría y una dogmática general de los derechos fundamentales más allá de las fronteras de los sistemas jurídicos nacionales. En ellos se abordan problemas relativos a la teoría de las normas justiciables, al método de la ponderación o a los principios de proporcionalidad y de igualdad. Aparte del propio Alexy, los autores que participan en este libro han contribuido en diversa manera e incluso desde perspectivas críticas al desarrollo de esa teoría general y pueden ser considerados como especialistas en la teoría de los principios y de los derechos fundamentales.

Problemas de vida o muerte

Diez ensayos sobre bioética
Eduardo Rivera López

El avance de la medicina ha generado enormes desafíos éticos. En ellos está en juego la vida, la muerte, la dignidad y la autonomía de las personas. ¿Debería permitirse la eutanasia voluntaria? ¿Tenemos el deber de garantizarles a nuestros futuros hijos un mínimo de calidad de vida? ¿Deberían las personas con problemas de fertilidad adoptar, en lugar de utilizar técnicas de reproducción asistida? ¿Generará el avance de la genética desigualdades sociales inaceptables? ¿Quién debería tener acceso a la información genética de una persona? ¿Debería permitirse la venta de órganos?

Estas son algunas de las preguntas que se exploran en los diez ensayos reunidos en este volumen. Se trata de artículos independientes, publicados en general en revistas especializadas, que ofrecen, sin embargo, un panorama amplio de algunos de los problemas de la bioética contemporánea.

Dilemas constitucionales

Un debate sobre sus aspectos jurídicos y morales
L. Zucca, G. Languet, D. Martínez Zorrilla y S. Álvarez

Los conflictos entre valores son uno de los temas más apasionantes y difíciles de la filosofía moral, en la medida en que su existencia, planta, un serio desafío a la coherencia de la ética como sistema normativo. Cuando este tipo de conflictos se traslada, a través de los principios y las normas constitucionales, al ámbito del derecho, los desafíos se multiplican. Se trata de un ámbito en el que la teoría del derecho y la teoría constitucional necesitan recurrir inexorablemente a la filosofía moral para poder ofrecer un análisis certero del tipo de problema que se debe abordar.

En este libro se analizan los conflictos morales y se estudian los diversos escenarios jurídicos que se abren a partir de su incorporación al ámbito del derecho constitucional. Para ello se parte de un artículo de Lorenzo Zucca —autor que recientemente se ha ocupado del tema de una manera novedosa y audaz— sobre los «dilemas constitucionales». Dicho artículo es seguido de los comentarios de otros tres autores —Silvana Álvarez, Guillermo Languet y David Martínez Zorrilla— que también han trabajado sobre los conflictos de derechos fundamentales y que realizan un análisis minucioso y original de la obra de Zucca.

Instituciones del derecho

Neil MacCormick

Instituciones del derecho contiene una presentación con pretensiones de exhaustividad de una teoría institucional del derecho. Pero, a diferencia de buena parte de la literatura contemporánea sobre la materia, es un libro con el cual el lector puede también aprender de derecho. Aunque comienza ofreciendo una definición de derecho como un orden normativo institucional, luego muestra cómo desde esa perspectiva es posible arrojar luz sobre cuestiones fundamentales del derecho público y del derecho privado. En este el libro es un libro de «instituciones» en un sentido que lo vincula a una forma tradicional de literatura jurídica, de la cual las *Instituciones del Derecho de Escocia* (1581) de Lord Stair son un ejemplo paradigmático.

Este es un libro escrito por un autor que cree que la labor de la teoría del derecho no se reduce a aclarar las confusiones que afectan al lenguaje de los juristas; que, al contrario, cree importante hacer un examen atento y sofisticado de las prácticas jurídicas que existen realmente, porque a través de su reconstrucción racional podemos entender la forma en que un orden normativo se transforma, por así decirlo, en instituciones. *Instituciones del derecho* es un intento de reconstruir de este modo el derecho contemporáneo.

