



Quaestio facti

Revista Internacional sobre Razonamiento Probatorio
International Journal on Evidential Legal Reasoning

N. 2 | 2021

ISSN: 2604-6202

Michele Taruffo, in memoriam, Jordi Ferrer Beltrán

ENSAYOS

A critical method for the evaluation of evidence: François Gorphe's (1889-1959) contribution to a science of proof à la française, Olivier Leclerc

Predictibilidad conductual y proceso penal: algunos apuntes sobre el fundamento epistémico de los hechos futuros en las medidas cautelares, Jeffrey José Mora Sánchez

Evidence assessment and standards of proof: a messy issue, Giovanni Tuzet

Repensando las máximas de experiencia, Alan Limardo

La decisión sobre la quaestio facti en los acuerdos de culpabilidad, Diana Veleda

Las relaciones entre la práctica de los sistemas penales y la determinación de los estándares de prueba, Santiago Eyherabide

A privacidade digital posta à prova no processo penal, Paulo de Sousa Mendes

CONJETURAS Y REFUTACIONES

Naturalized Epistemology and the Law of Evidence Revisited, Ronald J. Allen

How to theorize about statistical evidence (and really, about everything else): a comment on Allen, David Enoch

Naturalized epistemology and the law of evidence: methodological reflections, Michael S. Pardo

Strong, weak, or apparent naturalization? Relative plausibility theory and conceptual analysis, Nicola Muffato

In defense of weird hypotheticals, Barbara A. Spellman

El testimonio único de la víctima en el proceso penal desde la perspectiva de género (2), José Luis Ramírez Ortiz

CIENCIA PARA EL PROCESO

Los conocimientos de la psicología más allá de la prueba pericial, Carmen Herrero

IURIS PRUDENTIA

¿Debería probarse el odio? Consideraciones en torno a la sentencia del caso Diana Sacayán, María Laura Manrique

Qf

Quaestio *facti*

Revista Internacional sobre Razonamiento Probatorio
International Journal on Evidential Legal Reasoning

N. 2 | 2021

www.quaestiofacti.com

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.

© CÁTEDRA DE CULTURA JURÍDICA

© MARCIAL PONS
EDICIONES JURÍDICAS Y SOCIALES, S. A.

San Sotero, 6 - 28037 MADRID

☎ (91) 304 33 03

www.marcialpons.es

ISSN: 2604-6202

Depósito legal: M. 3.518-2020

Fotocomposición: Ene Estudio

Impresión: ELECÉ, Industria Gráfica, S. L.

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

MADRID, 2021

DIRECTORES

Diego Dei Vecchi
Universitat de Girona, España

Jordi Ferrer Beltrán
Universitat de Girona, España

COMITÉ DE REDACCIÓN

Daniela Accatino
Universidad Austral de Chile, Chile

Daniel González Lagier
Universidad de Alicante, España

Edgar Aguilera
Universidad Autónoma del Estado de México.
Universitat de Girona, España

Carmen Vázquez
Universitat de Girona, España

COMITÉ EDITORIAL

Christian Dahlman
Lund University, Suecia

Antonio Manzanero
Universidad Complutense de Madrid, España

Mauricio Duce
Universidad Diego Portales, Chile

Paul Roberts
University of Nottingham, Reino Unido

Mercedes Fernández
Universidad de Alicante, España

Sarah Summers
University of Zurich, Suiza

Raymundo Gama
Instituto Tecnológico Autónomo de México

Giovanni Tuzet
Università Bocconi, Italia

Ho Hock Lai
National University of Singapore, Singapur

Jonatan Valenzuela
Universidad de Chile, Chile

CONSEJO ASESOR

- Ronald Allen
Northwestern University, EEUU
- Amalia Amaya
Universidad Nacional Autónoma de México, México
- Perfecto Andrés Ibáñez
Magistrado emérito de la Sala Segunda del Tribunal
Supremo Español
- José María Asencio
Universidad de Alicante, España
- Lorena Bachmaier
Universidad Complutense de Madrid, España
- Zhang Baosheng
University of Political Science and Law (CUPL),
China
- Juan Carlos Bayón
Universidad Autónoma de Madrid, España
- Lorenzo Bujosa
Universidad de Salamanca, España
- Rodrigo Coloma
Universidad Alberto Hurtado, Chile
- Margarita Diges
Universidad Autónoma de Madrid, España
- Gary Edmond
University of New South Wales, Australia
- Luigi Ferrajoli
Università degli Studi di Roma, Italia
- Paolo Ferrua
Università degli Studi di Torino, Italia
- Marina Gascón
Universidad de Castilla-La Mancha, España
- Joaquín González
Universidad Autónoma de Madrid, España
- Susan Haack
University of Miami, EEUU
- Juan Igartua
Universidad del País Vasco, España
- John Jackson
University of Nottingham, Reino Unido
- Larry Laudan
University of Texas, EEUU
- Giuliana Mazzoni
University of Hull, Reino Unido
- Dale Nance
Case Western Reserve University, EEUU
- Jordi Nieva-Fenoll
Universitat de Barcelona, España
- Eduardo Oteiza
Universidad Nacional de La Plata, Argentina
- Andrés Páez
Universidad de los Andes, Colombia
- Jairo Parra Quijano
Universidad Libre, Colombia
- Catherine Piché
Université de Montréal, Canadá
- Joan Picó i Junoy
Universitat Pompeu Fabra, España
- Geraldo Prado
Universidade Federal do Rio de Janeiro, Brasil
- Giovanni Priori
Pontificia Universidad Católica del Perú, Perú
- Vitor Lia de Paula Ramos
Uniritter, Brasil
- Frederick Schauer
University of Virginia, EEUU
- Paulo de Sousa Mendes
Universidade de Lisboa, Portugal
- Michele Taruffo
Università degli studi di Pavia, Italia
- William Twining
University College of London, Reino Unido
- Giulio Ubertis
Università Cattolica del Sacro Cuore, Italia
- Alan Uzelac
University of Zagreb, Croacia
- Adrian Zuckerman
University of Oxford, Reino Unido

ÍNDICE

FERRER BELTRÁN, JORDI: <i>Michele Taruffo, in memoriam</i> (Español)	9
FERRER BELTRÁN, JORDI: <i>Michele Taruffo, in memoriam</i> (English)	13
FERRER BELTRÁN, JORDI: <i>Michele Taruffo, in memoriam</i> (Italiano)	17
FERRER BELTRÁN, JORDI: <i>Michele Taruffo, in memoriam</i> (Português).....	21

ENSAYOS

LECLERC, OLIVIER: <i>A critical method for the evaluation of evidence: François Gorphe's (1889-1959) contribution to a science of proof</i> à la française	29
MORA SÁNCHEZ, JEFFRY JOSÉ: <i>Predictibilidad conductual y proceso penal: algunos apuntes sobre el fundamento epistémico de los hechos futuros en las medidas cautelares</i>	53
TUZET, GIOVANNI: <i>Evidence assessment and standards of proof: a messy issue</i> ..	87
LIMARDO, ALAN: <i>Repensando las máximas de experiencia</i>	115
VELEDA, DIANA: <i>La decisión sobre la quaestio facti en los acuerdos de culpabilidad</i>	155
EYHERABIDE, SANTIAGO: <i>La prueba ilícita ante la bifurcación del tribunal penal</i>	185
SOUSA MENDES, PAULO DE: <i>A privacidade digital posta à prova no processo penal</i>	225

CONJETURAS Y REFUTACIONES

ALLEN, RONALD J.: <i>Naturalized Epistemology and the Law of Evidence Revisited</i>	253
ENOCH, DAVID: <i>How to theorize about statistical evidence (and really, about everything else): a comment on Allen</i>	285
PARDO, MICHAEL S.: <i>Naturalized epistemology and the law of evidence: methodological reflections</i>	299
MUFFATO, NICOLA: <i>Strong, weak, or apparent naturalization? Relative plausibility theory and conceptual analysis</i>	313
SPELLMAN, BARBARA A.: <i>In defense of weird hypotheticals</i>	325
RAMÍREZ ORTIZ, JOSÉ LUIS: <i>El testimonio único de la víctima en el proceso penal desde la perspectiva de género (2)</i>	339

CIENCIA PARA EL PROCESO

HERRERO, CARMEN: <i>Los conocimientos de la psicología más allá de la prueba pericial</i>	363
---	-----

IURIS PRUDENTIA

MANRIQUE, MARÍA LAURA: <i>¿Debería probarse el odio? Consideraciones en torno a la sentencia del caso Diana Sacayán</i>	411
---	-----

MICHELE TARUFFO, *IN MEMORIAM*

Jordi Ferrer Beltrán

Mientras estábamos haciendo el trabajo de cerrar este segundo número de Quaestio facti, nos cayó como jarro de agua fría la noticia del fallecimiento de Michele Taruffo. Él fue, sin duda, una de las anillas más importantes de la cadena causal que llevó a la fundación de esta primera revista dedicada al razonamiento probatorio en el área de los sistemas romano-germánicos. A él se debe en gran medida el cambio radical de tendencia que va de un cierto abandono y desprecio por la relevancia de los problemas probatorios a que hoy estén cada vez más en el foco del análisis y del estudio en estos sistemas. Se han multiplicado exponencialmente las publicaciones al respecto, se han creado programas de estudio, se apuntan cambios sustanciales en la jurisprudencia y poco a poco también en la legislación.

Pero nada de todo esto sucedía cuando Taruffo escribió *Studi sulla rilevanza della prova* (1970) o su magistral *La prova dei fatti giuridici* (1992). Desde esos años hasta su fallecimiento ha sido un autor prolífico en libros y artículos, siempre escritos con la pretensión de claridad y precisión de un filósofo analítico. No en vano quiso estudiar con Amedeo Conte en Pavía, con Uberto Scarpelli en Milán (quien le recomendó la lectura de los positivistas lógicos para sus primeros estudios sobre la relevancia de la prueba) y fue ávido lector de Giovanni Tarello. Pero si hay algo que le caracteriza en su obra escrita es la permanente ambición de traspasar fronteras. Lo hizo en el sentido físico toda su vida, pues amaba viajar y descubrir lugares recónditos, lejos de las aglomeraciones, ingresando en selvas a la aventura, siempre a la búsqueda de maravillas arqueológicas que le permitieran comprender mejor los orígenes culturales. Gozaba de experimentar con tradiciones culinarias y, sobre todo, de la conversación de sobremesa con los amigos que hizo por todo el mundo. No me inviten a viajar, que corren el riesgo de que acepte, decía.

Esa ambición de traspasar fronteras la llevó también a su trabajo teórico: quería entender las instituciones jurídicas con profundidad y, siguiendo el ejemplo de su maestro, Vittorio Denti, siempre pensó que la mejor manera de hacerlo era con el método comparado, escudriñando sus orígenes y contextos culturales, conociendo los debates, la literatura y la jurisprudencia de otros países, comprendiendo porqué un mismo problema era afrontado y resuelto de modo distinto en unos países y otros. Pero Taruffo no pretendió solo traspasar las fronteras físico-políticas entre países y sistemas jurídicos, sino también huir del modo aislacionista de trabajar en el ámbito jurídico. Por eso, se interesó por la filosofía, la epistemología, la sociología y hasta por la neurociencia. Despreciaba, en sus palabras, a los glosadores del artículo 15, segundo párrafo, letra a) de cualquier ley, que le parecían productores de discursos miopes y faltos de interés. Comentando algunas de sus primeras obras, un conocido procesalista italiano dijo: “¡Bueno Taruffo, aunque lea cosas extrañas!”. Taruffo, en cambio, amaba el debate y, en especial, el ejercicio de pensar sin esos límites estrechos: “si un problema es serio afectará a todos o a muchos, independientemente de la lengua de cada uno. (...) [S]i un problema existe solo para quien habla una determinada lengua, con toda probabilidad no será un problema serio”¹.

Tuve la oportunidad de conocerle personalmente en el año 1998 en el congreso italo-español de teoría analítica del derecho. Presentarme a Michele Taruffo, siendo yo un joven doctor, es una de las muchas cosas que debo a Paolo Comanducci. Ese mismo día acordamos que me haría cargo de la traducción de *La prova dei fatti giuridici* al español. Fue un arduo y largo trabajo de casi 3 años, pero durante ese tiempo tuvimos centenares de intercambios de correo y conversaciones personales y aprendí tanto que me sentí muy ignorante. Me entusiasmó hasta tal punto su trabajo que hoy sigo estudiando esos mismos temas. En nuestros diálogos, que se hicieron cada vez más frecuentes en viajes conjuntos y estancias suyas en Girona, Taruffo mostraba una extraña combinación de asertividad y humildad intelectual. Sus opiniones solían ser tajantes, firmes y expresadas hasta con vehemencia (así era su carácter), pero a la vez era capaz de tomar en serio el argumento contrario y no le temblaba el pulso para ponerse incluso a sí mismo como ejemplo de quien había sostenido una idea que ahora le parecía equivocada. Seriedad intelectual se llama.

Sus visitas en Girona fueron cada vez más frecuentes y por algún tiempo tuvimos la suerte de que se instalara a trabajar con nosotros por cuatro o cinco meses al año. Aquí dio clases en el máster de derecho de daños y estuvo también desde el inicio involucrado en la creación del máster en razonamiento probatorio, del que ha sido profesor hasta sus últimos días. Abandonaba totalmente la formalidad superflua y se dejaba llamar Michelino por algunos de nuestros atrevidos doctorandos, tomó tequila y ron mientras cantamos al son de una guitarra al final de nuestros encuentros de discusión de los avances de tesis doctorales. En Girona, como en otros lugares, dejó una escuela, cuyos integrantes esperamos honrar sus enseñanzas.

¹ Véase «La torre de Babel», de Michele Taruffo, más abajo.

Casi dos meses después de su prematuro adiós, escribo estas primeras líneas sobre quien fue un maestro y un gran amigo para mí. La tristeza y una sensación de orfandad vuelven a invadirme. Se nos fue uno de los mejores juristas de estos últimos 50 años, que cambió totalmente la comprensión de la prueba en el área cultural romano-germánica. Por eso, queremos dedicar a su memoria este número de la revista que de un modo u otro él inspiró.

* * *

Gracias a la generosidad de Cristina de Maglie, esposa de Taruffo, y de Luca Passanante, hemos tenido acceso a este pequeño cuento inédito, en el que Taruffo presenta una imagen irónicamente parroquial de los juristas, antítesis del modelo que él representaba. Con el permiso de la Prof. De Maglie lo publicamos aquí como póstumo homenaje.

Y para luchar contra la metáfora de la Torre de Babel que él presenta, lo hacemos en las cuatro lenguas oficiales de nuestra revista.

LA TORRE DE BABEL

Michele TARUFFO

No está clara la identidad de quienes vivieron y trabajaron en la Torre de Babel, especialmente después del desafortunado evento conocido como la “confusión de las lenguas”. Las fuentes son muy fragmentarias al respecto. En cualquier caso, antes del desafortunado evento, se encontraban allí gentes de muy diversa procedencia y de variadas costumbres lingüísticas. Habría también muchos intérpretes, dado que, antes de que el desafortunado evento se produjera, las cosas funcionaban razonablemente bien, hasta el punto de provocar en alguno el conocido gesto de exagerado orgullo.

Se puede también presumir que allí habría muchos juristas, aunque solo sea por los diversos y serios problemas jurídicos que surgen en la construcción de una relevante edificación manufacturada. Derechos de propiedad, licencias administrativas, licitaciones, contratos de suministros, derecho del trabajo y de las relaciones industriales y filosofía del derecho (en particular por las relaciones con la Divinidad) estaban necesariamente involucrados en la empresa. Debió tratarse de juristas políglotas (o que disponían de hábiles traductores) y expertos comparatistas, dado que -siempre según las fuentes- no se conoce que se dieran particulares problemas en esos ámbitos, a pesar de la presencia de personas de la más diversa procedencia étnica y cultural.

Vino después la confusión de las lenguas y fue un *shock* para todos. La construcción de la torre acabó mal, como todos sabemos, pero eso no importa ahora. Un aspecto poco estudiado hasta el momento es el referido a las consecuencias a largo

plazo que la confusión de las lenguas produjo sobre las personas que estaban allí y sus sucesores. En el fondo, enfrentarse de golpe a la situación de no entender el lenguaje de los colegas y colaboradores debió ser una experiencia trágica, como saben muchos de los que se aventuran a unas vacaciones “por su cuenta” en el extranjero. Debido siempre a la conocida carencia de fuentes directas, no sabemos nada sobre este relevante problema. Sin embargo, argumentando *ex post*, es posible llegar a algunas conclusiones plausibles.

Una de ellas es que no todos reaccionaron del mismo modo. Otra es que los que se encontraron peor fueron seguramente los juristas. Pero quizás es mejor explicar en qué se fundan estas audaces inferencias.

Muchos de los profesionales presentes supieron reaccionar de forma positiva frente al problema, o al menos lo hicieron -antes o después- sus sucesores. Una cuidadosa investigación de historia de la cultura científica podría ofrecer muchos detalles al respecto, pero centrémonos ahora en lo importante. El hecho es que ingenieros, físicos, médicos, expertos en economía y finanzas, comerciantes y prostitutas aprendieron rápidamente a entenderse, a pesar de que sus lenguas, las de sus colegas y sus clientes se hicieran diversas y casi incomprensibles. Se entendieron con gestos, muecas, símbolos matemáticos, gruñidos y suspiros, o aprendiendo inglés, pero lo hicieron. Más aún: de las condiciones “anteriores al desafortunado evento” conservaron la actitud de no distinguir las capacidades, valores, métodos, estilos y resultados en función de la lengua de las personas. Y peor aún: se inclinaron a pensar que si un problema es serio afectará a todos o a muchos, independientemente de la lengua de cada uno. Empezaron además a pensar que si un problema existe solo para quien habla una determinada lengua, con toda probabilidad no será un problema serio.

Al generalizarse esta tendencia, sin embargo, se corría el riesgo de frustrar el objetivo que Alguien se había propuesto con la confusión de las lenguas, esto es, hacer imposible la comunicación y, en consecuencia, imposibilitar la formación de una cultura común difundida, la elaboración de grandes y demasiado ambiciosos proyectos que implicaran a la humanidad entera y así sucesivamente.

Un solo grupo de personas se sustrajo a la tendencia general, fue capaz de conjurar el peligro y, por ello, se ganó la estima de Quien todo valora sabiamente: el de los juristas.

La confusión de las lenguas, y su separación, resultó para ellos sacra e inviolable. El uso de una sola lengua, a ser posible hablada por pocos, devino un principio metodológico fundamental. De ello derivó un ulterior principio, también fundamental, que reza más o menos así: “evita conocer lo que hacen y dicen los juristas que hablan otra lengua; si acaso no consigues evitarlo, no lo tomes en cuenta. Sepas en todo caso que, si violas este principio, no tendrás mucho éxito en tu carrera”.

Traducción de Jordi Ferrer Beltrán

MICHELE TARUFFO, *IN MEMORIAM*

Jordi Ferrer Beltrán

While we were finishing up this second issue of *Quaestio facti*, the news of Michele Taruffo's death fell upon us like a cold shower. He was undoubtedly one of the most important links in the causal chain that led to the foundation of this first journal dedicated to evidentiary reasoning in the area of the Roman-Germanic legal traditions. It is largely to his merit that evidential issues, which were once neglected and marginalized, are now increasingly at the center of analysis and academic study. Publications have multiplied exponentially, study programs have been created, substantial changes can be identified in judicial reasoning and are registered in case law and, little by little, in legislation itself too.

But none of this had happened yet when Taruffo wrote *Studi sulla rilevanza della prova* (1970) or his masterful *La prova dei fatti giuridici* (1992). From then on, and until his death, he authored several books and articles, and always wrote with the clarity and precision of an analytic philosopher. It is no coincidence that, in Pavia, he studied with Amedeo Conte and Uberto Scarpelli (who advised him to focus on the work of logical positivists while he was completing his first monograph) and that he was an avid reader of Giovanni Tarello.

But if there is something that characterized his work, it is his relentless ambition to cross borders. He did it in a physical sense all his life, because he loved to travel and discover remote places, far from the crowds, and venture into jungles for adventure, always in search of archaeological wonders that would allow him to better understand cultural origins and developments. At the table, he enjoyed experimenting all culinary traditions and, above all, chatting with the many friends he had all over the world. He used to say: «Don't invite me to travel, I may very well accept». This ambition to cross borders characterized his academic work too: he wanted to know

in depth legal institutions and, following the example of his master Vittorio Denti, he always thought that the best way to do it was through the comparative method; he would scrutinize the origins of legal institutions and their cultural contexts, and would know the debates, literatures and case law of other countries, trying to understand why the same problem is faced and solved differently in different countries.