Constitucionalismo popular y control de constitucionalidad

Larry D. Kramer

La obra de Larry Kramer propone una revisión de la historia social del control de constitucionalidad de las leyes que recupera el papel del pueblo como piedra angular del programa constitucional estadounidense. Su revisión muestra los límites de las reconstrucciones convencionales que han asignado al poder judicial, y en él a la Corte Suprema, el rol de intérprete último de la Constitución. Más bien, en la medullosa labor historiográfica realizada por Kramer, el modelo de supremacía judicial que hoy conocemos aparece como sólo uno entre los mecanismos disponibles para la resolución de los conflictos interpretativos de los departamentos del gobierno. Según nos muestra Kramer en esta interesante obra, la historia constitucional estadounidense es la historia del constitucionalismo popular. Ahora bien, una vez aceptada la invitación a repasar dos siglos de constitucionalismo desde el prisma de la soberanía popular, queda la reflexión sobre las invocaciones del autor a repositonar al pueblo, la democracia y el autogobierno en el centro del proyecto constitucional contemporáneo. ¿Será ello posible? ¿Es ello deseable?

Compendio de una teoría analítica del derecho

Alchourrón y Bullygin en sus textos
Daniel Mendonca (ed.)

Este libro intenta reunir fragmentos de distintos textos de Alchourrón y Bullygin, ordenados temáticamente, a fin de dar una idea abarcadora, pero resumida, del sistema de pensamiento que ellos elaboraron. En esos fragmentos, los dos autores hablan directamente a la mente del lector, en tanto su recopilación sirve de atajo simplificador. Si la síntesis llega donde antes no había llegado, si sirve para rectificar preconcepções, aclarar dudas y remover objeciones, la intención habrá sido satisfactoria. Si, además, tiene por virtud generar ideas, conceptos y argumentos divergentes, que puedan plantearse y debatirse con el mismo rigor de los que aquí se presentan, la comunidad de los lectores habrá cooperado con los autores en la más importante tarea común: hacer avanzar la reflexión filosófica en dirección al perfeccionamiento de la hoy vetusta práctica del derecho.

La teoría de la democracia en el mundo real

Ian Shapiro

Isaiah Berlin gustaba de clasificar en dos grandes grupos a los pensadores. Por un lado los zorros, para los que la realidad es compleja, inasimilable, diversa; por otro los erizos, para los que ha de existir un principio organizador que posibilite una explicación coherente de la pluralidad, un orden. Ian Shapiro se define a sí mismo como «aspirante a erizo», y este libro es uno de los capítulos de su viaje intelectual en ese sentido.

Tal y como el título nos adelanta, en él encontramos teoría y práctica. Y, tanto en la una como en la otra, la pluma de Shapiro va más allá de los discursos habituales y obliga al lector a revisar presupuestos y a adoptar una perspectiva distinta a la establecida.

Locke se presenta aquí como defensor de la legitimidad democrática, y no (o no solo) como liberal. Madison aparece como defensor de una teoría pluralista de la democracia que Shapiro defiende como el mejor recurso a nuestro alcance contra la dominación. Pero esas y otras singularidades teóricas no se presentan aisladas, redactadas en exclusiva para la torre de marfil. Muy al contrario, se entremezclan y se traen a colación al alimón con la realidad de nuestro mundo.

Y, así, la teoría de la democracia se esgrime para lidiar con el aborto y el papel de la revisión judicial, o con los procesos de paz y las transiciones a la democracia, o con el cosmopolitismo y el derecho a la injerencia, o con la eliminación del impuesto de patrimonio y los turbios procedimientos previos tendientes a encauzar la opinión pública mediante la preparación de encuestas sesgadas. Estamos ante un teórico, y uno de primera línea. Pero un teórico que, a diferencia de otros, desciende al mundo real y se las ve con él.

Técnicas de interpretación jurídica

Breviario para juristas
Pierluigi Chiassoni

La interpretación es un difícil banco de pruebas para el jurista; este libro se propone ofrecer a quien aborde el estudio del derecho o a quien en cualquier caso está interesado en adquirir una mayor conciencia metodológica, un breviario de las nociones fundamentales de técnicas de la interpretación jurídica, acompañado de las oportunas coordinadas teóricas. Para tal fin, se ilustran algunos tipos de silogismos judiciales útiles para el análisis de la jurisprudencia, los pasajes claves de la interpretación de los textos normativos (desde la identificación de las disposiciones a su traducción en normas jurídicas), los instrumentos de la argumentación interpretativa, los resultados de la interpretación, el problema de la fidelidad del intérprete al derecho, la disciplina positiva de la interpretación de la ley, algunas concepciones de la interpretación constitucional, las legunas y los conflictos normativos, con sus modos de identificación, integración y subsunción.