But Taruffo didn't only cross physical-political boundaries between countries and legal systems: he tried to escape the, so to speak, «isolationist» way of working in the legal field. For this reason, he became interested in philosophy, epistemology, sociology and even neuroscience. He had no esteem, in his words, for the «exegetes of article 15, second paragraph, letter a)» of any piece of legislation, who seemed to him to produce short-sighted and uninteresting arguments. Commenting on some of his earliest works, a well-known procedural scholar once said: «Taruffo is very good, but he reads strange things». Michele, instead, loved to debate and, above all, the exercise of thought without restricted limits: «if a problem is serious, it affects everyone or many, regardless of their idiom [...]; if there is a problem only for those who speak a certain idiom, then it is very likely that it is not a serious problem at all»¹.

I had the opportunity to meet him personally, for the first time, in 1998 at the Italo-Spanish congress of analytic philosophy of law. I had just finished my doctorate, and Paolo Comanducci, to whom I am indebted for this among many other things, introduced me to Michele Taruffo. That same day we decided that I would work on the Spanish translation of *La prova dei fatti giuridici*. It was a long and tiring job, which lasted almost three years, but during that time we exchanged hundreds of emails and personal conversations; I have learned so much that I had the constant feeling of being very ignorant indeed. I was so impressed with his work that I continue to study those same topics to this day. In our dialogues, which became more and more frequent in the various trips we made together and during his stays in Girona, Taruffo showed a rare combination of assertiveness and intellectual humility. His views were usually blunt, firm and even vehemently expressed (this was his character), but at the same time he was able to take the opposite argument seriously and had no qualms in citing himself as an example of someone who had supported an idea that now seemed wrong. This is what intellectual integrity is all about.

Over the years, his visits to Girona became more frequent and for some time we were fortunate enough to have him working with us for four or five months a year. Here he lectured in our graduate course on civil liability law and contributed from the very beginning to the creation of our graduate course on evidential reasoning, where he taught until his final days. He completely abandoned unnecessary formalities and even allowed himself to be called Michelino by some of our most daring PhD students; he drank tequila and rum while we sang to the sound of a guitar at the end of our meetings to discuss the progress of our PhD students and their theses.

¹ See Taruffo's text below.

In Girona, as in many other places, he left a real school; and we all hope to honor his teachings.

Almost two months after his premature farewell, I write these first lines about a master and a great friend of mine. A great sadness pervades me, having the feeling of being orphaned. We lost one of the greatest jurists of the last fifty years, someone who profoundly changed the understanding of evidence law in the Roman-Germanic cultural area. This is why we want to dedicate to his memory this issue of the journal, which he inspired in so many ways.

* * *

Thanks to the generosity of Cristina de Maglie, Taruffo's wife, and Luca Passanante, we had access to this little unpublished story, in which Taruffo presents an ironically parochial image of lawyers, as an antithesis of the model that he represented. With the permission of Prof. de Maglie, we publish it here as a posthumous tribute.

And, in order to fight against the metaphor of the Tower of Babel that he presents, we do so in the four official languages of our journal.

Translated by Marco Segatti

THE TOWER OF BABEL

Michele TARUFFO

It is not really clear the identity of those who lived and operated on the Tower of Babel, especially after the unfortunate event known as the «confusion of the tongues». On this matter, sources are sketchy. However, before the unfortunate event there were people of different origins and various linguistic habits. There must have been many interpreters too; for, before the unfortunate event, things were running decently well, so much so that someone was inspired to the well-known gesture of exaggerated pride.

It is reasonable to assume that there were many lawyers, if only because of the serious legal problems that arise when one builds a significant construction. Property law, administrative tenders, supply contracts, subcontracts, labor and industrial relations law, and philosophy of law (particularly in relation with the Divinity) were necessarily involved in the ordeal. They must have been polyglot lawyers (or had access to skilled translators) and expert comparative scholars, since—again according to the sources—it appears that there weren't particular problems on this score, the presence of people from the most diverse ethnic and cultural backgrounds notwithstanding.

And then came the confusion of the tongues, and it shocked everyone. The construction of the tower ended badly, as everyone knows, but that's not what matters.

An aspect that has been neglected so far concerns the long-term consequences which the confusion of the tongues had for those who were there and for their progeny. After all, suddenly finding oneself to no longer understand the language of colleagues and collaborators must have been a tragic experience, as many of those who risk a «do-it-yourself» holiday abroad know. Again, given the known lack of direct sources, we know nothing about this not insignificant problem. However, arguing ex-post, it is possible to arrive at some plausible conclusions.

One of them is that not everyone reacted in the same way. Another is that lawyers were certainly the worst-off. But perhaps, it is better to explain on what these bold inferences are based.

Many of the professionals who were present reacted positively to the problem; or, at any rate, their successors did—sooner or later. An accurate research on the history of scientific culture might provide many details about this, but let's get to the point. The fact of the matter is that engineers, physicists, medical doctors, economics and finance wizards, traders and prostitutes rapidly learnt to understand each other anyway, even though their languages, those of their colleagues and of their clients had become different and almost incommunicable. They did it with gestures, grimaces, mathematical symbols, grunts and sighs, or by learning English, but they did it. Even more: of the conditions «prior the unfortunate event», they kept the habit of not differentiating one's capacities, values, methods, styles and results in light of his language. Worse still: they came to think that if a problem is serious, then it affects all or many, regardless the idiom of anyone. They came to think, too, that if a problem exists only for those who speak a particular idiom, then it is most likely that it is not a serious problem at all.

This tendency, when generalized, threatened the purpose that Someone had pursued with the confusion of the tongues; namely, to make communication impossible and, consequently, to make the formation of a widespread common culture, the development of large and too ambitious projects which would involve humanity as a whole, impossible too.

Only one group managed to evade the general trend, was able to avert the danger, and therefore became dear to the heart of the One who wisely judges everything: lawyers.

The confusion of the tongues, and their separation, became sacred and inviolable for them. The use of a single tongue, possibly spoken by a few, became a fundamental methodological problem. From this, a further and equally fundamental principle was logically derived. It goes something like this: «Avoid knowing what lawyers who speak another language do and say; if you can't really avoid it, do not take it into account. Know too that, if you violate this principle, you will not be very lucky in your career».

Translated by Marco Segatti

MICHELE TARUFFO, *IN MEMORIAM*

Jordi Ferrer Beltrán

Mentre stavamo chiudendo questo secondo numero di *Quaestio facti*, la notizia della morte di Michele Taruffo è caduta su di noi come una doccia d'acqua fredda. È stato senza dubbio uno degli anelli più importanti della catena causale che ha portato alla fondazione di questa prima rivista dedicata al ragionamento probatorio nell'area dei sistemi romano-germanici. È in gran parte dovuto a lui il cambiamento radicale di tendenza per cui le questioni probatorie, un tempo trascurate e marginalizzate, sono oggi sempre più al centro dell'analisi e dello studio accademico. Le pubblicazioni in tal senso si sono moltiplicate in modo esponenziale, sono stati creati programmi di studio, si notano cambiamenti sostanziali nella giurisprudenza e, poco a poco, anche nella stessa legislazione.

Ma niente di tutto questo era ancora successo quando Taruffo scrisse *Studi sulla rilevanza della prova* (1970) o il suo magistrale *La prova dei fatti giuridici* (1992). A partire da allora, e fino alla sua morte, egli è stato un autore prolifico di libri e articoli, scritti sempre con la chiarezza e precisione di un filosofo analitico. Non è un caso che, a Pavia, abbia studiato con Amedeo Conte e Uberto Scarpelli (che lo indirizzò alla lettura dei positivisti logici nei suoi primi studi sulla rilevanza della prova) e che fosse un avido lettore di Giovanni Tarello. Ma se c'è qualcosa che lo caratterizzava nel suo lavoro scritto, questa era l'ambizione permanente di attraversare i confini. Lo ha fatto in senso fisico per tutta la vita, perché amava viaggiare e scoprire luoghi remoti, lontani dalla folla, addentrarsi nelle giungle all'avventura, sempre alla ricerca di meraviglie archeologiche che gli permettessero di comprendere meglio le origini culturali. A tavola, gli piaceva sperimentare le tradizioni culinarie e, soprattutto, le conversazioni con gli amici che aveva in tutto il mondo. Non invitatemi a viaggiare, correreste il rischio che accetti, diceva.

Questa ambizione di varcare le frontiere ha contraddistinto anche il suo lavoro accademico: voleva conoscere a fondo le istituzioni giuridiche e, seguendo l'esempio del suo maestro Vittorio Denti, ha sempre pensato che il modo migliore per farlo fosse attraverso il metodo comparativo, scrutando le origini degli istituti e i loro contesti culturali, conoscendo i dibattiti, la letteratura e la giurisprudenza di altri paesi, cercando di capire perché lo stesso problema è affrontato e risolto in modo diverso in paesi differenti.

Ma Taruffo non ha solo cercato di oltrepassare i confini fisico-politici tra paesi e sistemi giuridici: ha anche provato a sfuggire al modo, per così dire, "isolazionista" di lavorare in campo giuridico. Per questo motivo si è interessato alla filosofia, all'epistemologia, alla sociologia e persino alle neuroscienze. Non aveva stima, nelle sue parole, dei "glossatori dell'articolo 15, secondo comma, lettera a)" di qualsiasi legge, che gli sembravano produttori di discorsi miopi e poco interessanti. Commentando alcune delle sue prime opere, un noto processualista disse: "Bravo Taruffo, ma legge cose strane". Michele, invece, amava il dibattito e, soprattutto, l'esercizio del pensiero senza limiti ristretti: "se un problema è grave, riguarda tutti o molti, indipendentemente dalla lingua [...]; se esiste un problema solo per chi parla una determinata lingua, allora con ogni probabilità non è un problema serio".¹

Ho avuto l'opportunità di incontrarlo personalmente, per la prima volta, nel 1998 al congresso italo-spagnolo di teoria analitica del diritto. Avevo da poco terminato il dottorato, e devo a Paolo Comanducci, tra le tante cose, il fatto di avermi presentato Michele Taruffo. Quello stesso giorno abbiamo deciso che mi sarei occupato della traduzione in spagnolo de *La prova dei fatti giuridici*. È stato un lavoro lungo e faticoso, durato quasi tre anni, ma durante quel periodo abbiamo avuto centinaia di scambi di e-mail e conversazioni personali e ho imparato talmente tanto da aver la frequente sensazione di essere assai ignorante. Sono rimasto così colpito del suo lavoro che continuo a studiare quegli stessi argomenti ancora oggi. Nei nostri dialoghi, che diventavano sempre più frequenti nei vari viaggi che abbiamo fatto in comune e nei suoi soggiorni a Girona, Taruffo ha mostrato una strana combinazione di assertività e umiltà intellettuale. Le sue opinioni erano solitamente schiette, ferme e persino espresse con veemenza (questo era il suo carattere), ma allo stesso tempo era in grado di prendere sul serio l'argomento contrario e non aveva alcuna remora a citarsi come esempio di qualcuno che avesse sostenuto un'idea che ora sembrava sbagliata. Si chiama serietà intellettuale.

Negli anni, le sue visite a Girona sono diventate più frequenti e per qualche tempo abbiamo avuto la fortuna che si stabilisse a lavorare con noi per quattro o cinque mesi all'anno. Qui ha tenuto lezioni al nostro *Master* di diritto della responsabilità civile e ha contribuito fin dall'inizio alla creazione del *Master* in ragionamento probatorio, di cui è stato docente fino ai suoi ultimi giorni. Abbandonava completamente le formalità superflue e addirittura si lasciava chiamare Michelino da alcuni dei no-

¹ V. il testo che segue.

stri dottorandi più audaci; beveva tequila e rum mentre cantavamo al suono di una chitarra al termine dei nostri incontri per discutere il percorso delle tesi di dottorato. A Girona, e in molti altri posti, ha lasciato una vera e propria scuola; e tutti noi speriamo di onorare i suoi insegnamenti.

Quasi due mesi dopo il suo prematuro addio, scrivo queste prime righe su un mio maestro e grande amico. Mi pervade una grande tristezza e la sensazione di essere rimasto orfano. Ci ha lasciato uno dei migliori giuristi degli ultimi cinquant'anni, che ha cambiato in profondità la comprensione dei problemi probatori nell'area culturale romano-germanica. Per questo vogliamo dedicare alla sua memoria questo numero della rivista, che in un modo o nell'altro lui stesso ha ispirato.

* * *

Grazie alla generosità di Cristina de Maglie, moglie di Taruffo, e Luca Passanante, abbiamo avuto accesso a questo piccolo racconto inedito, in cui Taruffo presenta un'immagine ironicamente campanilistica dei giuristi, come antitesi del modello che lui stesso rappresentava. Con il permesso della prof.ssa De Maglie, lo pubblichiamo come un omaggio postumo.

E per lottare contro la metafora della Torre di Babele che ci presenta, lo pubblichiamo nelle quattro lingue ufficiali della nostra rivista.

Traduzione di Marco Segatti

LA TORRE DI BABELE

Michele TARUFFO

Non è ben chiara l'identità di coloro che vissero e operarono sulla Torre di Babele, soprattutto dopo lo sfortunato evento noto come "confusione delle lingue". Le fonti sono molto lacunose a questo proposito. Comunque, prima dello sfortunato evento vi si trovava gente di assai diversa provenienza e di svariate abitudini linguistiche. Dovevano anche esserci molti interpreti dato che, prima che lo sfortunato evento si verificasse, le cose funzionavano discretamente, tanto da indurre taluno al ben noto gesto di esagerato orgoglio.

È comunque lecito presumere che vi fossero molti giuristi, se non altro a causa dei vari gravi problemi giuridici che sorgono nel caso di costruzione di un rilevante manufatto edile. Diritto di proprietà, concessioni amministrative, appalti, contratti di fornitura, diritto del lavoro e delle relazioni industriali, e filosofia del diritto (in particolare per le relazioni con la Divinità) erano necessariamente coinvolti nell'impresa. Doveva trattarsi di giuristi poliglotti (o con abili traduttori) ed esperti comparatisti, poiché –

sempre stando alle fonti – non risulta che siano sorti particolari problemi sotto questo profilo, malgrado la presenza di genti della più diversa provenienza etnica e culturale.

Venne poi la confusione delle lingue, e fu uno shock per tutti. La costruzione della torre finì male, come tutti sanno, ma non è questo che importa. Un aspetto poco indagato finora riguarda le conseguenze di lungo periodo che la confusione delle lingue produsse su coloro che c'erano, e sui loro successori. In fondo, trovarsi di colpo a non capire più il linguaggio di colleghi e collaboratori doveva essere una esperienza tragica, come sanno molti di coloro che azzardano una vacanza “fai-da-te” all'estero. Sempre per la nota carenza di fonti dirette, non sappiamo nulla su questo non irrilevante problema. Tuttavia, argomentando *ex post*, è possibile pervenire ad alcune plausibili conclusioni.

Una di queste è che non tutti reagirono allo stesso modo. Un'altra è che a trovarsi peggio di tutti furono sicuramente i giuristi. Ma forse è meglio spiegare su che cosa si fondano queste ardite inferenze.

Molti dei professionisti presenti reagirono in modo positivo al problema, o almeno lo fecero – prima o poi – i loro successori. Un'accurata ricerca di storia della cultura scientifica potrebbe fornire molti dettagli proposito, ma veniamo alla sostanza. Il fatto è che ingegneri, fisici, medici, esperti di economia e finanza, commercianti e prostitute impararono rapidamente ad intendersi lo stesso, benché le loro lingue, quelle dei loro colleghi, e dei loro clienti, fossero diventate diverse e pressoché incommunicabili. Lo fecero a gesti, a smorfie, a simboli matematici, a grugniti e a sospiri, o imparando l'inglese, ma lo fecero. Più ancora: delle condizioni di “prima dello sfortunato evento” conservarono l'attitudine a non differenziare capacità, valori, metodi, stili e risultati, in funzione della lingua delle persone. Peggio ancora: diventarono inclini a pensare che se un problema è serio, esso riguarda tutti o molti, indipendentemente dalla lingua di ognuno. Cominciarono inoltre a pensare che, se un problema esiste solo per chi parla una determinata lingua, allora con ogni probabilità non è un problema serio.

Questa tendenza, generalizzandosi, rischiava tuttavia di vanificare lo scopo che Qualcuno si era proposto con la confusione delle lingue, ossia di rendere impossibile la comunicazione e, di conseguenza, di rendere impossibile la formazione di una cultura comune diffusa, l'elaborazione di grandi e troppo ambiziosi progetti che coinvolgessero l'intera umanità, e così via.

Un solo gruppo di persone si sottrasse alla generale tendenza, fu capace di sventare il pericolo, e perciò divenne caro al cuore di Colui che tutto saggiamente valuta: i giuristi.

La confusione delle lingue, e la loro separazione, divenne per loro sacra ed inviolabile. L'uso di una sola lingua, possibilmente parlata da pochi, divenne un principio metodologico fondamentale. Da questo venne logicamente derivato un ulteriore principio altrettanto fondamentale, che suona all'incirca così: “evita di conoscere ciò che fanno e dicono i giuristi che parlano un'altra lingua; se proprio non riesce ad evitarlo, non tenerne conto. Sappi inoltre che, se violi questo principio, non sarai molto fortunato nella carriera”.

MICHELE TARUFFO, *IN MEMORIAM*

Jordi Ferrer Beltrán

Enquanto fazíamos o trabalho de fechamento deste segundo número de *Quaestio Facti*, caiu-nos como um balde de água fria a notícia do falecimento de Michele Taruffo. Ele foi, sem dúvida, um dos elos mais importantes da corrente causal que levou à fundação desta primeira revista dedicada ao raciocínio probatório na área dos sistemas romano-germânicos. A ele se deve, em grande medida, a mudança radical de tendência, que vai de um certo abandono e desapareço pela relevância dos problemas probatórios até a que, hoje em dia, esses estejam cada vez mais no foco da análise e do estudo nesses sistemas. Multiplicaram-se exponencialmente as publicações a respeito, criaram-se programas de estudo, apresentaram-se mudanças substanciais na jurisprudência e, pouco a pouco, também na legislação.

Nada disso estava acontecendo quando Taruffo escreveu *Studi sulla rilevanza della prova* (1970), ou seu magistral *La prova dei fatti giuridici* (1992). Desde esses anos até seu falecimento ele foi um autor prolífico em livros e artigos, sempre escritos com a preocupação com a clareza e a precisão, própria de um filósofo analítico. Não por acaso quis estudar com Amedeo Conte em Pavia, com Uberto Scarpelli em Milão (que lhe recomendou a leitura dos positivistas lógicos para seus primeiros estudos sobre a relevância da prova) e foi ávido leitor de Giovanni Tarello. Entretanto, se há algo que o caracteriza em sua obra escrita é a permanente ambição de traspasar fronteiras. Fez isso no sentido físico durante toda a sua vida, pois amava viajar e descobrir lugares recônditos, distantes de aglomerações, ingressando em selvas por aventura, sempre na busca de maravilhas arqueológicas que lhe permitissem compreender melhor as origens culturais. Desfrutava de experimentar tradições culinárias e, sobretudo, das conversas à mesa com os amigos que fez por todo o mundo. “Não me convidem para viajar, pois correm o risco de que eu aceite”, dizia.

Essa ambição de traspasar fronteiras ele levou também a seu trabalho teórico: queria entender as instituições jurídicas com profundidade e, seguindo o exemplo de seu mestre, Vittorio Denti, sempre pensou que a melhor maneira de fazer isso era com método comparado, esquadrinhando suas origens e contextos culturais, conhecendo os debates, a literatura e a jurisprudência de outros países, compreendendo o porquê de um mesmo problema ser enfrentado e resolvido de modo diferente em uns países e em outros. Mas Taruffo não pretendeu somente traspasar fronteiras físico-políticas entre países e sistemas jurídicos, mas também fugir do modo isolacionista de trabalhar no âmbito jurídico. Por isso, interessou-se por filosofia, economia, epistemologia, sociologia e até por neurociência. Desgostava, em suas palavras, dos glosadores do artigo 15, segundo parágrafo, letra a) de qualquer lei, que lhe pareciam produtores de discursos míopes e carentes de interesse. Comentando algumas de suas primeiras obras, um conhecido processualista italiano disse: “Taruffo é bom, ainda que leia coisas estranhas!”. Taruffo, de seu turno, amava o debate e, em especial, o exercício de pensar sem esses limites estreitos: “se um problema é sério, isso diz respeito a todos ou a muitos, independentemente da língua de cada um. (...) [S]e um problema existe somente para quem fala uma determinada língua, então muito provavelmente não se trata um problema sério”¹.

Tive a oportunidade de conhecê-lo pessoalmente no ano de 1998, no congresso ítalo-espanhol de teoria analítica do direito. Ter me apresentado a Michele Taruffo, sendo eu um jovem doutor, é uma das muitas coisas que devo a Paolo Comanducci. Nesse mesmo dia concordamos que eu me encarregaria da tradução ao espanhol de *La prova dei fatti giuridici*. Foi um árduo e grande trabalho de quase 3 anos, mas durante esse tempo tivemos centenas de intercâmbios de mensagens e conversas pessoais, e aprendi tanto que me senti muito ignorante. Entusiasmou-me a tal ponto seu trabalho que hoje sigo estudando esses mesmos temas. Em nossos diálogos, que se tornaram cada vez mais frequentes em viagens conjuntas e em estadas suas em Girona, Taruffo mostrava uma singular combinação de assertividade e humildade intelectual. Suas opiniões costumavam ser contundentes, firmes e expressadas até mesmo com veemência (assim era sua personalidade), mas, ao mesmo tempo, era capaz de levar a sério o argumento contrário e sua mão não tremia para incluir a si mesmo como exemplo de quem sustentara uma ideia que agora lhe parecia equivocada. Isso se chama seriedade intelectual.

Suas visitas a Girona foram cada vez mais frequentes e por algum tempo tivemos a sorte de que se instalasse para trabalhar conosco por quatro ou cinco meses por ano. Aqui deu aulas no mestrado de direito de danos e esteve também, desde o início, envolvido na criação do mestrado em raciocínio probatório, do qual foi professor até seus últimos dias. Abandonada totalmente a formalidade supérflua, deixava-se chamar de Michelino por alguns de nossos atrevidos doutorandos; tomou tequila e rum enquanto cantamos ao som de um violão, ao final de nossos encontros de discussão

¹ Veja-se «A torre de Babel», de Michele Taruffo, mais abaixo.

dos avanços de teses doutorais. Em Girona, como em outros lugares, deixou uma escola, e nós, seus integrantes, esperamos honrar seus ensinamentos.

Quase dois meses depois de seu prematuro adeus, escrevo essas primeiras linhas sobre quem para mim foi um mestre e um grande amigo. A tristeza e uma sensação de orfandade tornam a invadir-me. Perdemos um dos melhores juristas dos últimos 50 anos, que mudou totalmente a compreensão da prova na área cultural romano-germânica. Por isso, queremos dedicar à sua memória este número da revista que, de um modo ou outro, ele inspirou.

* * *

Graças à generosidade de Cristina de Maglie, esposa de Taruffo, e de Luca Passanante, tivemos acesso a este pequeno conto inédito, em que Taruffo apresenta uma imagem ironicamente paroquial dos juristas, antítese do modelo que ele representava. Com a autorização da Profa. De Maglie publicamos aqui, como homenagem póstuma.

E, para lutar contra a metáfora da Torre de Babel que ele apresenta, fazemos isso nas quatro línguas oficiais de nossa revista.

Tradução de Vitor de Paula Ramos

A TORRE DE BABEL

Michele TARUFFO

Não é totalmente clara a identidade daqueles que viveram e laboraram na Torre de Babel, sobretudo depois do desafortunado evento conhecido como “confusão das línguas”. As fontes são muito lacunosas a respeito. De qualquer sorte, antes do desafortunado evento ali havia gente de proveniência bastante diversa e de diferentes hábitos linguísticos. Devia haver também no local muitos intérpretes, visto que, antes de que o desafortunado evento ocorresse, as coisas funcionavam discretamente, a ponto de induzir alguém ao bem conhecido gesto de exagerado orgulho.

Seja como for, é lícito presumir que ali houvesse muitos juristas, senão por outra razão, por conta dos vários graves problemas jurídicos que surgem no caso de construção de um relevante edifício feito à mão. Direito de propriedade, concessões administrativas, empreitadas, contratos de fornecimento, direito do trabalho e das relações industriais, e filosofia do direito (em particular pelas relações com a Divindade) estavam necessariamente envolvidos no desígnio. Deveria se tratar de juristas políglotas (ou com tradutores hábeis) e comparatistas experientes, visto que – sempre de acordo com as fontes – não resulta que tenha havido problemas particulares

sob esse aspecto, apesar da presença de povos da mais diversa proveniência étnica e cultural.

Veio, então, a confusão das línguas e foi um choque para todos. A construção da torre acabou mal, como todos sabem, mas isso não é o que importa. Um aspecto pouco questionado até hoje diz respeito às consequências de longa duração que a confusão das línguas produziu naqueles que lá estavam, e em seus sucessores. No fundo, ver-se repentinamente não entendendo mais a linguagem dos colegas e colaboradores devia ser uma experiência trágica, como sabem muitos daqueles que amargam um período de férias “faça-você-mesmo” no exterior. Ainda por conta da carência de fontes diretas, nada sabemos sobre esse não irrelevante problema. Todavia, argumentando *ex post* é possível chegar a algumas plausíveis conclusões.

Uma dessas é que nem todos reagiram do mesmo modo. Outra é que os que se viram pior de todos foram seguramente os juristas. Mas talvez seja melhor explicar no quê se fundam essas impávidas inferências.

Muitos dos profissionais presentes reagiram de modo positivo ao problema, ou pelo menos o fizeram – mais cedo ou mais tarde – os seus sucessores. Uma investigação cuidadosa de história da cultura científica poderia fornecer muitos detalhes a propósito, mas passemos ao cerne. O fato é que engenheiros, físicos, médicos, *experts* de economia e finanças, comerciantes e prostitutas aprenderam rapidamente a, de qualquer sorte, se entenderem, apesar de a sua língua, a de seus colegas e a de seus clientes terem se tornado diferentes e quase incomunicáveis. Fizeram-no com gestos, caras, símbolos matemáticos, grunhidos e suspiros, ou aprendendo inglês, mas o fizeram. Mais do que isso: algumas condições de “antes do desafortunado evento” conservavam a postura de não diferenciar capacidades, valores métodos, estilos e resultados, em função da língua das pessoas. Pior do que isso: tornaram-se inclinados a pensar que se um problema é sério, isso diz respeito a todos ou a muitos, independentemente da língua de cada um. Começaram, ademais, a pensar que, se um problema existe somente para quem fala uma determinada língua, então muito provavelmente não se trata de um problema sério.

Generalizando-se essa tendência, todavia, corria-se o risco de tornar vão o escopo que Alguém pretendia com a confusão das línguas, ou seja, de tornar impossível a comunicação; e, conseqüentemente, de tornar impossível a formação de uma cultura comum difusa, a elaboração de grandes e demasiadamente ambiciosos projetos que envolvessem toda a humanidade, e assim por diante.

Somente um grupo de pessoas apartou-se da tendência geral, sendo capaz de afastar o perigo e, por isso, tornando-se caro ao coração Daquele que tudo valora sabiamente: os juristas.

A confusão das línguas, e a separação dessas, tornou-se para esses sagrada e inviolável. O uso de uma só língua, possivelmente falada por poucos, tornou-se um princípio metodológico fundamental. Disso se derivou logicamente um ulterior princípio, tão fundamental quanto o anterior, que soa mais ou menos assim: “evita

de conhecer aquilo que fazem e dizem os juristas que falam outra língua; se realmente não consegues evitar isso, não lhe leva em conta. Sabe, ainda, que se violares este princípio não serás muito exitoso na carreira”.

Tradução de Vitor de Paula Ramos

ENSAYOS

A CRITICAL METHOD FOR THE EVALUATION OF EVIDENCE: FRANÇOIS GORPHE'S (1889-1959) CONTRIBUTION TO A SCIENCE OF PROOF À LA FRANÇAISE

Olivier Leclerc

CNRS Senior researcher

Center for Critical Legal Research (CERCRID UMR 5137)

University of Lyon, University Jean-Monnet (France)

olivier.leclerc@cnrs.fr

ABSTRACT: In his 1924 book *La critique du témoignage*, the French judge François Gorphe (1889-1959) advocated the need that evaluation of evidence be based on the best scientific knowledge available. Drawing upon an extensive knowledge of the experiments conducted by contemporary forensic psychologists, Gorphe urged judges to consider forensic psychology when interrogating witnesses and, ultimately, when evaluating evidence. Gorphe imposed himself as one of the main promoters of a science of proof in France. This paper examines Gorphe's contributions and the audience they received among forensic psychologists and criminologists in France and abroad. Despite their originality and anchoring in international scientific discussions, Gorphe's views have not encouraged the creation of an academic domain interested in evidential reasoning in French law faculties. The paper concludes with an examination of the reasons for the limited posterity of Gorphe's approach.

KEYWORDS: history of evidence theory; evidential reasoning; science of proof; forensic psychology; psychology of testimony.

SUMMARY: 1. INTRODUCTION.—2. BRIDGING FORENSIC PSYCHOLOGY AND THE ANALYSIS OF EVIDENCE.—3. INTERESTING THE SPECIALISTS OF CRIMINAL LAW AND CRIMINOLOGY IN FORENSIC PSYCHOLOGY.—4. TRAINING SOCIETY TO GIVE TESTIMONY, TRAINING JUDGES TO EVALUATE WITNESSES.—5. CONCLUSION.

RECOMMENDED CITATION: LECLERC, O., 2020: «A Critical Method for the Evaluation of Evidence: François Gorphe's (1889-1959) Contribution to a Science of Proof *à la française*», in *Quaestio facti*, 2: 29-52. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22389

«Dans une conception rationnelle de la justice, et notamment des preuves, la conviction qui emporte la décision doit être l'aboutissement logique d'un examen analytique des faits et d'une appréciation critique des éléments de preuve. Elle passe ainsi de l'état de simple croyance subjective à celui d'une véritable connaissance objective, communicable et contrôlable»¹.

1. INTRODUCTION²

In 1928, at the height of his academic renown, the American specialist in the law of evidence and criminologist John Henry Wigmore (1863-1943) gave a laudatory account³ of a book recently published in France by the publisher Dalloz, entitled *La critique du témoignage* [The Critique of the Testimony]⁴. Wigmore praised it lavishly: «May we have such a book in English!». The book, published in 1924 and quickly reissued for a second edition, was based on the thesis of a young French magistrate, François Gorphe (1889-1959). Wigmore's praise was not without ulterior motives, for Gorphe's book explicitly followed the path opened up by Wigmore himself⁵, who found a relay for his ideas in France, while reaffirming the anteriority of his own work: «And yet in France and the Continent generally the profession at large has not been appreciably in advance of the Anglican profession in realizing the possibilities of psychological and logical science as applied to judicial evidence. This is the first French treatise on the subject». Indeed, Gorphe later returned the compliment to Wigmore, noting in relation to *The Principles of Judicial Proof* (1913), «There is no equivalent in Continental law»⁶.

¹ GORPHE, 1947: 18.

² I am grateful to Catherine Fillon for her valuable help in identifying archives on François Gorphe and for her extensive knowledge on the history of the law faculties in France. Any mistake obviously remains all mine.

³ WIGMORE, 1928: 605.

⁴ GORPHE, 1924.

⁵ WIGMORE, 1913b (in his last published pieces Gorphe referred to the third edition published in 1937 under the title *The Science of Judicial Proof* [WIGMORE, 1937]).

⁶ GORPHE, 1947: 21.

François Gorphe was a pioneering author in France in the development of a science of proof. He affirmed the need for judges to draw on the knowledge developed in the field of psychology to reinforce the rigour of their investigations, suspects interrogation techniques and, eventually, evaluation of the evidence presented to them. As Gorphe put it, French legislation and case-law recognized that judges have «an absolute power of appreciation» of evidence: «Here is the judge left by the legislator to his own means, to the “flair” that he has or is supposed to have, we would say to his divination instinct. This is truly a heavy honour, indeed!»⁷. Therefore, the (incomplete but far-reaching) suppression of the legal proof system in France has broken «the old chains», but also created the risk that the judges’ reasoning be guided by «emotional intuition»⁸. In order to guide this work and to avoid errors of reasoning as much as possible, «it is indeed to science that one must deliberately resort»⁹.

Gorphe was born in Bordeaux (France) on 14 June 1889. He attended secondary school in Périgueux, his parents having settled in the Dordogne (commune of Douzillac), in the Southwestern part of France. After passing his baccalaureate at the age of 16 (1905), he studied law at the University of Bordeaux, from which he graduated in 1908. In the years preceding the First World War, he continued his legal training with a view to exercising a judicial profession (he graduated from the *Institut Pratique de Droit* of Bordeaux in 1913, then worked successively as a trainee lawyer and a trainee judge at the Court of Bordeaux from that date¹⁰), in parallel with a job at the *Banque de France*, obtaining a postgraduate degree [*diplôme d’études supérieures*] in philosophy and fulfilling his military obligations (Gorphe entered the army in 1910 and was placed on auxiliary service. From 1912 he was placed on standby duty until the general mobilisation decree of 1 August 1914)¹¹. Following his demobilisation in 1919, Gorphe was appointed deputy judge at the civil court of Libourne. His career¹² continued in various courts of first instance in western France: Marennes (1922-1926), Bressuire (1926), and Poitiers (1926-1927). In 1927, he was appointed judge at the Lille court of first instance, where he practised until 1932, when he became a judge at the Poitiers Court of Appeal, then President of Chamber until his retirement on 10 June 1959.

During his career, Gorphe’s worked mainly as an active sitting judge, but during the years 1920-1926 he was an investigating judge and, for a short period of time

⁷ GORPHE, 1924: 11.

⁸ GORPHE, 1947: 17. Wigmore had made a very similar assumption earlier: «For centuries, lawyers and judges had evidenced and proved by the artificial numerical system; they had no training in any other, – no understanding of the living process of belief; in consequence, when “legal proof” was abolished, they were unready, and judicial trials have been carried on for a century past by uncomprehended, unguided, and therefore unsafe mental processes. Only in recent times, under the influences of modern science, are they beginning to develop a science of proof» (WIGMORE 1913a: 78).

⁹ GORPHE, 1924: 11.

¹⁰ Arch. Nat. 19800035/417/55805.

¹¹ *Military service record* (regimental number: 1909/86).

¹² *Annuaire rétrospectif de la magistrature, XIXe-XXe siècles*, V° François Joseph Gorphe.

in 1926, prosecutor. During this period he also prepared his first doctoral thesis in political and economic sciences¹³. The dissertation, entitled *La critique du témoignage*, was defended at the University of Paris in 1924 and published shortly after by Dalloz¹⁴. In 1928, Gorphe defended a second thesis, in legal sciences, entitled *Le principe de bonne foi* [The Principle of Good Faith], also published by Dalloz in 1928¹⁵. If his second thesis earned Gorphe the fairly prestigious *Limandour Prize* from the *Académie des Sciences morales et politiques* (Legislation Section)¹⁶ on the proposal of Henri Capitant¹⁷, it was *La critique du témoignage* that established Gorphe's reputation. In 1927, Dalloz published a second edition of the book—which was unusual for a doctoral thesis—which resulted in the publication of reviews in scientific journals both in France and abroad. In this book, Gorphe defended the idea that although testimony is fallible, it is possible to develop a critical method of analysis that allows for a determination of its value. His analyses were nourished by a vast literature in forensic psychology, in French and other languages, and by the personal experience he had accumulated as an investigating judge in courts of first instance since 1920. According to Gorphe, the assessment of testimony cannot be handled solely as a matter of logic (internal coherence, concordance or contradiction between testimonies), but also of psychology, «especially affective psychology»¹⁸. The mental habits of judges play a central role in the evaluation of evidence. However, testimony may lead to truth as much as to error. Hence, testimonial criticism must classify and analyse errors: «Errors, like diseases, must, in our opinion, be carefully analysed and determined in their character, causes and symptoms; their nosology, etiology and semiology must be established»¹⁹. This endeavour led Gorphe to analyse the types of errors affecting testimony according to their object, the cognitive functions to which they relate (perception, memory, imagination, judgement, will), their form (errors stemming from substitution, modification, transposition, fusion or confusion, dissociation, entrenchment, addition, invention, errors of comprehension, errors in estimating quality or quantity), their causes (inventions, hallucinations, confabulations, misinterpretations, illusions). The identification of these errors by judges (Gorphe coined the term «testimonial diagnosis») presupposes not only that they have good

¹³ In 1856, the doctorate in law was divided into two separate diplomas, one in political and economic sciences the other in legal science. Defending two theses was not uncommon in law faculties. Holding a doctorate was not compulsory for judges, yet it could boost their career. In contrast, those hoping to become academics would usually defend two theses. As he already worked as a judge, Gorphe had no obligation to even defend one thesis, not to mention defending two of them. At some point, he might have considered joining the university although no other indication of such a project has been found.

¹⁴ GORPHE, 1924.

¹⁵ GORPHE, 1928.

¹⁶ *Séances et travaux de l'Académie des sciences morales et politiques*, 1929, p. 387.

¹⁷ Capitant sat, together with Lévy-Ullmann and Julliot de la Morandière, in the thesis defence jury. The three of them had a certain academic prestige.

¹⁸ GORPHE, 1924: 40.

¹⁹ *Ibid.*, 54.

knowledge of the possible errors but also that they apply analytical methods to detect them. Gorphe thus presented a wide range of appropriate interrogation techniques and examination methods available to judges (mental examination, psychological tests, psycho-diagnostic apparatus, method of associating ideas, psychometric tests). The analysis of testimony also requires, Gorphe asserted, an analysis of the personality of the witness, since although the witness is an «instrument of proof [...] infinitely superior to those constructed by men, it does not have the security or precision of such an instrument»²⁰. It is therefore incumbent on judges to analyse the morality of witnesses, their intellectual capacities, affective dispositions (interest, passions) and psychological states (alterations and mental illnesses). Likewise, the material conditions in which a witness perceived the facts he or she is reporting must be carefully examined as they bear a decisive influence on sensory perception and recognition. As experimental psychology had demonstrated, witness identifications are «an essentially affective phenomenon, residing in an impression of the familiar, of the already known, before being expressed in a judgement of identity»²¹. Gorphe also examined memorization processes and warned that judges must evaluate how the erasure of memories affects the reliability of the testimony offered by a witness. Gorphe finally recommended practical techniques for hearing witnesses and evaluating the value of their statements based on the lessons of experimental psychology.

Throughout his career as a magistrate, Gorphe continued to promote the approach initiated in his 1924 work, regularly publishing articles and books until his death in 1959. In 1947, Gorphe published a book entitled *L'appréciation des preuves en justice: essai d'une méthode technique* [Assessing Evidence in Court: a technical method]²², in which he generalized the analysis carried out in 1924 on testimony by including all the evidence submitted to judges. In this book, which is the most complete of those published by Gorphe²³, he took up the essence of the demonstration he proposed in 1924; completing analyses relating to the logical investigation of evidence (here mostly based on Wigmore [WIGMORE, 1913]). Gorphe not only presented the psychological theories useful for the analysis of testimony, but he also analysed the scientific principles that must guide the evaluation of other evidence like written documents, confessions and statements, clues and presumptions. Shortly afterwards, in 1952, Gorphe published a last book entitled *Les Décisions de justice: Etude psychologique et judiciaire* [Judicial Decisions: A Psychological and Judicial Study]²⁴, which essentially combines his previous work on the evaluation of evidence with insights on the psychology of judges when performing the act of judging. Gorphe's analysis on

²⁰ *Ibid.*: 120.

²¹ *Ibid.*: 312.

²² GORPHE, 1947.

²³ For this book Gorphe was awarded the *Dupin Aîné* prize of the *Académie des Sciences morales et politiques*, together with Paul Roubier's *Théorie générale du droit* (*Revue des travaux de l'Académie des sciences morales et politiques et comptes rendus de ses séances*, 1947, p. 141 and p. 181).

²⁴ GORPHE, 1951.

the reasoning of judges focused mainly on the interpretation of the law and on legal qualification and proved to be less innovative. In this book, Gorphe mostly referred to the professional discipline that he claimed judges must adhere to (e.g. rising above individual preferences, the love of justice). Although Gorphe extended his analysis not only to the evaluation of evidence but also to judicial decision-making, his arguments on legal reasoning and the psychology of judges didn't attract as much interest as his work on evidence.

Gorphe's publications display a real continuity in his analysis. The author's primary originality lies in the fact that he has been able to overcome two limitations inherent in the literature on which it was based: beyond mere testimony, he generalized the evaluation to all types of evidence; beyond the mere appreciation of evidence²⁵, he investigated the contributions of forensic psychology to the analysis of all operations involved in the determination of facts, including the decision of judges. In spite of the limits and inadequacies that can be found in Gorphe's proposals, his project remains highly topical and open to potential development. Gorphe himself very quickly became clearly aware of the ambition and scope of his project—a project whose publication would take him a lifetime. Indeed, most of his analyses had been in fact developed, if not published, before World War II: his 1947 book was written before the war²⁶ and a good part of the examination published in his 1952 book had been published separately in the early 1930s²⁷.