El legado de H. L. A. Hart

Filosofía jurídica, política y moral
M. H. Kramer, C. Grant, B. Colburn y A. Hatzistavrou

En julio de 2007 se celebró un simposio de la Academia Británica en homenaje a los cien años del nacimiento de H. L. A. Hart, el insigne jurista inglés cuyas contribuciones al ámbito de la filosofía del derecho, la filosofía política y la filosofía moral se cuentan entre las más importantes del siglo xx. En este volumen se recopilan diecisiete trabajos cuyas versiones preliminares fueron presentadas como ponencias en dicho simposio. La riqueza y amplitud de la obra de Hart se ve reflejada en los diferentes capítulos en los que se ha dividido la obra: teoría general del derecho y positivismo jurídico, responsabilidad penal, causalidad en el derecho, justicia, derechos, tolerancia y libertad. Por otra parte, la decisiva y perdurable influencia del

pensamiento de Hart en el estado de discusión actual de esos temas se pone de manifiesto en la jerarquía de los autores que han aportado sus colaboraciones: R. A. Duff, Cécile Fabre, John Finnis, John Gardner, Leslie Green, Brad Hooker, David Lyons, Susan Mendus, Phillip Pettit, Gerald J. Postema, Alan Ryan, Hillel Steiner, Judith Jarvis Thomson, Jeremy Waldron, W. J. Waluchow, Leif Wenar y Richard W. Wright.

Positivismo jurídico y sistemas constitucionales

Claudia Crunescu

En los últimos años se ha popularizado la tesis de que el positivismo jurídico no constituye una teoría adecuada para dar cuenta de los sistemas jurídicos tal como se presentan en las democracias actuales. Este libro aspira a mostrar que el positivismo jurídico sí está a la altura del desafío. Con ese cometido son analizados algunos de los temas más relevantes —y complejos— que ofrece el diseño institucional de las democracias constitucionales: ¿cómo se relacionan los derechos básicos? ¿cómo se relacionan los derechos básicos con los derechos constitucionales? ¿cómo opera el proceso tendiente a garantizar la aplicación de las normas constitucionales? ¿está justificada la imposición de límites a la legislación ordinaria a través de una constitución y, si es así, quién está legitimado para garantizar su supremacía? ¿es necesario un acercamiento teórico diferente del positivismo jurídico para dar cuenta del funcionamiento de los sistemas jurídicos y, en especial, del control de constitucionalidad?

Compendio de filosofía del derecho

Rafael Hernández Marín

Esta obra es esencialmente una exposición, abreviada en muchos casos, de los diversos temas en que se divide, en opinión del autor, la teoría general del derecho. Aunque resulta más detallada en algunos de los temas más importantes para la práctica del derecho, como son la interpretación del derecho y la aplicación del derecho, dos temas que no suelen ser debidamente distinguidos, pero que en esta obra son diferenciados con toda nitidez.

Los intersticios del derecho

Indeterminación, validez y positivismo jurídico

Ángeles Ródenas

La expresión intersticios del derecho, empleada en el título de este libro, hace referencia a la zona, de límites borrosos, situada entre aquellas pautas que son inequívocamente reconocidas como derecho y aquellas otras que claramente no lo son. La indagación sobre esta área de penumbra es la constante que vertebra los tres núcleos temáticos del libro: la indeterminación del derecho, la validez jurídica y la crisis actual del positivismo jurídico. El conjunto de los tres ensayos permite contemplar, bajo una luz nueva, esta área intersticial o penumbrosa del derecho.