Gorphe's analysis reflects the way in which developments of positivism in the social sciences in the late 19th century pervaded legal scholarship. In a 1950 newspaper article, the then-famous specialist of international criminal law and French judge at the Nuremberg trials, Henri Donnedieu de Vabres²⁸ referred to Gorphe as opening a new scientific era in the analysis of evidence:

In the history of repression, a scientific phase follows the metaphysical phase and the legal phase. The Paris congress [second International Congress of Criminology, 1950] has only broadened a development that the recent works of Mr. Gorphe, president of chamber at the Poitiers Court of Appeal, have highlighted with regard to the theory of evidence. This movement, born of free scientific research, creates guarantees against arbitrariness. It is itself a generator of freedom²⁹.

²⁵ Twining emphasized the «narrowness of focus» of Bentham, and Wigmore after him, who concentrated on «formal determinations of liability and guilt in respect of questions of fact by judges or juries in trial» and mostly left aside the determination of facts «at other phases of legal process (such as negotiation and sentencing) or other arenas (such as arbitration and administrative tribunals)» (TWING, 1985: 176). Although Gorphe did not deliver such an all-encompassing vision, he sought to highlight the contributions of forensic psychology at all stages of the trial.

²⁶ GORPHE, 1947: fn 1, p. 13.

²⁷ Some articles are simply repeated in *Les décisions de justice* (1952) (GORPHE, 1931b, 1931c), others are substantially rewritten while remaining very similar in content (GORPHE, 1931a, 1931d, 1935a). An initial footnote in these articles indicates that they are part of a book «in preparation» or «to be published». Twenty years passed before the announced book eventually appeared.

²⁸ HALPERIN, 2007.

²⁹ *Le Monde*, 27 Sept. 1950.

The wording alludes to the law of three stages of Auguste Comte³⁰, but it also takes up the ideas developed by the Italian sociologist Enrico Ferri (1856-1929), who had a notable influence on criminologists at the beginning of the 20th century and heralded the time of proof in justice based on science³¹. Gorphe himself, moreover, inscribed his work in continuity with the then-influential historical method defended by Langlois and Seignobos³² and the efforts they made to give historiography a scientific foundation. Gorphe saw the work of jurists on testimony as a variation of the approach undertaken by historians³³: «historians, who have preceded jurists in the criticism of testimony, have established a triple critique of documents which is based on three main modes of alteration of the truth: the critique of interpretation, the critique of sincerity and the critique of exactitude»³⁴.

Singular magistrate, Gorphe is one of the few in France who actively participated in the intellectual ferment of the years 1900-1930 in the development of a science of proof. He displayed a vast knowledge of experimental psychology, both from European and American sources, as the numerous references he quoted attest to (more than 600 in *La critique du témoignage*), not only in French but also in German, Italian, and English. In his numerous writings, on which he worked in addition to his activity as a judge, Gorphe advocated that the experimental knowledge of forensic psychology can improve forensic techniques and must be taken into account by judges when evaluating evidence (section 1). In Gorphe's view, «such a method has a general scope and is applicable in any judicial field». Nevertheless, he admitted, «its importance is much greater in criminal matters»³⁵. Therefore, Gorphe devoted much of his efforts to interesting criminologists and criminal law specialists, particularly the post-World War II promoters of the «social defence», in contributions to forensic psychology (section 2). In so doing, Gorphe not only highlighted the contributions of forensic psychology for evidence; he also advocated changes in judicial training and a widespread diffusion of psychology-based techniques in courts, and even beyond. Hence, Gorphe posited his contribution at an equal distance from a theoretical contribution to the evaluation of evidence and a repertoire of professional technique, with all the awkwardness associated with this interstitiality (section 3).

³⁰ COMTE, 1830.

³¹ FERRI, 1884, translated in French in 1905. Ferri identified five stages in the evolution of evidence systems: ethnic, religious, legal, sentimental and scientific (FERRI, 1905: 507).

³² LANGLOIS, SEIGNOBOS, 1898.

³³ This idea had already been supported by the psychologist and physiologist Jean Larguier des Bancel's in a text Gorphe had read and cited (GORPHE, 1924: 14) with an inaccurate reference. Larguier des Bancel's mention of Langlois and Seignobos appears in LARGUIER DES BANCEL'S, 1905.

³⁴ GORPHE, 1924: 57. The same idea is advocated again in GORPHE 1947b: 146, with further reference to BERNHEIM, 1903 and DE SMEDT, 1883: «Historians were the first to feel the need for [a scientific analysis of the testimony]: it is true that nothing is less certain than historical accounts, based on disappeared, unverifiable testimonies and [testimonies] not dictated by a pure concern for truth».

³⁵ GORPHE, 1947: 23.

2. BRIDGING FORENSIC PSYCHOLOGY AND THE ANALYSIS OF EVIDENCE

At the turn of the 20th century, social psychology developed as a specific branch of psychology. Under the title of forensic psychology, part of this research was devoted to the study of psychological phenomena in the context of trials. The psychology of testimony was one of the main objects of analysis. In several European countries, physicians and psychologists set up experimental protocols in order to assess the reliability of testimony. In 1902, the German psychologist William Stern (1871-1938) published a study on the psychology of testimony³⁶, followed by a book³⁷ and shortly after by the creation of an ephemeral review (1903-1906) devoted to the same topic: *Beiträge zur Psychologie der Aussage*. In 1906, he founded the Institute for Applied Psychology in Berlin together with the psychologist Otto Lipmann. The institute hosted research on the psychology of testimony and suggestibility³⁸. Stern's 1902 and 1903 results soon gained interest outside Germany. In France, his experiments echoed Alfred Binet's (1857-1911) interest in the risk that investigators and judges unduly direct the answers of witnesses with leading questions; a risk Binet had already pointed out in 1900 in his noted book on suggestibility³⁹. Incidentally, he also expressed interest in creating a «practical science of testimony»⁴⁰. It does not come as a surprise, therefore, that the experiments carried out in Berlin aroused the interest of Binet, who drew the attention of the French-speaking community to this work⁴¹. In 1904, Binet published in *L'Année psychologique*—the journal he founded in 1894⁴²—a paper where he highlighted the need for a «science of the testimony»⁴³. This call was shortly followed by the publication of two other contributions in *L'Année psychologique* offering detailed accounts of the experiments carried out in Switzerland by the psychologist and physician Édouard Claparède (1873-1940), in his own name⁴⁴ or in his laboratory in Geneva by Marie Borst⁴⁵. The same year, the psychologist and physiologist from the University of Lausanne Jean

³⁶ STERN, 1902a.

³⁷ STERN, 1902b.

³⁸ LIPMANN, 1908; STERN, 1910.

³⁹ BINET, 1900.

⁴⁰ *Ibid.*: fn 1, 285.

⁴¹ Binet also took the opportunity to reaffirm his pioneering role in this field of research. His 1900 call for a science of testimony had not, he lamented, received the response it deserved: «One is never a prophet in one's own country, it is understood. This part of my book on suggestibility had no echo in France. [...] It was in Germany that the seed germinated» (BINET, 1904: 130). Over these precedence disputes, see WOLFFRAM, 2018.

⁴² NICOLAS, SEGUI, LEFRAND, 2000.

⁴³ BINET, 1904.

⁴⁴ CLAPARÈDE, 1905a, 1905b.

⁴⁵ BORST, 1904. On the laboratory of psychology in Geneva and Claparède's turn to experimental psychology, see RATCLIFF, RUCHAT, 2006: 40.

Larguier des Bancelles published a voluminous summary⁴⁶ of the work carried out in this field, with particular emphasis on the decisive contributions of Stern⁴⁷. This research paved the way for further investigation on memory and on the capacity of individuals to describe a scene or an image according to their personal characteristics, particularly when the witnesses are children⁴⁸. The reliability of ordinary testimony was strongly questioned, hence the need to analyse in detail the psychological mechanisms hampering perception, recognition and memory.

Forensic psychology received mixed interest from legal scholars. Bentham analysed the «psychological causes of the falsity or truth of the testimony»⁴⁹, but he grounded his views in universal experience much more than in science⁵⁰. The Austrian criminologist Hans Gross (1847-1915) is often praised for his early interest in experimental psychology as a necessary toolkit for investigating judges (his major contributions appeared in the last decade of the 19th century⁵¹). In the United-States, the route opened up by Gross was followed by Wigmore, who even dedicated *The Principles of Judicial Proof* to the memory of Gross, «who did more than any other man in modern times to encourage the application of science to judicial proof»⁵². Gross and Wigmore shared much in common. Both began their careers as legal practitioners (the former an investigating judge and the latter a lawyer) before embracing academic careers, at the University of Gratz (Austria) and Northwestern University (United States), respectively. Both devoted a significant part of their professional activity to the development of criminology⁵³ and investigated the impact of psychology on witnesses and judges. According to Gross, psycho-criminology was a branch of criminology («criminal phenomenology»)⁵⁴, with two main sub-fields: the knowledge of crimes and their investigation and subjective criminal psychology. Gross developed the latter subject in greater length in his 1898 *Kriminal Psychologie*, later translated into English at Wigmore's instigation⁵⁵.

⁴⁶ LARGUIER DES BANCELLES, 1905.

⁴⁷ Larguier des Bancelles had already published in *L'Année psychologique* a laudatory review of Stern's 1902 article on the psychology of testimony (LARGUIER DES BANCELLES, 1902).

⁴⁸ The Belgian psychologist and pedagogue Julien Varendonck published several studies on the reliability of children's testimony: e.g. VARENDONCK, 1914.

⁴⁹ BENTHAM, 1823: 44; BENTHAM, 1827:150.

⁵⁰ BENTHAM, 1827: 110.

⁵¹ GROSS, 1906 (1893); GROSS, 1911 (1898).

⁵² WIGMORE, 1913b. Wigmore also paid explicit homage to Bentham. The proximities between Wigmore and Bentham have been analysed in depth by William Twining (TWINING, 1985).

⁵³ In 1898, Gross founded a scientific journal on criminology: *Archiv für Kriminal-Anthropologie und Kriminalistik*. Wigmore participated in the creation of the *Journal of Criminal Law & Criminology* in 1910 and chaired the *American Institute of Criminal Law and Criminology* (1909-1910).

⁵⁴ GROSS, 1906: xxv.

⁵⁵ GROSS, 1911.

By contrast, in France, forensic psychology received very limited attention in legal scholarship. With rare exceptions⁵⁶, this research was also ignored by legal practitioners. The publication of Gorphe's thesis in 1924 broke with this situation. Gorphe explicitly placed his book in the tradition of these authors:

But if testimony is as old as the world, the science of testimony is as young as our 20th century and it is not yet finished being born. It is true that some jurists have previously had some general ideas drawn from judicial practice, but that is all [...]. It was then that Hans Gross published his famous *Practical Manual of Judicial Investigation*, a chapter of which is devoted to testimonies and contains the most valuable advice for investigating judges. Shortly afterwards, under the impulse of Binet in France, Gross and Stern in Germany, with the help of the recent progress of experimental psychology and clinical psycho-pathology, the first elements of the science of testimony were formed⁵⁷.

In his book, he extensively cited psychologists, whom he believed offered valuable resources for the improvement of the work of judges and investigators. This science of testimony has, he praised,

begun by destroying the «superstition» of the authority of testimony and demonstrating beyond doubt that error in testimony is the rule (in proportions to be determined), not the exception, and that the fidelity of testimony depends not only on the qualities of the witness, but on many factors relating to their mentality, the subject matter of their testimony and the conditions of their testimony⁵⁸.

Gorphe therefore advocated the development of psychological expertise in court, alongside the mental examinations already carried out by specialists in psychiatric disorders. Indeed, like many French psychologists in his time, notably Piéron, he believed that the results of pathological psychology shed light on the psychology of ordinary individuals. From then on, psychological expertise had to «follow the development of experimental psychology»⁵⁹ and required the training of specialized experts. Gorphe described in detail the results of the most important experiments in the field of word association («associative psychodiagnosis», the claim that experimentally provoking the association of ideas could reveal the thoughts hidden by a suspect) carried out in his time, following the pioneering work of the German psychologist W. M. Wundt and further developed by Wertheimer and Klein [1904] in Germany and by Yerkes and Berry in the United States [1909]. The use of the method of word association in trials found in Hugo Münsterberg a fervent promoter⁶⁰, and gave rise to lively exchanges with Wigmore, who undertook to demonstrate

⁵⁶ In 1906, the inspector general of administrative services of the Ministry of the interior, Camille Granier, published a «Critique of oral evidence», which drew heavily on forensic psychology studies: GRANIER, 1906.

⁵⁷ GORPHE, 1924: 12.

⁵⁸ *Ibid.*: 16.

⁵⁹ GORPHE, 1947: 79.

⁶⁰ The psychologist Hugo Münsterberg worked in Germany in Leipzig, where he had been a pupil of Wundt, and in the United States, at Harvard. In 1908, he strongly advocated the use of word association methods in courts [MÜNSTERBERG, 1908].

the limits of this method⁶¹. Gorphe agreed with Wigmore that the promises of the associative method should not be exaggerated, but nevertheless noted the benefits that judges could derive from it: [this method] «is nonetheless superior to empirical means of superficial observation and practical intuition, and deserves to be included among the modes of investigation of psychological expertise»⁶², provided that it is practiced by trained professionals and is supplemented by other evidence. Gorphe also devoted attention to psychological evidence through observation of symptoms (the observation and interpretation of the symptoms of the feelings manifested by the suspect and witness) promoted by the German psychologist C. Leonhardt⁶³ as well as to the then-developing «psychometric tests». Gorphe was particularly interested in the idea that physiological reactions could enable the detection of a lie. He gave a detailed account of the experiments led by the Italian psychologists V. Benussi and C. Musatti on «the respiratory symptoms of sincerity and lying»⁶⁴ and of their later development by Martson and Larson to develop what would become the polygraph. The lie detector «complementing quite well», Gorphe asserted, verbal association methods⁶⁵.

Although he necessarily occupied a marginal place in the field, not being a psychologist by profession, Gorphe earned recognition among psychologists as well. In 1924, Henri Piéron (1881-1964) reviewed *La critique du témoignage* in *L'Année psychologique*. Piéron succeeded Binet at the head of the Laboratory of Psychology at the Sorbonne and of *L'Année psychologique* in 1912 and gained wide fame as a psychologist. In his review of Gorphe, he emphasized the «considerable documentation» gathered by the author and «the conscientious and clear presentation of the main studies relating to the science of testimony»⁶⁶. The professional relationship between Piéron and Gorphe continued throughout their careers. At the end of his life, Gorphe wrote a voluminous contribution to the treatise of applied psychology edited by Piéron⁶⁷. In this chapter, Gorphe gave a comprehensive overview of the contributions of forensic psychology to the evaluation of evidence. This was to be his last published text. Shortly before, Gorphe had urged younger psychologists to consider orienting their career to the service of justice. Deploring the fact that psychologists had not yet been given the role they deserved in the courts and criticizing their «subordinate role» to physicians, Gorphe exhorted psychologists to affirm the

⁶¹ WIGMORE, 1909.

⁶² GORPHE, 1947: 97.

⁶³ *Ibid.*: 99.

⁶⁴ *Ibid.*: 115.

⁶⁵ *Ibid.*: 128

⁶⁶ PIÉRON, 1925.

⁶⁷ GORPHE, 1959. In a short tribute published on the occasion of Gorphe's death in 1959, Piéron recalled his «fundamental work» on the critique of testimony, his book devoted to judicial decisions (1952) and the large chapter devoted to psychology in justice in the volume VII of the *Traité de psychologie appliquée, Livre VII: Les grands domaines d'application*. (PIÉRON, 1959).

usefulness of their knowledge for trials: «Can psychologists accept this disappointing situation? Aren't they better?»⁶⁸.

Gorphe's views on forensic psychology were expressed in consideration of French criminal procedure where the witnesses are examined by the judges, hence his focus on the utility of forensic psychology for judges. As Gorphe put it, judges should not be interested in clinical studies only when witnesses suffer from a mental illness. They must also consider «that part of applied psychology that is now called *forensic psychology*»⁶⁹ which is concerned with the psychology of normal individuals and they should modify their professional techniques accordingly. Crucially, Gorphe pointed to identifications of an individual by a witness (even in good faith) as the source of innumerable miscarriages of justice: «even more than by any other kind of testimony, a serious judgement on the value of a recognition can only be made after a critical examination is made, not at random, but according to a scientific method of psychological analysis that needs to be developed»⁷⁰. In order to reduce the risk of errors, it was therefore necessary that the written minutes of the testimony—the sole document on which judges base their evaluation of the testimony when it is not presented orally in court—describe precisely the conditions under which the testimony was given, whether the witnesses was hesitant, embarrassed, angry, etc. Judges must also pay attention, Gorphe argued, to the conditions under which witnesses were confronted with the person they supposedly recognized in order to avoid any suggestive questions or practices that could interfere with true identification⁷¹: «It is therefore not useless that, following the example of notorious foreign judges, we should take advantage of the new results of forensic psychology to try to establish an interrogation technique [...] and a reporting technique [...] [on] scientific basis»⁷².

Given Gorphe's attention to criminal procedure and given the functions he had occupied in the early years of his career as an investigating judge, it was logically among legal scholars specialising in criminal procedure and in criminology that he particularly sought to find an audience.

3. INTERESTING THE SPECIALISTS OF CRIMINAL LAW AND CRIMINOLOGY IN FORENSIC PSYCHOLOGY

The history of criminology as a scholarly enterprise taking crime and the criminal as its object has been widely undertaken⁷³. The development of criminology

⁶⁸ GORPHE, 1957c.

⁶⁹ GORPHE, 1924: 14.

⁷⁰ GORPHE, 1929a: 293. Because of the limitations of identifications, Gorphe also criticised the full faith given by the law to certain identifications: GORPHE, 1933.

⁷¹ GORPHE, 1929b. This text was also published in English the following year: GORPHE, 1930.

⁷² GORPHE, 1929b: 175.

⁷³ MUCCHIELI, 1994; DEBUYST et al., 1995; KALUSZYNSKI, 2002.

spread throughout the 19th century in Europe, under various forms, and led to the creation of two journals almost simultaneously: in Italy, the *Archivio di psichiatria, antropologia criminale e scienze penali* [Archives of Psychiatry, Criminal Anthropology and Criminal Sciences] (1880-1917) on the initiative of Cesare Lombroso and, in France, the *Archives de l'anthropologie criminelle et des sciences pénales* [Archives of Criminal Anthropology and Criminal Sciences] (1886-1914) founded in Lyon by the physician Alexandre Lacassagne. Neither of these two journals continued after the First World War. After the conflict, the French school of criminology associated with the name of Lacassagne⁷⁴ partly detached from forensic medicine and turned towards investigation techniques. In 1910, Lacassagne's former student and assistant, Edmond Locard, created a laboratory for police science in Lyon. In 1920, he published *L'enquête criminelle et les méthodes scientifiques* [Criminal Investigation and Scientific methods]⁷⁵, in which he called for the transformation of criminal investigation through scientific methods: «scientific methods tend to penetrate the entire criminal trial: whether it is a question of responsibility, observations, clues or the assessment of testimony, everywhere technical considerations are needed as necessary progress»⁷⁶.

Criminology was hence a field where Gorphe's views could find their way and ultimately attract the interest of legal scholarship. Very quickly, Gorphe endeavoured to publish articles defending his ideas in the main criminology and criminal law journals. As his materials and sources of inspiration were far from being only based in France, Gorphe published extensively abroad. In the United States, it is the relationship Gorphe established with Wigmore following the latter's review of *La critique du témoignage* that gave him access to *The American Journal of Police Science*⁷⁷ (the journal of the *American Institute of Criminal Law and Criminology* founded by Wigmore at Northwestern University). Wigmore himself translated into English three pieces by Gorphe and had them published in the *Journal of Criminal Law and Criminology* (also founded by Wigmore)⁷⁸. Gorphe published in European criminal law and criminology journals as well, especially in the main European countries where forensic psychology had developed: Belgium (*Revue de droit pénal et de criminologie*⁷⁹), Switzerland (*Revue de criminologie et de police technique*⁸⁰), Italy (*Giustizia Penale*⁸¹), with the notable exception of Germany.

In France, Gorphe sought to find an audience for his ideas among jurists interested in comparative law (Gorphe was a full member of the French Society for

⁷⁴ KALUSZYNSKI, 2005.

⁷⁵ LOCARD, 1920.