Acción, dolo eventual y doble efecto

Un análisis filosófico sobre la atribución de consecuencias probables

María Laura Manrique Pérez

En el derecho penal contemporáneo, debido a la creciente importancia de la doctrina del dolo eventual, la distinción entre resultados intentados y consecuencias previstas no es tomada como normativamente relevante. De acuerdo a esta doctrina, ambos efectos de nuestras acciones son equivalentes en el momento de atribuir responsabilidad penal. Por el contrario, en filosofía moral se utiliza esta distinción por medio de la doctrina del doble efecto, que entiende que nuestra responsabilidad es mayor cuando provocamos un daño de manera directa que cuando el daño es producto de nuestra intención oblicua. Nuevas aproximaciones a la doctrina del doble efecto conectan esta estrategia con el liberalismo moral y político. Si el derecho penal necesita estar justificado por principios liberales —como las constituciones y los códigos penales parecen exigir—, la doctrina del doble efecto genera un importante desafío a nuestras concepciones de responsabilidad penal. Más específicamente, este trabajo sostiene que la doctrina del doble efecto provee de una mejor solución al problema de la distinción entre resultados intentados y consecuencias previstas que la doctrina del dolo eventual.

Particularismo y derecho

Un abordaje positivista en el ámbito práctico

Hernán G. Bouvier

Las afirmaciones del tipo «matar está mal» o «Pedro debe responder por el mal causado» suelen ser sometidas a escrutinio exigiendo su justificación. Según una concepción arraigada en el terreno práctico y que se denomina «universalismo», los juicios o acciones o acciones están justificadas sólo si consti-

yen el seguimiento o aplicación de un parámetro, regla o principio de cierto tipo. El universalismo es objetado de dos maneras. Por un lado por el escepticismo, para el cual nunca tiene sentido hablar de afirmaciones o acciones justificadas. Por el otro por el particularismo, para el que puede hablarse de acciones o afirmaciones justificadas aunque no impliquen la aplicación de un principio o regla. Según la particularista ninguna regla o principio pueden ser lo suficientemente atentos o receptivos a la complejidad de los casos sobre los que habrá que decidir. La discusión entre universalismo, escepticismo y particularismo se proyecta directamente en el plano del derecho. En el terreno del derecho se puede sostener o bien que la decisión de un juez está justificada sólo si se basa en una regla o principio general, o bien que nunca está justificada, o bien que puede existir una justificación jurídica a pesar de que la decisión no se base en una regla o principio. Este libro analiza la discusión entre universalismo, escepticismo y particularismo y su proyección en la teoría del derecho. Se pretende defender una variante del particularismo abordando en tono crítico los postulados de la filosofía analítica clásica. Si se concede que nuestro conocimiento es necesariamente parcial y algunas veces verdadero, y se acepta la existencia de libertad, entonces puede haber buenas razones para ser particularistas o, más bien, para otorgar prioridad a la experiencia frente a la mera forma y a la técnica.

Naturalismo y teoría del derecho

Brian Leiter

En el ámbito de la teoría del derecho contemporánea, en particular de la anglosajona, el realismo jurídico es a menudo considerado un movimiento que resulta atractivo sólo a juristas filosóficamente superficiales y que no ha brindado ninguna contribución sustancial a la reflexión teórica acerca del derecho. En contra de tal opinión, este libro ofrece una reinterpretación original y novedosa de los realistas jurídicos norteamericanos como precursores del «giro naturalista» en filosofía. A la vez que reubraza su concepción con una serie de argumentos finamente elaborados para rebatir las críticas que le han sido dirigidas por diferentes movimientos teórico-jurídicos, como, por ejemplo, el positivismo hartiano o los *Critical Legal Studies*. El volumen, además, reaviva el programa de investigación ilustrado, vinculándolo al debate metodológico contemporáneo y utilizándolo, en el ámbito de dicho debate, a favor de la defensa de una teoría del derecho de corte genuinamente descriptivo, concebida, de la manera naturalista, como la rama abstracta de las ciencias jurídicas predictivas. El volumen explora también las implicaciones de una concepción de la filosofía de estricta observancia naturalista acerca de las cuestiones de la objetividad del derecho y de la moral: en este marco, asume una particular relevancia filosófica el análisis demolidor que Leiter reserva a la concepción de Ronald Dworkin.