⁷⁶ *Ibid.*: 25.

⁷⁷ GORPHE, 1930.

⁷⁸ GORPHE, 1936a, 1936b, 1936c.

⁷⁹ GORPHE, 1931b, 1934, 1946.

⁸⁰ GORPHE, 1950a; GORPHE, 1951a.

⁸¹ GORPHE, 1938a.

Comparative Legislation and acted as a rapporteur at some of its congresses⁸²). At first, Gorphe directed his contributions to criminologists involved in the French tradition of criminology of Lacassagne and Locard. As early as 1929, Gorphe published his first paper in *Études criminologiques*⁸³, a journal edited by the Institute of Criminology of the University of Paris (1925-1930). He also soon published in the *Revue internationale de criminalistique*⁸⁴ founded by Locard in 1929 and in the *Revue pénitentiaire et de droit pénal*⁸⁵ published by the *Société générale des prisons*⁸⁶. However, after the early 1930s, Gorphe published mainly in journals aimed at legal scholars interested in international criminal law. The contact he established with Henri Donnedieu de Vabres was instrumental in the dissemination of his ideas in this milieu. Gorphe published in the *Revue internationale de droit pénal*⁸⁷ founded in 1924, with the participation of Donnedieu de Vabres. He also had access to the *Revue de science criminelle et de droit pénal comparé*⁸⁸, to whose creation Donnedieu de Vabres had contributed in 1926. Donnedieu de Vabres also agreed to write the prefaces to Gorphe's books published in 1947 and 1952. In these prefaces, Donnedieu de Vabres referred to his experience as a French judge at the Nuremberg Tribunal and underlined the practical interest of Gorphe's research. If the general tone is laudatory («the book of M. Gorphe is in its genre a “Discourse on Method”, a treatise on the art of thinking well, and, by thinking well, of judging well»⁸⁹), a few hints point to some reservations due to the fact that the book is «full of anecdotes» and might lack systematic conclusions.

After World War II, Gorphe's conversations with criminal law specialists in France found a renewed flourishing in the new social defence movement developed in the wake of Marc Ancel⁹⁰. The new social defence movement resonated with much of Gorphe's ideas on the individualization of punishment and the rehabilitation of offenders based on the analysis of their personality. The fact that social defence was attaching great importance to criminology also contributed to the diffusion of Gorphe's thesis, as he had already gained a foothold in criminology circles. Even before the war, the humanist views on criminal law Gorphe had defended were fully in line with those of the new social defence, particularly on the legality of offences⁹¹ or the psychological care of prisoners⁹². From the early 1950s on, Gorphe explicitly referred to these «new criminological conceptions that led to a focus in repression

⁸² See GORPHE, 1952a.

⁸³ GORPHE, 1929a.

⁸⁴ GORPHE, 1929b.

⁸⁵ GORPHE, 1933.

⁸⁶ KALUSZYNSKI, 1997.

⁸⁷ GORPHE, 1932a, 1935b, 1935c, 1937a.

⁸⁸ GORPHE, 1938b, 1937b, 1938c, 1947a, 1957a.

⁸⁹ GORPHE, 1947: 12 (foreword).

⁹⁰ ANCEL, 1954.

⁹¹ GORPHE, 1937a.

⁹² GORPHE, 1938b.

much less on the objective criminal fact than on the personality of the agent»⁹³. The «scientific humanism»⁹⁴ that characterizes the new social defence was thus nicely suited to Gorphe's views.

The social defence promoters strongly advocated for the individualisation of the criminal sentence. Gorphe supported this view, insisting on the need for judges to rely on psychological studies in order to have «a thorough knowledge of the accused and [to decide] on the appropriate treatment to be applied to him»⁹⁵. The repressive individualisation promoted in social defence prompted Gorphe to advocate for the creation of a «real personality file» for each accused person, this being «the only way out of arbitrariness, especially with psychotechnical tests, which allow safe comparisons»⁹⁶. In 1952, Gorphe contributed to the first international criminology course. In his contribution, he insisted on the need that interrogation methods be adapted to the person being interrogated and suggested that instruments such as psychomorphology (linking body shape and character) and graphology were appropriate⁹⁷. In 1953, Gorphe took part in the *Journées de défense sociale* [Social Defence Conference] organised in Poitiers, his city of residence. His contribution appeared in the first volume published by the newly founded *Centre d'études de défense sociale* [Center for the Study of Social Defence] within the *Institut de droit comparé de Paris*, led by Ancel⁹⁸. This paper did not develop any new ideas compared to Gorphe's previous publications but it allowed him to disseminate them to an audience more favourable to his views than any other before. However, this new opportunity arrived late in Gorphe's life. Most of the texts he published in the 1950s, up to his death in 1959, were essentially reprints and syntheses of his contributions of the 1920s and 1930s. The only innovations he put forward at that time were practical recommendations regarding the activity of judges, quite distant from the concerns of the social defence.

4. TRAINING SOCIETY TO GIVE TESTIMONY, TRAINING JUDGES TO EVALUATE WITNESSES

Beginning with the defence of his thesis in 1924, Gorphe tried to reconcile a reflection that the analysis of evidence be grounded on the best scientific knowledge available with the design of practical recommendations directed to judges in order to reinforce their professional skills. In his terms, referring to the «technical method» he

⁹³ GORPHE, 1952b: 1.

⁹⁴ SIZAIRE, 2017. For a subjective account of social defence within the field of criminology, see RADZINOWICZ, 1999.

⁹⁵ GORPHE, 1950a: 84.

⁹⁶ GORPHE, 1951b: 168.

⁹⁷ GORPHE, 1952c.

⁹⁸ GORPHE, 1954.

was offering to judges, «this method, like justice, must strike a balance between the modern scientific spirit and the practical concerns of the courts»⁹⁹.

The first practical lesson Gorphe drew from forensic psychology was that testimony is fragile but still «educable». The psychologists Borst¹⁰⁰ and Claparède¹⁰¹ had experimentally demonstrated that individuals, both adults and children, trained in testifying, significantly improve the accuracy and precision of their statements. Gorphe echoed this idea in March 1929 at the International Congress of Applied Psychology held in Paris. His contribution was published in the journal *La psychologie et la vie* under the title «L'éducabilité du témoignage»¹⁰². In this paper, he argued that «learning to provide testimony includes a whole series of tasks»¹⁰³ (to observe, to fix one's attention, to remember, to reproduce faithfully). Therefore, proper training for these tasks must be provided, not only to children in schools, but also to «every person». Hence Gorphe strongly advocated that everyone should learn «the elementary notions of the art of acting as a witness and, if need be, to practise it without waiting to be called to justice and to risk the worst errors to the detriment of others»¹⁰⁴.

This proposal, which Gorphe pointed out as constituting a true «psycho-pedagogical program», easily found its place in *La psychologie et la vie: Revue de psychologie appliquée*. Indeed, this journal was edited from 1927 onwards by the commercial company Pelman Institute, which was committed to spreading in France the Pelman method of mental training and personal development¹⁰⁵. The Pelman method had been developed in the United Kingdom in the late 19th century by Christopher Louis Pelman and popularised by William Joseph Ennever as a method of training physical and mental capacities dedicated to the reinforcement of self-confidence and success. In the 1920s and 1930s, the method had been quite popular in France. The distance course «Twelve lessons of Pelmanism» received a widespread audience and the publication of *La psychologie et la vie* (1927-1939), under the supervision of the philosopher and psychologist Paul Masson-Oursel, offered a platform for the diffusion of these ideas. Gorphe's view on the educability of the testimony echoed the confidence of Pelmanism in training methods and self-improvement. Just as the Pelman method—which Gorphe did not mention in his paper—offered to develop personal capacities through mental training, testifying in courts should be based on an ordered method: this program, Gorphe argued, «is realizable only with the contact of the facts and with the help of a practical method»¹⁰⁶. This eventually demands a moral education of the citizen, which must be based first and foremost on

⁹⁹ GORPHE, 1947: 472.

¹⁰⁰ BORST, 1904.

¹⁰¹ CLAPARÈDE, 1905a, 1905b.

¹⁰² GORPHE, 1929c.

¹⁰³ *Ibid.*: 55.

¹⁰⁴ *Ibid.*: 56.

¹⁰⁵ LE SONN, 2017.

¹⁰⁶ GORPHE, 1929c: 55.

self-knowledge: «It is thus important to know oneself: critical control of oneself and moral conscience»¹⁰⁷.

Beside his interest in the education of the general public to the art of testifying, Gorphe devoted much attention to the education of judges. As early as the 1930s, Gorphe insisted on the need to train judges in «ancillary forensic sciences» (criminology, penitentiary science, forensic psychology, forensic psychiatry, evidence technology) and suggested that these topics should be included in the curriculum of the professional examination for admission to the judiciary¹⁰⁸. But, in his view, the training of judges also required special moral qualities based on a good «knowledge of men» and «a psychological penetration that is not provided by the abstract science of law or the practice of civil affairs»¹⁰⁹. This contention led Gorphe to strongly oppose the intervention of popular juries¹¹⁰, these «improvised judges»¹¹¹ administering an «arbitrary and fanciful»¹¹² justice, who lacked both experience and moral qualities. Nevertheless, Gorphe's views on the qualities of judges can be considered as pertaining much more to professional identity building than a really thorough work based on experimental psychology. In 1951, when the occasion was given to him to express his views in the *Académie des sciences morales et politiques*, Gorphe referred only to very broad notions of the moral qualities that the judges should possess more than anyone else: «qualities of intelligence and method, and qualities of will and firmness»¹¹³.

If Gorphe advocated the need for a method in the education of witnesses and judges, he did not himself trace its contours. Here again, he turned to psychologists. Gorphe's interest in the training of personal capacities, in the general public and in judges, turned him into a firm promoter of the psychotechnical methods that had been recently designed in the field of vocational orientation. His above-mentioned acquaintance with Piéron, as well as the interest he had developed in Claparède's work in the 1920s, had enhanced Gorphe's familiarity with these methods. Both authors had, indeed, an interest in education and promoted new methods taking into account the individuality and capacities of each child¹¹⁴. Both promoted, as well, the

¹⁰⁷ GORPHE, 1929c: 55.

¹⁰⁸ GORPHE, 1932b: 241; GORPHE, 1957c: 167.

¹⁰⁹ GORPHE, 1932b: 243.

¹¹⁰ GORPHE, 1934; GORPHE, 1935b; GORPHE, 1952d: 111. Gorphe also objected to witnesses being questioned by people who had not been specially trained to criticize testimony, such as journalists or insurance agents (GORPHE, 1924: 393).

¹¹¹ GORPHE, 1951b: 165.

¹¹² *Ibid.*: 169.

¹¹³ *Ibid.*: 166.

¹¹⁴ Claparède strongly criticised the uniformity of pedagogical methods and their lack of consideration of the tastes, creativity and personality of the pupils (RUCHAT, 2015). He supported the pedagogical experiences that took place under the label of the *éducation nouvelle* and promoted their development, notably by training the educators in the school for learning sciences (*Ecole libre des sciences de l'éducation. Institut Jean-Jacques Rousseau*) he founded in 1912 in Geneva (HOFSTETTER, SCHNEUWLY,

contributions of psychological research to professional orientation¹¹⁵. Gorphe collaborated with the *Société psychotechnique de Paris* [Psychotechnical Society of Paris] which was directed by Guy de Beaumont, a great promoter of psychotechnical tests in recruitment. In 1929, Gorphe was also a member of the *Institut général psychologique*¹¹⁶, a rather eclectic group where mediums interested in telepathy would meet with academic psychologists striving to give a scientific base to their discipline like Piéron¹¹⁷. Gorphe took up these ideas and suggested that psychotechnical tests should be used not only as an important tool for professional orientation but also as a way to assess personalities in courts (whether the accused persons or the witnesses). It would be advisable, he suggested, to «constitute in schools an anthropological record for each pupil or student: this would serve, not only for justice eventually, but also and above all for the subject himself for his professional and general orientation in life»¹¹⁸. The same holds for judges, Gorphe claimed. Psychotechnical examinations developed by specialists of professional orientation can help verify that aspiring judges bear the moral qualities needed to perform a balanced and effective evaluation of the evidence. Indeed, if Gorphe had underlined the limitations of these methods at an early stage of their development¹¹⁹, his caution was all gone in the 1950s: «Psychotechnical science is now in a position to discern aptitudes of all kinds»¹²⁰; it is «one of the most useful applications of experimental psychology, used for professional orientation, diagnosis of aptitudes, selection of candidates in various branches of activity»¹²¹. Psychotechnical examinations should thus be included in professional examinations to determine the aptitudes of judges: this would avoid «the deplorable spectacle of a magistrate failing to judge properly»¹²².

5. CONCLUSION

Despite the originality of Gorphe's work and although he was closely connected to various professional fields, including law, psychology and criminology, his impact on legal scholarship in France has been limited. Apart from Donnedieu de Vabres—who wrote the preface to his last two books—Gorphe received very limited attention within French universities. After Gorphe's death in 1959, shortly after he retired, the path he had opened up was not followed by others. Forensic psychology

2006). Piéron embraced these ideas as well. He contributed to the institutionalisation of the *éducation nouvelle* by sitting on the board of the *Groupe français d'éducation nouvelle* created in 1929, thereby enhancing the scientific credentials of this institution (SAVOYE, 2007: 240).

¹¹⁵ HUTEAU, 2014; HUTEAU, 2018; MARTIN, 2014: 112.

¹¹⁶ MURCHISON, 1929: 379.

¹¹⁷ PLAS, 2012: 98 ; MARMIN, 2001: 157.

¹¹⁸ GORPHE, 1950a: 88.

¹¹⁹ GORPHE, 1924: 114.

¹²⁰ GORPHE, 1950a: 90.

¹²¹ GORPHE, 1951b: 171.

¹²² GORPHE, 1950a: 90.

and criminology kept developing after World War II, but this happened in isolation from law schools and without any connection to academic lawyers. The post-war period did not see the emergence of figures that would be willing to work—like Gross, Wigmore or Gorphe did—at the interface of the theory of evidence, criminology and psychology. The opportunity to lay the foundations of a legal scholarship interested in the evaluation of evidence and evidential reasoning was missed.

This lack of academic interest in Gorphe's work has several explanations. The very practical nature of his work is one of them. Indeed, Gorphe was concerned with the actual practice in the courts and assumed an orientation towards the professional skills of judges. Just like Gross¹²³ and Wigmore¹²⁴, Gorphe did not hesitate to go into very mundane details (e.g. the recommendation that judges draw the layout of the premises or that witnesses draw the person they have seen before proceeding with confirmations of identity)¹²⁵. Moreover, the psychological theories and experiments on which Gorphe relied may now seem dated and even obsolete. The faith he placed in psychotechnical tests, psychomorphology or methods of mental development inspired by Pelmanism have been partly discredited over the years. Moreover, some of his recommendations may seem unrealistic, such as filing an anthropological record for each student in the schools aimed at providing information about his or her personality. Finally, Gorphe's appeal to the personal qualities of the judges gives a moralizing stance to his last pieces that strongly contrast with his claim to base the evaluation of evidence on science.

Gorphe's failure to secure a lasting influence in the universities is also due to institutional reasons. Gorphe had always worked as a judge, moreover in provincial jurisdictions. Apart from a few occasional lectures, he never held a university position. In France, law professors have maintained a distant relationship with judges, by whom they have sought information on the practice of courts but rarely theoretical inspiration. Similarly, the efforts Gorphe made throughout his career to get criminologists interested into the contributions of forensic psychology suffered from the development of criminology outside law faculties in France¹²⁶ and from the weakening of the «social defence» doctrine in 1980s France¹²⁷. Furthermore, Gorphe's interest in the psychological foundation of the evaluation of evidence was far from resonant with legal scholarship. Indeed, French academics inherited from the Napoleonic reconstruction of the universities in the early 19th century a legal culture centred on dogmatic analysis¹²⁸. Now, as Gorphe himself pointed out at the outset of *La critique du témoignage*, French law does not impose on judges what form evidential reasoning should take. The free proof system that mainly prevails since the

¹²³ GROSS, 1906: 143.

¹²⁴ WIGMORE, 1913a: 755.

¹²⁵ GORPHE, 1929b: 167.

¹²⁶ COLSON, 2013.

¹²⁷ SIZAIRE, 2017.

¹²⁸ JAMIN, JESTAZ, 2004; AUDREN, HALPERIN, 2013.

French Revolution—as the intimate conviction in criminal trials epitomize—, leaves judges free to assess the evidence, without the law imposing the means by which their conviction must be formed or the degree to which they must be convinced to consider a fact proven¹²⁹. In such a context, Gorphe ultimately invited jurists to turn their attention to the evaluation of evidence, a domain where French law remains silent. Consequently, such an invitation had little chances of finding any echo among academics essentially interested in dogmatic analysis of the law. While Anglo-American scholarship gave far more attention to Wigmore than to Gorphe, it is only in Spanish-speaking countries where Gorphe received continued interest from legal doctrine. Gorphe's books have all been translated into Spanish¹³⁰, and still receive attention today¹³¹. This situation does not reflect a dismissal of dogmatic analysis in Spanish-speaking scholarship. As in Spain and some Latin American countries the rules of procedure explicitly require judges to comply with *sana crítica*, the evaluation of evidence became a possible subject of analysis for jurists interested in the dogmatic approach. Gorphe was fully aware that the rule of *sana crítica* offered a possible reconciliation of his views with dogmatic doctrinal analysis. When he depicted this provision of the Spanish Code of Civil Procedure of 1881, he noticed with a hint of envy that this «wise prescription [...] is in line with the ideas set out [in his 1924 book]»¹³².

6. BIBLIOGRAPHY

- ABEL LLUCH, X., 2015: *Las reglas de la sana crítica*, Madrid: La Ley.
- ANCEL, M., 1954: *La défense sociale nouvelle*, Paris: Cujas.
- AUDREN, F., HALPERIN, J.-L., 2013: *La culture juridique française. Entre mythes et réalités XIXe-XXe siècle*, Paris: LGDJ.
- BENTHAM, J., 1823: *Traité des preuves judiciaires*, Paris: Bossange Frères.
- 1827: *Rationale of Judicial Evidence*, vol. 1, London: Hunt and Clarke.
- BERNHEIM E., 1903: *Lehrbuch des historischen Methode und der Geschichtsphilosophie*, Leipzig: Duncker & Humblot.
- BINET, A., 1900: *La suggestibilité*, Paris: Schleicher Frères.
- 1904: «La science du témoignage», in *L'Année psychologique*, 11:128-136.
- BORST, M., 1904: «Recherches expérimentales sur l'éducabilité et la fidélité du témoignage», in *Archives de psychologie*, 3:233-314.
- CLAPARÈDE, E., 1905a: «La psychologie judiciaire», in *L'Année psychologique*, 12: 275-302.
- 1905b: «Expériences collectives sur le témoignage», in *Archives de psychologie*, 5: 344-387.
- COLSON, R., 2013: «Criminology à la française», in *Brit J Criminol*, 53(4): 552-567.
- COMTE, A., 1830: *Cours de philosophie positive*, Paris: Rouen Frères.
- DE SMEDT, Ch., 1883 : *Principes de la critique historique*, Liège: Société bibliographique belge.
- DEBUYST, C., DIGNEFFE F., LABADIE J.-M., PIREN, A. P., 1995: *Histoire des savoirs sur le crime et la peine (vol. 1: Des savoirs diffus à la notion de criminel né)*, Bruxelles: De Boeck and Larcier.

¹²⁹ TARUFFO, 2003.

¹³⁰ GORPHE, 1949; GORPHE, 1950b; GORPHE, 1955.

¹³¹ e.g. ABEL LLUCH, 2015.

¹³² GORPHE, 1924: 32.

- FERRI, E., 1905 (1881): *Sociologie criminelle* (trad. Terrier), Paris: F. Alcan.
- GORPHE, F., 1924: *La critique du témoignage*, Paris: Dalloz.
- 1928: *Le principe de la bonne foi*, Paris: Dalloz.
- 1929a: «Comment apprécier les reconnaissances d'identité», in *Études criminologiques*, 1: 287-293.
- 1929b: «Comment procéder aux confrontations d'identité», in *Revue internationale de criminalistique*, 4-5: 165-175.
- 1929c: «L'éducabilité du Témoignage», in *La Psychologie et la Vie: Revue de psychologie appliquée*, 53-55.
- 1930: «Showing Prisoners to Witnesses For Identification», in *Am. J. Police Sci.*, 1: 79-88.
- 1931a: «Formation de la nouvelle jurisprudence sur la responsabilité des accidents», in *Revue critique de législation et de jurisprudence*, 1-2: 177-196.
- 1931b: «La conviction des juges», in *Revue de droit pénal et de criminologie*, 557-568.
- 1931c: «La fixation des faits de la cause par les juges», in *Revue internationale de criminalistique*, 2: 113-129.
- 1931d: «Les opérations du jugement judiciaire», in *Revue Philosophique de la France et de l'Étranger*, 111: 187-203.
- 1932a: «Convient-il de considérer l'accusé comme un témoin dans sa propre cause?», in *Revue internationale de droit pénal*, 1-2: 190-198.
- 1932b: «De quelles manières pourrait-on obtenir une meilleure spécialisation du juge pénal?», in *Revue internationale de droit pénal*, 1-2: 241.
- 1933: «La force probante des reconnaissances d'identité énoncées dans les procès-verbaux faisant foi jusqu'à inscription de faux», in *Revue pénitentiaire et de droit pénal. Bulletin de la Société générale des prisons*, 57: 171-180.
- 1934: «Le procès autrichien Halsmann et les leçons à en tirer», in *Revue de Droit Pénal et de Criminologie*, 1: 650-667.
- 1935a: «La responsabilité du fait des automobiles en droit comparé: jurisprudence française, lois étrangères et projet de loi», in *Revue critique de législation et de jurisprudence*, 150-195.
- 1935b: «Les Réformes du jury à l'étranger», in *Revue internationale de droit pénal*, 370-424.
- 1935c: «Les Réformes du jury aux États-Unis», in *Revue internationale de droit pénal*, 425-431.
- 1936a: «Reform of the Jury System in European Countries: England», in *Journal of Criminal Law and Criminology*, 27(1): 17-30.
- 1936b: «Reforms of the Jury-System in Europe: France and Other Continental Countries», in *Journal of Criminal Law and Criminology*, 27(2): 155-168.
- 1936c: «Reform of the Jury-System in Europe: Results and Conclusions», in *Journal of Criminal Law and Criminology*, 27(4): 473-479.
- 1937a: «Est-il désirable que le juge puisse retenir et punir un fait qui ne rentre pas expressément sous les termes d'une disposition légale?», in *Revue internationale de droit pénal*, 1-4: 228-240.
- 1937b: «La valeur probante des indices», in *Revue de science criminelle et de droit pénal comparé*, 2: 418-437.
- 1938a: «Comment peut-on appliquer la psychologie en justice?», in *Giustizia Penale*, 465-473.
- 1938b: «Une visite à des pénitenciers agricoles suisses», in *Revue de science criminelle et de droit pénal comparé*, 34-53.
- 1938c: «Variétés et difficultés dans l'appréciation des indices», in *Revue de science criminelle et de droit pénal comparé*, 2: 215-241.
- 1946: «Les rapports entre les preuves», in *Revue de droit pénal et de criminologie*, 465-479.
- 1947a: «La méthode générale d'examen critique des preuves», in *Revue de science criminelle et de droit pénal comparé*, 1: 69-79.
- 1947b: *L'appréciation des preuves en justice: essai d'une méthode technique*, pref. H. Donnedieu de Vabres, Paris: Sirey.
- 1949: *La Critica del Testimonio* (trad. Mariano Ruíz-Funes), Madrid: Instituto Editorial Reus.
- 1950a: «Peut-on réaliser une justice scientifique?», in *Revue de criminologie et de police technique*, 4(2): 83-91.

- 1950b: *De la apreciación de las pruebas* (trad. Luis Alcalá Zamora y Castillo), Buenos Aires: Ediciones Jurídicas Europa-América / Bosch y Cía.
- 1951a: «Le sens de l'aveu criminel dans une critique scientifique des preuves», in *Revue de criminologie et de police technique*, V(1): 9-13.
- 1951b: «Quelques aperçus psychologiques sur l'élaboration de la justice», in *Revue des travaux de l'Académie des sciences morales & politiques et comptes rendus de ses séances*, 164-171.
- 1952a: «L'aveu dans la procédure pénale», in *Revue Internationale de Droit comparé*, 4: 776-780.
- 1952b: «L'explication psychologique des crimes par l'examen des inculpés», in *Bulletin de la Société de philosophie de Bordeaux*, 6(34): 1-17.
- 1952c: «Une méthode technique de l'interrogatoire judiciaire», in G. HEUYER et J. PINATEL (coord.), *L'examen médico-psychologique et social des délinquants; Premier cours international de criminologie*, Paris: Société internationale de criminologie, 212-222.
- 1952d: *Les Décisions de justice: Étude psychologique et judiciaire*, pref. H. Donnedieu de Vabres, Paris: Puf (ed. Recueil Sirey, 1952).
- 1954: «Comment peut se faire l'individualisation judiciaire?», in M. ANCEL et J.-B. HERTZOG (dir.), *L'Individualisation des mesures prises à l'égard du délinquant*, Paris: Cujas, 363-392.
- 1955: *De la apreciación de las pruebas*, (trad. Santiago Sentís Melendo), Buenos Aires: Ed. jurídicas Europa.
- 1957a: «L'influence du cinéma sur la délinquance et les questions qu'elle pose», in *Revue de Science criminelle et de Droit pénal comparé*, 599-605.
- 1957b: «Le rôle des psychologues dans les examens en justice», in *Bulletin de psychologie*, XI(12): 149.
- 1957c: «Les applications philosophiques en justice», in *Archives de philosophie du droit*, 3:167.
- 1959: «La psychologie appliquée en justice», in H. PIÉRON et al. (dir.), *Traité de psychologie appliquée, Livre VII: Les grands domaines d'application (Domaine social, scolaire, industriel, judiciaire, militaire)*, Paris: PUF: 1395-1790.
- GRANIER, C., 1906: «Aveu et témoignage. Critique de la preuve orale», in *Journal du Ministère public*, Marchal et Billard.
- GROSS, H., 1906 (1893): *Criminal Investigation. A Practical Handbook for Magistrates, Police Officers, and Lawyers* (trans. J. Adam and J. C. Adam), Madras: Krishnamachari.
- 1911 (1898): *Criminal Psychology. A Manual for Judges, Practitioners, and Students* (trans: Kallen H M), Boston: Little, Brown & Co.
- HALPERIN, J.-L., 2007: «Donnedieu de Vabres, Félix-Auguste-Henri», in P. ARABEYRE, J.-L. HALPERIN, J. KRYNEN, *Dictionnaire historique des juristes français (XIIe-XXe siècle)*, Paris: PUF.
- HOFSTETTER, R., SCHNEUWLY, B., 2006: «Le laboratoire: matrice des sciences de l'éducation – l'exemple de Genève», in RATCLIFF M. J., RUCHAT, M., 2006: *Les laboratoires de l'esprit. Une histoire de la psychologie à Genève, 1892-1965*, Genève: Musée d'histoire des sciences.
- HUTEAU, M., 2014: «Henri Piéron, la psychologie de l'orientation professionnelle», in *Bulletin de psychologie*, 67(5):363-384.
- 2018: «Édouard Claparède (1873-1940) et l'orientation professionnelle», in *Bulletin de psychologie*, 71(2): 623-631.
- JESTAZ, Ph., JAMIN, C., 2004: *La doctrine*, Paris: Dalloz.
- KALUSZYNSKI, M., 1997: «Réformer la société. Les hommes de la société générale des prisons, 1877-1900», in *Genèses*, 28: 76-94.
- 2002: *La République à l'épreuve du crime. La construction du crime comme objet politique (1880-1920)*, Paris: LGDJ.
- 2005: *Histoire de la criminologie*, Criminocorpus. Available via OpenEdition. <https://journals.openedition.org>.
- LANGLOIS, Ch.-V., SEIGNOBOS, Ch., 1898: *Introduction aux études historiques*, Paris: Hachette.
- LARGUIER DES BANCELS, J., 1902: «Review: Stern, Zur Psychologie der Aussage. Experimentelle Untersuchungen über Erinnerungstreuung», in *L'Année psychologique*, 9: 331-338.
- 1905: «La psychologie judiciaire», in *L'Année psychologique*, 12:157-232.

- LE SONN, L., 2017: «L'institut Pelman de gymnastique mentale: une officine de consultations psychologiques durant l'Entre-deux-guerre», in *Corps*, 15: 67-77.
- LIPMANN, O., 1908: *Die Wirkung von Suggestivfragen*, Leipzig: Barth.
- LOCARD, E., 1920: *L'enquête criminelle et les méthodes scientifiques*, Paris: Flammarion.
- MARMIN, N., 2001: «Métapsychique et psychologie en France (1880-1940)», in *Revue d'Histoire des Sciences Humaines*, 4: 145-171.
- MARTIN, J., 2014: «De l'orientation professionnelle à l'orientation scolaire: l'Association générale des orienteurs de France et la construction de la profession de conseiller d'orientation (1931-1956)», in *Histoire de l'éducation*, 142: 109-128.
- MUCCHIELLI, L. (ed.), 1994: *Histoire de la criminologie française*, Paris: L'Harmattan.
- MÜNSTERBERG, H., 1908: *On the Witness Stand. Essays on Psychology and Crime*, New York: The McClure Company.
- MURCHISON, C. A., 1929: *The Psychological register*, Vol. 2, Worcester: Clark University Press.
- NICOLAS, S., SEGUI, J., LEFRAND, L., 2000: «Les premières revues de psychologie: la place de L'Année Psychologique», in *L'Année psychologique*, 100(1): 71-110.
- PIÉRON, H., 1925: «Review: F. Gorphe, La critique du témoignage», in *L'Année psychologique*, 579.
- 1959: «Nécrologie», in *L'Année psychologique*, 59(2): 659-660.
- PLAS, R., 2012: «Psychology and psychical research in France around the end of the 19th century», in *History of the Human Sciences*, 25(2): 91-107.
- RADZINOWICZ, L., 1999: *Adventures in Criminology*, London: Routledge.
- RATCLIFF M. J., RUCHAT, M., 2006: *Les laboratoires de l'esprit. Une histoire de la psychologie à Genève, 1892-1965*, Genève: Musée d'histoire des sciences.
- RUCHAT, M., 2015: *Edouard Claparède. A quoi sert l'éducation?*, Lausanne: Antipodes.
- SAVOYE, A., 2007: «L'Education nouvelle en France. De son irrésistible ascension à son impossible pérennisation», in OHAYON, A., OTTAVI, D., SAVOYE, A. (eds), *L'Education nouvelle, histoire, présence et devenir*, Bern: Peter Lang., 235-269.
- SIZAIRE, V., 2017: «Que reste-t-il de la défense sociale nouvelle?», in *Revue de science criminelle et de droit pénal comparé*, 2: 261-272.
- STERN, W., 1902a: «Zur Psychologie der Aussage», in *Zeitschrift für die gesamte Strafrechtswissenschaft*, 22(2-3): 362-418.
- 1902b: *Zur Psychologie der Aussage. Experimentelle Untersuchungen über Erinnerungstreu*, Berlin: Guttentag.
- 1910: «Abstracts of Lectures on the Psychology of Testimony and on the Study of Individuality», in *The American Journal of Psychology*, 21(2): 270-282.
- TARUFFO, M., 2003: «Rethinking the Standards of Proof», in *The American Journal of Comparative Law*, 51(3): 659-677.
- TWINING, W., 1985: *Theories of Evidence: Bentham and Wigmore*, London: Weidenfeld & Nicolson.
- VARENDONCK, J., 1914: *La psychologie du témoignage*, Gand: Maison d'éditions & d'impression.
- WERTHEIMER, M., KLEIN, J., 1904: «Psychologische Tatbestandsdiagnostik», in *Archiv für Kriminal-Anthropologie und Kriminalistik*, XV: 73.
- WIGMORE, J. H., 1909: «Professor Muensterberg and the Psychology of Testimony», in *Illinois Law Review*, 3: 399-345.
- 1913a: «The Problem of Proof», in *Illinois Law Review*, 8(2): 77-103.
- 1913b: *The Principles of Judicial Proof as Given by Logic, Psychology, and General Experience and Illustrated in Judicial Trials*, Boston: Little, Brown & Co.
- 1928: «Review of F. Gorphe, La Critique du Témoignage», in *Journal of the American Institute of Criminal Law and Criminology*, 18(4): 605-606.
- 1937: *The Science of Judicial Proof*, Boston: Little, Brown & Co.
- WOLFFRAM, H., 2018: *Forensic Psychology in Germany: Witnessing Crime, 1880-1939*, London: Palgrave Macmillan.
- YERKES, R. M., BERRY, CH., 1909: «The association-reaction method of mental diagnosis», in *American Journal of Psychology*, XX: 22-37.

PREDICTIBILIDAD CONDUCTUAL Y PROCESO PENAL: ALGUNOS APUNTES SOBRE EL FUNDAMENTO EPISTÉMICO DE LOS HECHOS FUTUROS EN LAS MEDIDAS CAUTELARES

Jeffrey José Mora Sánchez

Juez Penal, Poder Judicial de Costa Rica

jmorasa@poder-judicial.go.cr

jjmsabogado@gmail.com

RESUMEN: La determinación de hechos futuros en un proceso penal implica un juicio predictivo conductual que, como cualquier otra decisión, debe basarse en elementos epistémicos sólidos y generarse a partir de un razonamiento inferencial válido. Desde la perspectiva señalada, la inferencia probatoria que el juzgador realice, para avalar la disposición de una medida cautelar, deberá incorporar una garantía, constituida por máximas de experiencia, presunciones o definiciones. El presente texto se ocupa de evaluar, partiendo del diseño procesal penal costarricense, cómo debe presentarse ese razonamiento inferencial, así como los aportes que la psicología cognitiva, los instrumentos actuariales y la inteligencia artificial pueden brindar, con el objetivo de construir resoluciones judiciales racionales e intersubjetivamente controlables, en un rubro tan relevante como el de las medidas cautelares y la prisión preventiva.

PALABRAS CLAVE: Hechos futuros; inferencia probatoria; psicología cognitiva; instrumento actuarial; garantía.

BEHAVIORAL PREDICTABILITY AND CRIMINAL PROCEEDINGS: SOME NOTES ON THE EPISTEMIC FOUNDATION OF FUTURE EVENTS IN PRECAUTIONARY MEASURES

ABSTRACT: The determination of future events in a criminal process implies a predictive behavioral judgment that, like any other decision, must be based on solid epistemic elements and generated from valid inferential reasoning. From the indicated perspective, the evidentiary inference that the judge makes, to endorse the provision of a precautionary measure, must incorporate a guarantee, consisting of maxims of experience, presumptions or definitions. The present text deals with evaluating, based on the Costa Rican criminal procedural design, how this inferential reasoning should be presented, as well as the contributions that cognitive psychology, actuarial instruments and artificial intelligence can provide, with the objective of building rational and intersubjectively controllable judicial resolutions, in an area as relevant as that of precautionary measures and preventive detention.

KEYWORDS: future facts; evidentiary inference; cognitive psychology; actuarial instrument; warrant.

SUMARIO: 1. INTRODUCCIÓN.—2. JUSTIFICACIÓN DE LA PREMISA FÁCTICA DE LAS MEDIDAS CAUTELARES: 2.1. Factores procesales positivos: 2.1.1. Incompletitud del elenco probatorio. 2.1.2. Devaluación de la garantía del derecho a la prueba. 2.2. Factores de resiliencia interna del decisor. 2.3. Factores estructurales de la pretensión.—3. LA INFERENCIA PROBATORIA EN LOS HECHOS FUTUROS: 3.1. ¿Jueces gurús, *bouche de la loi* o decisores racionales?: 3.1.1. Presunciones *iusris* o definiciones. 3.1.2. Máximas de experiencia.—4. EL *BACKING* DEL *WARRANT* EN LOS HECHOS FUTUROS: 4.1. La psicología cognitiva y la estadística. 4.2. La inteligencia artificial. 4.3. ¿Cuál es el rendimiento de la estadística?—5. CONCLUSIONES.—6. BIBLIOGRAFÍA.

RECOMMENDED CITATION: MORA SÁNCHEZ, J. J., 2020: «Predictibilidad conductual y proceso penal: algunos apuntes sobre el fundamento epistémico de los hechos futuros en las medidas cautelares», in *Quaestio facti*, 2: 53-85. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22463

*«Man can believe the impossible, but man can never believe the improbable»
Intentions, Oscar WILDE.*

*«A medida que se moderen las penas, que se supriman la sordidez y el hambre de las cárceles, que la compasión y la humanidad traspongan las puertas con cerrojos e inspiren a los inexorables y empedernidos ministros de justicia, las leyes podrán contentarse con indicios cada vez más débiles para capturar»
De los delitos y de las penas, Cesare BECCARIA.*

*«[E]l derecho penal debe esforzarse por aniquilar no solo a quien efectivamente haya perturbado la paz en casos concretos (...) toda futura perturbación de la paz y todas las predisposiciones que se correspondan con ella deberán combatirse con la totalidad de las armas idóneas para la defensa»
Derecho penal de voluntad, Roland FREISLER.*

1. INTRODUCCIÓN

En el año 2002 el director estadounidense Steven Spielberg llevó a la pantalla grande la distopía futurista *Minority Report*, basada en el cuento de Philip Kindred Dick. La película narra el ocaso de una unidad policial (*PreCrime*) que se vale de tres hermanos psíquicos (*precogs*) cuya habilidad innata les permite predecir homicidios; sin embargo, la promisoría institución represiva se desploma cuando se evidencia la falibilidad de las precogniciones. El escenario presentado en el filme representa uno de los implícitos anhelos de cualquier modelo de justicia penal, al punto que, aunque no se perciba con especial facilidad, algunas instituciones jurídicas funcionan desde larga data bajo la lógica de la predictibilidad conductual.

Los esfuerzos que encumbraron desde el positivismo criminológico Enrico Ferri, Cesare Lombroso y Rafael Garófalo, arguyendo la predisposición anatómica de ciertas personas para incurrir en conductas ilícitas, ha sido afablemente receptada por algunos Estados totalitarios como andamiaje de políticas públicas de criminalización de enemigos¹, pero también perviven en la forma en que opera el aparato represivo en algunos países de semblanza liberal y democrática, pero con altísima selectividad criminal, dirigida a ciertas *clases* de individuos². En la actualidad, el desarrollo de las neurociencias, especialmente los estudios relacionados con la vinculación del lóbulo frontal del cerebro y la toma de decisiones (SILVA, MERCURIO, y LÓPEZ 2008: 76–112) plantea enormes tensiones desde el punto de vista ético, al resultar plausible que en algún momento el sueño de Lombroso y sus acólitos tenga aún mayor acogida, esta vez refrendado por la autoridad de la ciencia.

Pero al margen de idealismos y posturas extremas superadas, al menos formalmente, el proceso penal ha dado buen recaudo a ciertos actos que requieren un ejercicio prospectivo y no retrospectivo, dentro de los cuales, por su nivel de impacto en la esfera de derechos fundamentales de las personas, la determinación de medidas cautelares de carácter personal³, sin duda alguna, es la más prominente, en especial cuando de prisión preventiva se trata⁴. La decisión judicial que se pronuncia sobre la

¹ Con una mera intención ejemplificativa puede verse el panfleto político elaborado por Roland FREISLER, el llamado «soldado político de Hitler», de donde se ha extraído la cita colocada en el epígrafe de esta monografía (FREISLER 2017: 117–118).

² Quizá el caso estadounidense sea el más paradigmático, por su autoconcepción como la *meca* de la libertad y la democracia. Refiriéndose a la anormal proporción de gente afrodescendiente en prisión en este país, puede verse: CHRISTIE 2004: 171–175.

³ En adelante, cuando se haga referencia a medidas cautelares deberá entenderse que se trata de cualesquiera medidas de carácter personal y no real, incluyendo por supuesto la prisión preventiva.

⁴ Aunque en el presente trabajo se ha preferido el abordaje general de las medidas cautelares, es evidente que los mayores problemas prácticos se dan con relación a la prisión preventiva. No obstante, debe hacerse hincapié en que los modelos argumentativos y la construcción de la inferencia probatoria será, en esencia, idéntica para uno u otro caso, pues lo que marca la diferencia en la aplicación de las diversas tipologías de medidas precautorias es la ponderación en cuanto a la entidad del riesgo y el estándar probatorio determinado para cada una de estas.

determinación de medidas cautelares contra un individuo requiere poco menos que la clarividencia del juzgador, dado que este necesariamente debe pronunciarse sobre dos aspectos medulares⁵: la probabilidad de que contra el investigado se emita una sentencia condenatoria⁶ y la probabilidad de que la libertad del imputado entrañe la afectación del proceso, sea por fuga o por alteración o desaparición de pruebas⁷.