La idea de los derechos humanos

Charles R. Beitz

Desde su creación tras la Segunda Guerra Mundial, la suerte que han corrido los derechos humanos ha sido paradójica. Por un lado, se han transformado en el lenguaje en el que se expresa la política global. Formular reclamos fundados en los derechos humanos se ha vuelto algo usual a nivel internacional, y no existe nadie que considere que tales reclamos carecen de importancia. Por otro lado, de modo creciente se han vuelto objeto de una creciente ola de escepticismo teórico y suspicacia política. Tal circunstancia hace que la tarea de clarificar el concepto de derechos humanos, evaluar su contenido y argumentar su justificación, sea importante y necesaria. Esta tarea es la que acomete Charles Beitz en la presente obra.

Con el fin de elucidar el concepto de derechos humanos presente en la práctica internacional, Beitz elabora un modelo según el cual los responsables primarios de la satisfacción de los derechos humanos son los Estados. Y en el que la comunidad internacional aparece como garante. Adicionalmente, presenta un esquema de cómo debería justificarse el contenido específico de los derechos humanos y lo aplica a algunos casos controvertidos como son los derechos humanos de las mujeres, los derechos humanos en contra de la pobreza o el derecho humano a la democracia.

Acordos y desacuerdos

Cómo y por qué los juristas discrepan
Paul Luque Sánchez y Giovanni Battista Ratti (eds.)

El volumen recoge una serie de ensayos alrededor de la cuestión de los desacuerdos jurídicos. Su razón de ser responde sobre todo a dos tipos de inquietudes: por un lado, reunir en un único texto las aportaciones más relevantes, obra de los autores más prestigiosos y consolidados de la teoría del derecho contemporánea, acerca de los desacuerdos jurídicos; por otro lado, mostrar algunos de los nuevos caminos por los cuales esta cuestión podría transitar en un futuro próximo. Los ensayos que componen el volumen han sido divididos en cinco secciones temáticas: «Para una primera panorámica

de los desacuerdos jurídicos», «El argumento del desacuerdo», «Los desarrollos del debate alrededor del argumento del desacuerdo», «Fallacious Disagreements» y «Positivism conventionalista, realismo jurídico y regla de reconocimiento».

Esta colección está pensada para un público heterogéneo. Por un lado, tiene como potencial lector al docente o investigador de filosofía del derecho, así como al jurista académico o de profesión en general. Pero por otro lado, este libro ha sido concebido como un manual para estudiantes de un curso avanzado de filosofía del derecho, pues al debate sobre los desacuerdos jurídicos subyacen algunos de los problemas fundamentales de la teoría del derecho, como por ejemplo la conexión entre derecho y moral, la tesis de las fuentes sociales, el rol de la interpretación en la práctica jurídica o el estatuto lógico de las reglas últimas del sistema jurídico. Este abanico de cuestiones es presentado de una manera analítica, ordenada y taxonómica, de forma que incluso aquellos que no estén familiarizados con las cuestiones específicas que aquí se discuten puedan acceder con cierta claridad al intrincado mundo teórico de las discrepancias jurídicas.

Neutralidad y teoría del derecho

J. Ferrer Beltrán, J.J. Morero y D. M. Papayannis (eds.)

Este libro reúne a algunos de los filósofos del derecho más importantes en las tradiciones anglosajona y continental. Se trata de la compilación de los artículos que estos autores presentaron en el Congreso «Neutralidad y teoría del derecho», que tuvo lugar en la Universidad de Girona en mayo de 2010. Aunque el libro aborda muchos de los debates filosóficos actuales, la pregunta principal en torno a la cual se articula es si la teoría del derecho puede ser neutral. Esta obra ofrece un análisis general de la conexión entre el derecho y la moral y la posibilidad de determinar el contenido del derecho sin apelar a argumentos normativos; en el aspecto metodológico, examina el tipo de proyecto que actualmente desarrollan los filósofos del derecho; estudia los diferentes enfoques para teorizar acerca de la naturaleza o el concepto de derecho, el rol de análisis conceptual y las propiedades esenciales del derecho. Asimismo, arroja luz sobre lo que puede aprenderse del estudio de las propiedades no esenciales del derecho. Finalmente, analiza la naturaleza de los enunciados jurídicos y sus valores de verdad. Sin duda, este libro permitirá al lector avanzar significativamente en la comprensión de la práctica jurídica.