Dado que técnicamente es imposible probar algo que no ha ocurrido, a no ser que se pueda vincular causalmente de forma necesaria con un hecho que ya ha acontecido (FERRER BELTRÁN 2019: 205), si se excluye la efectiva práctica de la nigromancia por parte del juzgador, resulta evidente la dificultad que este tendrá para apoyar su decisión. Pero entonces, ¿qué tipo de razonamiento efectúan los jueces cuando disponen la prisión preventiva u otra medida cautelar contra un individuo?; ¿es posible dotar de un sustrato epistémico robusto tales decisiones?

Siendo que, como bien señala GONZÁLEZ LAGIER, dentro de la estructura de cualquier inferencia probatoria el elemento esencial es la *garantía* (2013: 40), constituida por máximas de experiencia, presunciones o definiciones, resulta imprescindible explorar qué tipo de producto cognitivo puede alimentar una determinación tan relevante como la acá analizada. En esta empresa, el conocimiento derivado de la psicología cognitiva y los progresos relacionados con la inteligencia artificial podrían significar un avance, pero también, tal y como en el filme de SPIELBERG, un enorme retroceso en la búsqueda de decisiones judiciales más racionales en estos rubros. A estas y otras disquisiciones se destinarán las siguientes páginas.

2. JUSTIFICACIÓN DE LA PREMISA FÁCTICA DE LAS MEDIDAS CAUTELARES

Aunque la premisa normativa de las medidas precautorias, conforme con la concepción cautelar⁸, es susceptible de fundadas críticas⁹, es la premisa fáctica la que

⁵ El análisis parte del diseño procesal penal costarricense, el cual es, en líneas generales, similar al de la mayoría de los países latinoamericanos, al tener como base el Código Procesal Penal modelo para Iberoamérica de 1988. Sobre esta uniformidad relativa, particularmente en la regulación de las medidas cautelares, puede verse: DEI VECCHI 2015; LLOBET 2018: 590–591; MAIER 1996.

⁶ Este rubro implica que se corrobore la probabilidad de que se haya dado un hecho ilícito y la participación del investigado en este. Se ha querido presentar este aspecto material de la premisa normativa de esta forma a fin de evidenciar que se trata de una proposición diversa a la relativa a la decisión final, lo que permitiría establecer estándares de prueba idénticos en ambos casos, sin que signifique una «sentencia temprana». Sobre este punto se volverá en los apartados siguientes.

⁷ Se asume *stricto sensu* la concepción cautelar, con el único fin de facilitar el análisis de los hechos futuros en las medidas cautelares, pero sin inobservar que la mayoría de los países latinoamericanos han incorporado en sus codificaciones causales de orden punitivo, contrariando abiertamente los lineamientos que ha erigido la Corte Interamericana de Derechos Humanos. Sobre el tema puede consultarse: MORA 2019: 220–222; MORA 2015: 83–95.

⁸ Sobre la concepción cautelar *vid.*: DEI VECCHI 2015a; MORA 2019: 216–218.

Nota 9 en página siguiente

es la premisa fáctica la que comporta mayor dificultad operativa para los administradores de justicia penal. Las complicaciones pueden provenir de al menos tres tipos de variables: los factores procesales positivizados en las legislaciones; los factores de resiliencia interna del juzgador; y los factores estructurales de la pretensión.

2.1. Factores procesales positivos

Independientemente del diseño procesal del sistema penal que se trate, por la naturaleza de las medidas cautelares, su requerimiento se da en etapas tempranas de la investigación, lo que implica, habida cuenta de la necesidad de restringir potenciales detenciones arbitrarias, la fijación de plazos cortos para definir la situación jurídica del imputado, a partir del momento en que finalmente es aprehendido¹⁰.

Si a lo dicho se suma el exceso de carga laboral¹¹ y las presiones políticas para la *macdonalización*¹² de la justicia penal, se comprenderá sin dificultad que no nos encontramos ante un escenario ideal para la evaluación de restricciones en derechos fundamentales de personas que no han sido condenadas y que, en principio, se ven resguardadas por el principio de presunción de inocencia.

2.1.1. *Incompletitud del elenco probatorio*

Uno de los parámetros de justificación epistémica que la doctrina especializada ha desarrollado es precisamente la completitud (*comprehensiveness*)¹³ o inclusividad (FERRER BELTRÁN 2019: 273), que se refiere a la cantidad de pruebas relevantes dis-

⁹ Esencialmente, en cuanto al *fumus commissi delicti*, se cuestiona su compatibilidad con la presunción de inocencia como regla de trato (sobre este particular *vid.* FERNÁNDEZ LÓPEZ 2004: 218–222; FERRER BELTRÁN 2019: 195–199), mientras que en lo relativo al *periculum libertatis*, las principales críticas derivan del establecimiento de causales que persiguen fines punitivos y no precautorios (*vid.*: MORA 2019: 216–224).

¹⁰ El Código Procesal Penal costarricense (en adelante CPP) establece, en su art. 238: «(...) Si la persona se encontrare detenida, la solicitud de audiencia deberá pedirse dentro de las veinticuatro horas, contadas desde que el encausado se puso a la orden del juez; la audiencia deberá celebrarse dentro de las cuarenta y ocho horas y la resolución deberá ser dictada dentro de ese plazo».

¹¹ Aunque de acuerdo con el III Informe sobre el Estado de la Justicia, pese al incremento en la cantidad de jueces, entre los años 2016 y 2018 se muestra una disminución de los asuntos resueltos de forma definitiva en materia penal, con un total de 163.630 casos, lo que representa una disminución del -0,1 por 100, entre un año y otro (2020: 170).

¹² Las limitaciones de mi memoria me han impedido recordar dónde leí por primera vez esta frase, no obstante, la referencia corresponde a Manuel MIRANDA ESTRAMPES, quien se refirió a la «macdonalización y guantamización de los sistemas latinoamericanos por influencia norteamericana», durante las lecciones que impartió en el seno del Máster en Razonamiento Probatorio de la Universidad de Girona, en el año 2018.

¹³ «(...) *comprehensiveness is one determinant of quality of evidence that thorough inquiry requires not only sifting and weighing the available evidence, but also, when necessary, seeking out additional evidence*

ponibles para alcanzar una decisión. La premura con que se desarrolla la evaluación de medidas cautelares impide que se cuente, en la mayoría de los casos, con la totalidad de las pruebas relevantes para la determinación de la verdad, dado que, aunque la detención del imputado se haya dado luego de una exhaustiva investigación —piénsese, por ejemplo, en casos de delincuencia económica— y el Ministerio Público lograra acumular una enorme cantidad de evidencia, en este supuesto sería la defensa la que se encontraría en aprietos para procurar prueba de descargo, con las limitaciones que conlleva el tiempo para hacerlo, el estado psicológico del imputado —posiblemente en *shock*— y la restricción momentánea de su libertad ambulatoria.

La imposibilidad material de que la defensa procure prueba puede repercutir negativamente en los intereses del investigado, ya que hipótesis alternativas en favor de la libertad del imputado podrían no ser respaldadas epistémicamente. Pero en muchas otras ocasiones la situación es aún más grave, pues en supuestos de sucesos de violencia intempestiva, en donde no se ha podido realizar investigación previa, tanto la prueba de cargo como de descargo se ve sensiblemente constreñida.

Supongamos un caso de homicidio, en donde solamente existe un testigo que afirma haber visto al investigado corriendo desde el lugar de los hechos, con manchas rojas en su camisa, encontrándose pendiente toda la prueba pericial que usualmente se dispone en este tipo de asuntos, incluyendo análisis balísticos, de huellas dactilares, de elementos pilosos u otro material orgánico, etc. Resulta claro que los estudios científicos podrían dotar de un alto grado de confirmación a la hipótesis del ente acusador, o bien, descartarla, también con un importante grado de solidez; sin embargo, el juez deberá determinar la condición en que el investigado afrontará el resto del proceso -en libertad, en prisión, con un mecanismo electrónico de seguimiento u otra medida- prescindiendo de esa evidencia.

2.1.2. *Devaluación de la garantía del derecho a la prueba*

Si bien puede discutirse que el vocablo prueba se utilice en contextos distintos a la etapa de juicio, «desde el punto de vista epistemológico no hay distinción alguna (...) [t]odo elemento de juicio relevante para tomar una decisión sobre la ocurrencia de un hecho es una prueba»¹⁴ (FERRER BELTRÁN 2019: 194), con independencia de si se les denomina elementos de convicción, evidencia, indicios o cualquier otra nomenclatura. Empero, en las audiencias de medidas cautelares las garantías correlativas al derecho de defensa y vinculadas con la actividad probatoria se ven a menudo disminuidas, especialmente porque no se presenta una práctica de la prueba siquiera

(*which reminds me to remind you of the two meanings of “partial”: “biased” and “incomplete”*)» (HAACK 2014: 31).

¹⁴ Cfr. VALENZUELA 2018: 843, quien afirma que «pareciera que la prueba en el sistema procesal es aquella que tiene lugar en la audiencia de juicio oral pues el legislador usa ese vocablo (prueba o medios de prueba) cuando se diseñan las condiciones para la decisión del juicio».

semejante a la que se genera en la etapa de juicio, donde el principio de contradicción resulta determinante.

Aunque el CPP costarricense prevé la posibilidad de recibir (y practicar) prueba para la aplicación de las medidas cautelares, incluso de forma oficiosa por parte del juez¹⁵, esta no podría ser tomada en cuenta en juicio, salvo que se utilice el procedimiento de prueba anticipada (LLOBET 2017: 403). En la práctica judicial es el Ministerio Público el que suele recibir la denuncia y entrevistar testigos, sin presencia del encartado o su defensor¹⁶, de manera que lo que se lleva a la audiencia para sostener la gestión fiscal son los documentos en donde constan las manifestaciones de los involucrados.

El problema de la inclusión de manifestaciones de individuos mediante documentos puede visualizarse al menos desde dos frentes: en primer lugar, por la imposibilidad de ejercer el derecho de defensa y el principio de contradicción con respecto a la fuente de prueba, ya que no puede cuestionarse directamente la información vertida por el denunciante o el testigo¹⁷.

En segundo lugar, por la existencia de heurísticos¹⁸ que acarrearán sesgos de información como el denominado sesgo del entrevistador (CERDA LORCA y VALDIVIA CABRERA 2020), según el cual quien recepta la información interpreta la misma y la modula a la hora de reflejarla por escrito, con el potencial riesgo de variar el mensaje en aspectos relevantes, siendo esto notorio, por ejemplo, cuando el lenguaje utilizado no se corresponde con las características socio-educativas del entrevistado, entendiéndose la jerga judicial.

¹⁵ Art. 242: «Prueba para la aplicación de medidas cautelares. El fiscal o, en su caso el Tribunal, podrán recibir prueba, de oficio o a solicitud de parte, con el fin de sustentar la aplicación, revisión, sustitución, modificación o cancelación de una medida cautelar. Dicha prueba se agregará a un legajo especial cuando no sea posible incorporarla al debate. El tribunal valorará estos elementos de prueba conforme a las reglas generales establecidas en este Código y exclusivamente para fundar la decisión sobre la medida cautelar. Si el tribunal lo estima necesario, antes de pronunciarse, podrá convocar a una audiencia oral para oír a las partes o para recibir directamente la prueba. De dicha audiencia se levantará un acta».

¹⁶ Lo que ha sido avalado por la Sala Constitucional de Costa Rica, mediante el voto 13255-11 (citado por: LLOBET 2017: 403).

¹⁷ Este punto es fundamental si se repara en que, en muchas ocasiones, por ejemplo, en agresiones sexuales o violencia de género, solamente se contará en la audiencia con la denuncia por escrito de la presunta víctima.

¹⁸ De interés puede mencionarse el heurístico de la representatividad, que sugiere que las personas toman sus decisiones a partir de situaciones que les resultan familiares (NIEVA FENOLL 2010: 121), lo que se ajusta a la labor que realizan los secretarios o los fiscales que toman denuncias en multiplicidad de causas penales y que puede arrastrarlos a *acomodar* la información que reciben, aun de manera inconsciente.

2.2. Factores de resiliencia interna del decisor

Algunos estudios empíricos han mostrado cómo una multiplicidad de factores externos puede incidir de manera determinante en el juzgador, al momento de resolver una solicitud de imposición de medidas cautelares, relacionados con la participación y cobertura de medios de comunicación masiva, el criterio de la sociedad civil y los órganos políticos e incluso con la instrucción de causas disciplinarias en su contra¹⁹.

Cómo inciden estos factores en el juez dependerá de un amplísimo rango de vectores, que van desde su madurez, nivel de conocimiento, situación familiar, estabilidad laboral, etc. El llamado *miedo social al falso negativo* (MUÑOZ VICENTE y LÓPEZ-OSSORIO 2016: 134), concepto tomado de la psicología cognitiva y que denota el sesgo que llevaría al juez a disponer una medida cautelar ante la incertidumbre de que, de no hacerlo, pueda el imputado incurrir en una conducta indeseable, con lo que se incrementa, por razones obvias, la posibilidad de falsos positivos, también es un elemento que merece atención y que podría ser reflejo de la incidencia de los factores externos descritos, pero su génesis podría también encontrarse en la alineación ideológica del juez, por ejemplo, si ha resultado ser un seguidor de posturas autoritarias o expansionistas del derecho penal.

De cualquier forma, siendo inasequible lo que ocurre en la mente del decisor, los esfuerzos para morigerar posibles influencias externas o el influjo de corrientes ideológicas que choquen con un modelo de estado liberal, es preciso fortalecer el principio de independencia judicial, en donde ya se cuenta con un plexo normativo internacional amplio²⁰, así como el control de la racionalidad de las resoluciones, empresa esta última que, como se verá, no resulta simple, habida cuenta de las características que poseen los enunciados a probar para aplicar una medida cautelar.

¹⁹ *Vid.*, entre otros: CHINCHILLA CALDERÓN y GARCÍA AGUILAR 2003; MORA 2019; VEGA MONGE 2015. En el II Informe sobre el Estado de la Justicia, se presentaron entrevistas realizadas a 763 funcionarios judiciales (funcionarios de la judicatura, la defensa pública y la policía judicial), en donde se consultó sobre su percepción en cuanto a los criterios para la disposición de prisión preventiva, se señalaron aspectos como los prejuicios contra determinados grupos, el criterio de asociaciones de víctimas, la posibilidad de enfrentar procesos disciplinarios, la opinión pública y los medios de comunicación, entre otros (2017: 285–86). El informe puede consultarse en: www.estadonacion.or.cr

²⁰ La importancia de establecer un abordaje integral de la independencia judicial se imbuye en los «Principios de Bangalore sobre la Conducta Judicial» (aprobados por el Consejo Económico y Social de las Naciones Unidas en su resolución E/CN.4/2003/65/Anexo, en la Haya, en noviembre de 2002 y adoptados el 10 de enero de 2003), en donde se subrayó en su preámbulo que, dada la relevancia de la independencia judicial, los diferentes actores sociales (Poder Ejecutivo, Poder Legislativo, abogados y público en general) deben comprender y apoyar el ejercicio de la judicatura. Sobre el tema *vid.* además: MORA 2015: 38–48.

2.3. Factores estructurales de la pretensión

Conforme con el esquema elaborado por Stephen TOULMIN, la *pretensión* (GONZÁLEZ LAGIER 2013: 40) o *afirmación* (TOULMIN 2007: 132–39) constituye el punto de partida de cualquier argumentación. La particularidad de las medidas cautelares es que en estas esa pretensión, tanto en lo atinente al *fumus commissi delicti* como en lo relativo al *periculum libertatis*, es un enunciado formulado de manera prospectiva, es decir, se trata de una hipótesis sobre hechos futuros (VALENZUELA 2018: 841–42).

Como ya se ha adelantado, se asume en este trabajo la postura concebida por FERRUA, según la cual el *fumus commissi delicti* es una hipótesis diversa a la premisa fáctica relativa a la decisión final, en tanto, mientras que en el primer caso debe probarse la *probabilidad* de que el proceso culmine con una sentencia condenatoria contra el imputado, es decir, un pronóstico sobre el resultado del proceso en sí mismo²¹, en el segundo, la prueba será dirigida a la acreditación concreta del hecho y la participación del encartado.

Como puede verse, la diferencia radica en que la premisa fáctica del *fumus commissi delicti* («es probable que el proceso culmine con una sentencia condenatoria contra Fausto») incorpora una hipótesis probabilística que trasciende los aspectos probatorios relacionados con la justificación de la hipótesis de la sentencia. Al considerar que se trata de proposiciones diversas, se sortea el problema de mantener estándares probatorios idénticos y provocar una «sentencia temprana» (FERRUA 2018: 120) con el dictado de medidas cautelares altamente invasivas como la prisión preventiva²². En este sentido, NIEVA ha estimado como un «error común» que los jueces desatiendan este aspecto y se concentren exclusivamente en el *periculum libertatis*, con lo que ignoran «el presupuesto principal que se está comentando: la constancia de un hecho con caracteres de delito y su prácticamente indudable atribución a una persona» (2018: 72), es decir, se trata de un umbral probatorio asimilable al necesario para condenar.

La asunción de la concepción cautelar trae consigo, necesariamente, que a las medidas cautelares se le atribuyan las características de instrumentalidad, excepcionalidad, legalidad y jurisdiccionalidad (DEI VECCHI 2015b: 35), siendo a partir de esta última propiedad que se torna imprescindible la justificación, en el caso concre-

²¹ Agradezco a Diego DEI VECCHI por sugerir este punto.

²² FERRUA considera que la determinación de diversos estándares de prueba para distintas fases del proceso penal constituye «una perspectiva inaceptable porque, al disminuir el estándar de la prueba más allá de toda duda razonable, el concepto mismo de probar se omite y se entra en el ámbito de las ilusiones y sospechas» (2018:120; traducción libre). Cfr. FERRER BELTRÁN 2019: 208, para quien, en el caso de la prisión preventiva, «la exigencia probatoria deberá ser superior que la prevista para el inicio de actuaciones penales pero inferior al estándar de prueba establecido para la condena».

to, de la premisa fáctica de la medida²³, lo que requiere un ejercicio calificadorio y el despliegue del razonamiento inferencial probatorio²⁴.

Según lo señalado, mientras que la justificación del *fumus commissi delicti* coincide parcialmente con la justificación de la decisión final (existencia del hecho y participación del imputado), pero con la relevante diferencia apuntada, relativa a la incorporación de una hipótesis probabilística sobre el resultado del proceso, la gran dificultad que se presenta para justificar adecuadamente el *periculum libertatis* surge de las características del supuesto de hecho previsto en la premisa normativa, pues se trata de un hecho futuro²⁵, en donde se evalúa una peligrosidad de segundo orden, dado que es preciso realizar una prognosis sobre la producción de un evento relevante y, además, de la producción de un resultado concreto a partir de ese evento (DEI VECCHI 2015a: 3).

En otras palabras, se trata de una especie de *peligro de peligro*, una hipótesis probabilística doble, en donde el pronóstico no solamente se da con respecto a la creación de cierta secuela, sino también en cuanto al *hecho base* (DEI VECCHI 2015a: 3). Partiendo de los peligros procesales regulados en el CPP²⁶, sean el peligro de fuga²⁷ y el riesgo de obstaculización²⁸, la justificación de la premisa fáctica tendría que consi-

²³ De acuerdo con DEI VECCHI, para tener por probada una premisa fáctica se debe: probar un enunciado acerca de un hecho (acá se imbuyen las categorías de enunciado probatorio e inferencia probatoria); determinar si ese enunciado es una instancia particular del supuesto de hecho de la premisa normativa (incorpora las categorías enunciado calificadorio y razonamiento calificadorio) (2015b: 2).

²⁴ Sobre la problemática que entraña la diferenciación entre *quaestio iuris* y *quaestio facti*, puede verse: GONZÁLEZ LAGIER 2013: 29–36. El autor señala: «La calificación jurídica opera ciertamente con hechos, pero es una operación interpretativa que no podría realizarse sin normas; y -lo que es más grave y menos evidente- la prueba también está teñida de normatividad» (2013: 35).

²⁵ «La prueba de hechos futuros o los *future dangerousness assessments*, como se les conoce en la doctrina anglosajona, presenta una enorme trascendencia en los procesos penales, principalmente en la determinación de medidas cautelares y prisión preventiva (...) En cualquier caso, se trata de acontecimientos futuros, no susceptibles de prueba directa, que se asientan en la conjetura de su eventual existencia» (MORA 2018). También utilizando esta nomenclatura, *vid.* HUNTER AMPUERO 2015.

²⁶ Como ya se ha dicho, por no corresponder con la concepción cautelar, se prescindirá en este análisis de los riesgos de tipo punitivo.