El gobierno de las normas

Giovanni Battista Ratti

Este libro recoge una serie de estudios que, desde una perspectiva realista, afrontan algunos de los problemas más discutidos en el debate filosófico contemporáneo. En varios de los escritos el autor presenta nuevas estrategias de análisis de tales cuestiones. El volumen está compuesto por cuatro partes. En la primera se examina la dependencia del carácter sistemático atribuido al derecho, de las operaciones de los juristas sobre los materiales normativos, vinculándolo al análisis de las concepciones del derecho que pretenden explicar el fenómeno de la sistematicidad del derecho. La segunda parte se interroga acerca de las posibles relaciones entre derecho y lógica: relaciones que han sido reconstruidas, alternativamente, como el «gobierno» del derecho sobre la lógica, o de la lógica sobre el derecho. Analizando la segunda opción, se termina demostrando la primera, arguyendo que, en el ámbito jurídico, la «corrección» de las inferencias con normas, realizadas por los juristas, no depende de la preservación de algún valor semántico, sino de las «prácticas interpretativas e inferenciales» difusas entre los operadores del derecho. La tercera parte trata el problema de la derivabilidad de las normas como medio de explicación de distintos fenómenos jurídicos: el diverso tratamiento lógico al que se someten las normas, la separación entre el derecho que es y el derecho que debe ser, la identificación y la solución de las antinomias, los desacuerdos entre juristas, entre otros. Al mismo tiempo, esta parte analiza las cuestiones que surgen de entender las fuentes jurídicas como totalmente mediables (y, por ende, las normas como generalmente derivables) a la luz de su (presunta) corrección o incorrección moral, y del querer formalizar dicha propiedad mediante sistemas lógicos divergentes de los sistemas de lógica clásica. La cuarta parte, finalmente, concluye al «universo de lo normativo». Una tarea fundamental para los teóricos del derecho, sobre todo de la era moderna, ha consistido en hallar un método para distinguir, en el seno de dicho universo, lo jurídico de lo no jurídico. Esta parte del libro concentra la atención en los problemas conceptuales vinculados a dichos intentos de demarcación del ámbito de lo jurídico.

Las paradojas de la acción

Una introducción a la teoría de la acción humana desde el punto de vista del derecho y de la filosofía
Daniel González Lagier

¿Qué es una acción? ¿Y una omisión? ¿Podemos hacer varias acciones con un único movimiento corporal? ¿Se pueden contar las acciones? ¿Cómo podemos averiguar las intenciones de los demás?

¿Cuándo se nos puede imputar una acción no intencional? Este libro pretende sugerir respuestas a preguntas como éstas (o, al menos, proporcionar herramientas conceptuales para encontrarlas) y está dirigido a quienes —desde el campo de la ética, el derecho o la filosofía— estén interesados en las principales concepciones (filosóficas y jurídicas) de la acción y en el análisis de su estructura.

Pensar como un abogado

Una nueva introducción al razonamiento jurídico
Frederick Schauer

Este libro trata sobre pensar y razonar. Más específicamente, sobre las maneras de pensar y razonar y los métodos argumentativos de los abogados y los jueces, que pueden o no ser diferentes de las maneras de pensar y razonar y de los métodos argumentativos de las personas comunes. El que los abogados piensan, razonan y argumentan distinto de la gente común es una pregunta, no un axioma. No obstante, ciertas técnicas de razonamiento se consideran características de la toma de decisiones en el derecho.

En la facultad de derecho típica, en especial en los Estados Unidos, los docentes creen que enseñan pensamiento y razonamiento jurídicos por ósmosis, o intersticialidad, en el proceso de instruir sobre materias sustantivas como derecho de daños, contratos, derecho penal, derechos reales, procedimiento civil y derecho constitucional. Pero en realidad se enseña mucho menos sobre el pensamiento y el razonamiento jurídicos de lo que los profesores usualmente creen. Y aun si se enseñara en alguna medida, podría ser necesario hacerlo de manera unificada. Este libro busca satisfacer estas necesidades y, al mismo tiempo, ofrecer a los abogados y estudiantes de derecho algo en lo cual pensar —y estar o no de acuerdo— con relación a la mayoría de los temas que aborda.