²⁷ Art. 240 CPP: «*Peligro de fuga*. Para decidir acerca del peligro de fuga se tendrán en cuenta, especialmente, las siguientes circunstancias: a) *Arraigo en el país*, determinado por el domicilio, residencia habitual, asiento de la familia, de sus negocios o trabajo y las facilidades para abandonar definitivamente el país o permanecer oculto. La falsedad, la falta de información o de actualización del domicilio del imputado constituirá presunción de fuga. b) La pena que podría llegarse a imponer en el caso. c) La *magnitud del daño* causado. d) El *comportamiento del imputado* durante el procedimiento o en otro proceso anterior, en la medida que indique su voluntad de someterse a la persecución penal» (el destacado es propio).

²⁸ Art. 241 CPP: «*Peligro de obstaculización*. Para decidir acerca del peligro de obstaculización para averiguar la verdad se tendrá en cuenta, especialmente, la *grave sospecha* de que el imputado: a) Destruirá, modificará, ocultará o falsificará elementos de prueba. b) Influirá para que coimputados, testigos o peritos informen falsamente o se comporten de manera desleal o reticente, o inducirá a otros a realizar tales comportamientos. El motivo sólo podrá fundar la prisión hasta la conclusión del debate» (el destacado es propio).

derar, por ejemplo, la probabilidad de que Fausto intente salir del país para eludir la acción de la justicia, además de la probabilidad de que Fausto en efecto salga del país y eluda la acción de la justicia.

Desde la perspectiva señalada, prohiendo la postura sostenida por DEI VECCHI, se considera que solamente es posible probar, en estricto sentido, el *periculum libertatis*, enlazando la noción de riesgo o peligro con la ejecución de acciones concretas (con lo que se asimila el peligro procesal al *fumus commissi delicti*, en cuanto al mecanismo inferencial necesario para probarlo, como una probabilidad de primer orden), pese al coste político evidente que esa asunción acarrearía²⁹.

Lo anterior se explica en razón de que, como se verá en los capítulos siguientes, en cualquier esquema argumentativo, si las *razones* (conforme con la nomenclatura de TOULMIN) también son prospectivas, es decir, no tienen un referente empírico, sea que estas las defina el legislador mediante el recurso de los llamados *indicios tipo* o sea el juez el que seleccione cuales son las razones o los datos necesarios y suficientes, se establecería un vínculo conceptual que es ajeno al razonamiento probatorio como tal (DEI VECCHI 2015a: 8)³⁰.

Tomando en consideración lo expuesto hasta este punto, en el capítulo siguiente se trazarán algunas líneas con relación al diseño de la inferencia probatoria en hechos prospectivos, con especial énfasis en los elementos de la *garantía* y el *respaldo*.

3. LA INFERENCIA PROBATORIA EN LOS HECHOS FUTUROS

El modelo simple de análisis de los argumentos, de acuerdo con la teoría de Stephen TOULMIN, identifica cuatro elementos en cualquier género de razonamiento: pretensión (*claimt*), razones (*grounds*), garantía (*warrant*) y respaldo (*backing*) (ATIENZA 2005: 84–87; GONZÁLEZ LAGIER 2013: 40–44; TOULMIN 2007b). Conforme con este modelo, traspuesto al razonamiento judicial en materia de medidas cautelares, la *pretensión* sería la hipótesis inicial, por ejemplo, el enunciado «Fausto en libertad se fugará e impedirá la continuación del proceso»; si la pretensión se cuestiona, el proponente tendrá que ofrecer las *razones*, que son los hechos que constituyen el supuesto de hecho de la norma aplicable al caso en cuestión (v.gr. «Fausto no posee

²⁹ «El único modo de sortear este obstáculo, sin renegar de una genuina finalidad cautelar y manteniendo una noción de peligrosidad procesal susceptible de prueba, sería ligando esta última noción a la ejecución de acciones procesalmente peligrosas concretas por parte de ciertas personas, a modo de tipos coercitivos semejantes a las figuras que determinan las acciones punibles. Es decir, resulta insoslayable el abandono de la noción de *periculum libertatis* como peligro de segundo orden» (DEI VECCHI 2015a: 9). Cfr. FERRER BELTRÁN 2019:205; NIEVA FENOLL 2018: 79.

³⁰ En sentido similar, FERRER BELTRÁN ha considerado que «es extraordinariamente difícil acreditar suficientemente la hipótesis de riesgo procesal», salvo aquellos casos en que el investigado ha intentado ya acciones tendientes a su fuga o la desaparición de pruebas incriminatorias (2019: 207).

arraigo domiciliario en el país»); para autorizar o explicar el paso de un enunciado a otro, es decir, de las razones a la pretensión, se utiliza la *garantía*, que se presenta como una máxima de experiencia, una presunción o definiciones o teorías (GONZÁLEZ LAGIER 2013: 43) (v.gr. «Las personas que no tienen domicilio frecuentemente se fugan»); finalmente se presenta el *respaldo*, que serviría como andamiaje de la garantía (v.gr. «estudios efectuados desde la psicología cognitiva demuestran que las personas sin domicilio tienden a fugarse»).

Como puede verse, el esquema argumentativo descrito es formalmente aplicable al razonamiento judicial que se presenta al evaluar la aplicación de medidas cautelares. Sin embargo, en la práctica la *garantía* no suele explicitarse, con lo que la inferencia se convierte en un razonamiento entimemático o, como bien refiere Dei Vecchi, tendría que concluirse que se está ante resoluciones arbitrarias (2015a: 6), siendo una incógnita la manera en que los jueces razonan *realmente*³¹, lo que a su vez mengua la adecuada fundamentación de la resolución judicial.

Asimismo, el mayor inconveniente que se presenta al trabajar con hechos prospectivos es que, a diferencia de los hechos retrospectivos, en donde se trata de corroborar o afirmar como verdadera la ocurrencia de un hecho, en estos el enunciado a probar como tal es una hipótesis probabilística³² (DEI VECCHI 2015a: 3), cuestión que dificulta sobremano fundar suficientemente la *garantía*, máxime cuando se trate de medidas cautelares como la prisión preventiva o el monitoreo mediante dispositivos electrónicos, que requieren rebasar, al menos idealmente, un riguroso estándar de prueba, habida cuenta del enorme impacto que tienen en la esfera de derechos del imputado³³.

³¹ La ambivalencia en cuanto a la forma en que asumen los operadores de justicia los denominados *indicios tipo* imbuidos en el CPP costarricense se ha reflejado en una investigación realizada mediante una entrevista semiestructurada a 43 funcionarios vinculados con la administración de justicia (16 mujeres y 27 hombres), en donde el 77 por 100 de los entrevistados señaló que tales indicadores eran presunciones *iuris tantum*, el 8 por 100 que eran indicios taxativos que necesariamente debían acreditarse y solamente el 15 por 100 estimó que se trataban de indicios ejemplificativos (MORA 2019: 230). Los datos resultan relevantes en tanto, dependiendo de la forma en que se entienda el indicio tipo, así corresponderá la clase de garantía, sea una máxima de experiencia, una presunción o una definición.

³² Como ya se ha indicado, incluso de *segundo orden*, en el caso del *periculum libertatis*.

³³ Aunque la determinación del estándar de prueba aplicable es una decisión de política criminal, lo cierto es que el juicio de ponderación de intereses en juego parece hacer recomendable que se otorgue prevalencia a la afectación concreta al individuo, cuando «el mal que se trata de evitar sea uno de producción incierta respecto del cual no puede formularse más que un pronóstico de peligro más o menos fundado, mientras que el mal que se produce (...) es un menoscabo prolongado e indefinido del derecho fundamental a la libertad real, efectivo y actual, además de irreparable, es demostrativo por sí solo de la considerable mayor gravedad del mal causado con respecto al que se trata de evitar» (GRACIA MARTÍN, citado por MARTÍNEZ GARAY 2014: 55).

3.1. ¿Jueces gurús, *bouche de la loi* o decisores racionales?

Quizá el elemento más relevante en el razonamiento inferencial sea la *garantía* y su correspondiente *respaldo*, entendido este último como «el fundamento cognoscitivo en virtud del cual la inferencia puede considerarse fundada y apta para atribuir un grado considerable de fiabilidad a la respectiva conclusión» (TARUFFO 2010: 237). De acuerdo con González Lagier la garantía puede estar constituida por máximas de experiencia, presunciones o por definiciones o teorías (2013: 43), lo que está íntimamente relacionado con el diseño procesal que se disponga para un acto determinado.

El CPP costarricense establece, con relación a las causales para la aplicación de medidas cautelares, específicamente en cuanto al riesgo de fuga, algunos *indicios tipo* que podrían entenderse como presunciones legales, ya que incorporan circunstancias que el decisor deberá tener «especialmente» en cuenta, a saber: arraigo en el país (domiciliar, familiar y laboral); facilidades para abandonar definitivamente el país u ocultarse; falsedad o falta de información sobre el domicilio (lo que constituye «presunción de fuga»); la posible pena por imponer; la magnitud del daño causado; el comportamiento del imputado, en cuanto refleje su voluntad de someterse a la persecución penal³⁴.

El punto determinante es si dentro de un sistema de prueba libre, regido bajo la etérea noción de la sana crítica racional, pero que, en esencia, requiere razones epistémicas para fundar una decisión, podrían utilizarse presunciones o definiciones como enlace de la inferencia, más aún cuando en juego está uno de los bienes jurídicos de mayor valía, como lo es la libertad. Asimismo, si el enlace es una máxima de experiencia, es preciso establecer aquellas características del *respaldo* que permitirían dar solidez a la conclusión.

3.1.1. *Presunciones iuris o definiciones*

Si se estima que los *indicios tipo*, descritos en el acápite anterior, se tratan de presunciones jurídicas, estas operan como una fórmula de distribución de la carga de la prueba³⁵, si la presunción es refutable (presunción *iuris tantum*), o como una regla jurídica inderrotable, si la presunción es concluyente (presunción *iuris et de iure*) (SCHAUER 2013: 228–32). En cualquier caso, los datos constituidos por los referidos

³⁴ Art. 240 del Código Procesal Penal. *Vid.* la nota al pie número 31, en cuanto a la forma en que las personas vinculadas con la administración de justicia han entendido esta regulación normativa.

³⁵ Sobre la carga de la prueba, de forma crítica, *vid.* NIEVA FENOLL, FERRER BELTRÁN, y GIANNINI, 2019.

indicadores tendrían como *garantía*, no un componente fáctico, sino un elemento normativo, que los excluye como genuinas pruebas³⁶.

De esta forma, para que haya una genuina justificación probatoria de la premisa fáctica del *periculum libertatis*, necesariamente deben entenderse los indicadores referidos como presunciones *hominis*, es decir, como indicios³⁷ o datos cuyo respaldo se constituye por consideraciones empíricas, por máximas de la experiencia (DEI VECCHI 2015a: 6)³⁸. En síntesis, «[l]a diferencia entre presunciones [*iuris*] y máximas de experiencia es que las presunciones son enunciados revestidos de autoridad» (GONZÁLEZ LAGIER 2013: 44), que pueden verse también como «máximas de experiencia institucionalizadas y autoritativas» (GONZÁLEZ LAGIER 2013: 59), lo que hace recordar aquella famosa frase de MONTESQUIEU, plasmada en «El espíritu de las leyes», que propugnaba que el juez es la boca muda que pronuncia las palabras de la ley (FERNÁNDEZ SEPULVEDA 1985: 133).

En cuanto a las definiciones o teorías, de acuerdo con GONZÁLEZ LAGIER, estas pueden derivar de la doctrina, la jurisprudencia o inclusive de la ley, pero, en cualquier caso, al tratarse de un vínculo de carácter conceptual, en realidad no se estaría ante un razonamiento probatorio, sino más bien frente a una interpretación o calificación de hechos (2013: 44)³⁹.

Como puede verse, las características de este género de garantías las tornan inviables para servir de enlace en una inferencia relativa a medidas cautelares penales, en donde la Corte Interamericana de Derechos Humanos ha sido enfática sobre la necesidad de fundamentar adecuadamente las resoluciones que puedan afectar derechos humanos⁴⁰, así como verificar la existencia del *periculum libertatis* a partir de las circunstancias del caso concreto y no a través de presunciones⁴¹.

³⁶ «(...) el argumento determinante para descartar las presunciones *iuris*, en cualquier versión, como argumentos aptos para dar por acreditada la peligrosidad procesal tal que el encarcelamiento cautelar es indispensable, está en que esta clase de presunciones no son prueba en absoluto» (DEI VECCHI 2015a: 5–6).

³⁷ «El indicio, por tanto, no es equivalente a una presunción sino que constituye, insistimos, el hecho base de la presunción» (MIRANDA ESTRAMPES 1997: 227).

³⁸ «(...) su idoneidad inferencial en términos probatorios dependerá de que sean aptos como datos indiciarios, lo cual requerirá de la presencia de un respaldo fáctico justificado. En otras palabras, es insoslayable una fundamentación empírica, y no normativa, que cimiente la regla de paso del dato indiciario a la pretensión, tal que esa regla constituya una genuina máxima de experiencia» (DEI VECCHI 2015b:147). Desde una perspectiva similar, HUNTER AMPUERO habla de indicios tipo, enfatizando que «deben funcionar como guías y no como reglas» (2015: 246).

³⁹ En el mismo sentido: DEI VECCHI, 2015a: 8.

⁴⁰ Cfr. CIDH. Caso *Apitz Barbera y otros vs. Venezuela*, sentencia del 5 de agosto de 2008, párrafo 78; *Tristán Donoso vs. Panamá*, sentencia del 27 de enero de 2009, párrafo 153; *López Mendoza vs. Venezuela*, sentencia del 1 de septiembre de 2011, párrafo 148. Vid. también: LLOBET 2018.

⁴¹ Cfr. CIDH. Caso *Norín Catrimán y otros vs. Chile*, sentencia del 29 de mayo de 2014, párrafo 312. Se indica en esta resolución que «el peligro procesal no se presume, sino que debe realizarse la verificación de este en cada caso concreto, fundado en circunstancias objetivas y ciertas del caso concreto».

Ahora bien, el uso excesivo de las medidas cautelares de alto impacto, como la prisión preventiva o los mecanismos electrónicos de seguimiento, podría deberse precisamente a la automatización en su aplicación⁴², correlativa a la asunción de presunciones o la definición de los riesgos, operación que simplifica muchísimo una decisión de este tipo, con lo que parece que, en consonancia con lo dicho por DEI VECCHI, se trata de forzar argumentos basados en ciertas intuiciones comunes (2015a: 8), en donde el juez tendría más de gurú o nigromante, que de decisor racional.

Entonces, habiéndose mostrado que en el rubro objeto de este trabajo la *garantía* del razonamiento inferencial no puede tener la forma de una presunción *iuris* o una definición, queda por evaluar qué tipo de máximas de experiencia se pueden utilizar y, principalmente, cuál es (o debe ser) su *respaldo*.

3.1.2. Máximas de experiencia

Como ya se ha adelantado, las máximas de la experiencia son regularidades que constituyen a su vez la conclusión de una inducción ampliativa, probables en sentido inferencial, lo que implica que se basen en una inducción ampliativa sólida si se quiere que estas estén debidamente fundamentadas (GONZÁLEZ LAGIER 2013: 59). En ese sentido, se han delineado al menos cuatro modelos en que pueden fundarse las máximas de experiencia (TARUFFO 2010: 237–40): leyes científicas de carácter universal, de donde se sigue una inferencia bajo el modelo nomológico-deductivo que permite atribuir certeza deductiva a la conclusión; generalizaciones no universales con alto grado de probabilidad, de donde se sigue una inferencia bajo el modelo cuasi nomológico deductivo que permite atribuir certeza práctica a la conclusión; generalizaciones de sentido común (*id quod plerumque accidit*), las cuales aportan un grado de confirmación variable a la conclusión, pero no especialmente alto; y generalizaciones espurias (pseudoreglas), sin ningún fundamento empírico, sino más bien alimentadas por prejuicios.

Dado que las medidas cautelares, en especial la prisión preventiva, pueden afectar sensiblemente una variedad importante de derechos fundamentales del imputado⁴³,

⁴² En el ámbito costarricense, de acuerdo con el II Informe sobre el Estado de la Justicia, entre los años 2005 y 2015, el 54.2 por 100 de las personas detenidas preventivamente correspondieron a causas ligadas a la narcoactividad, concluyéndose en dicho informe que en las infracciones a la Ley de Psicotrópicos la prisión preventiva es casi automática, a pesar de que la cantidad de condenas es baja en ese tipo de casos (un 13 por 100 del total de condenas entre 2010 y 2015), aunado a que, de una muestra aleatoria de 159 expedientes en los que se dictó prisión preventiva en 2016, en despachos de la Gran Área Metropolitana, solamente en un 8 por 100 de estos asuntos las personas decisoras resolvieron de forma distinta a la peticionada por el ente fiscal, incluyendo un caso en que se solicitó una medida no privativa de libertad, pese a lo cuál la jueza dictó prisión preventiva (2017: 82–83).

⁴³ No solamente se trata de la libertad ambulatoria. El estado de las prisiones en la mayoría de los países latinoamericanos es deplorable, con alarmantes niveles de hacinamiento. En Costa Rica, para el año 2020, se ha establecido un hacinamiento carcelario superior al 40 por 100 encontrándose cárceles

es claro que el estándar de prueba requerido, tanto para la corroboración del *fumus commissi delicti* como del *periculum libertatis*, debe ser especialmente alto, lo que deriva de la vigencia del principio de proporcionalidad y se refleja en la excepcionalidad de la prisión preventiva, regulada convencionalmente y recogida en la mayoría de las codificaciones de los países latinoamericanos, aunque se trate en realidad de una *excepcionalidad de papel*, que no se condice con la realidad de nuestras prisiones y la práctica judicial imperante, según se ha visto⁴⁴.

Por tanto, frente a un estándar de prueba riguroso es menester contar con inferencias debidamente construidas y fiables en lo tocante a su conclusión. Pero ¿cómo hacemos eso en hechos prospectivos? En el siguiente capítulo se valorarán algunas opciones para otorgar *respaldo* a la *garantía* en esta tipología de hechos.

4. EL *BACKING* DEL *WARRANT* EN LOS HECHOS FUTUROS

En este punto es posible afirmar que los hechos prospectivos, en sentido estricto, no son susceptibles de ser probados, en tanto no se puede corroborar algo que no ha sucedido. No obstante, siendo imprescindible dotar de contenido epistémico a la hipótesis probabilística que constituye el *riesgo* de ocurrencia de un evento⁴⁵, es preciso establecer si es viable en la actualidad predecir la conducta humana, con suficiente fiabilidad como para afectar la esfera de libertades y derechos de un sujeto determinado.

La respuesta a dicha interrogante podría encontrarse en la psicología cognitiva y la estadística o en la inteligencia artificial, disciplinas que se encuentran íntimamente relacionadas pero que se analizarán de manera independiente para brindar mayor claridad, en tanto la expresión de cada una estas áreas de conocimiento presentan especiales limitaciones, que pueden tornar cuestionables las conclusiones a las que se arribe.

con niveles superiores al 100 por 100 de sobrepoblación (Informe, 2020:182–183). Esas condiciones implican un riesgo a la salud de los reclusos, además de la afectación a su derecho al trabajo, su entorno familiar, etc.

⁴⁴ En cuanto a medidas cautelares diversas a la prisión o la vigilancia electrónica, parece evidente que el estándar de prueba sea menor, no obstante, la estructura del razonamiento inferencial es la misma, pues en cualquier caso debe partirse de la «prueba» de hechos prospectivos.

⁴⁵ Hablamos acá del *periculum libertatis* como peligrosidad de segundo orden, pues en cuanto al *fumus commissi delicti* su tratamiento es mucho más simple, al coincidir su justificación parcialmente con la justificación de la decisión final, la cual esta constituida por hechos retrospectivos. De igual manera, si se asume el *periculum libertatis* como una fórmula de peligrosidad de primer orden (lo que no se condice con la regulación normativa actual en el CPP costarricense y en la mayoría de las codificaciones de los países latinoamericanos), los problemas se reducen, ya que existen máximas de experiencia, como se verá más adelante, que permitirían establecer una inferencia mucho más sólida.

4.1. La psicología cognitiva y la estadística

La psicología cognitiva se ocupa de la «explicación de la conducta a entidades mentales, a estados, procesos y disposiciones de naturaleza mental, para los que se reclama un nivel de discurso propio» (POZO 2006: 42), lo que permitiría, en principio, conocer si el comportamiento del individuo está determinado por tales representaciones mentales⁴⁶ y, en esa medida, anticipar eventos.

Pese a que los procedimientos que se utilizan para estimar la probabilidad de que se genere cierto comportamiento en el futuro pueden aplicarse a una gran variedad de contextos jurídico-penales, desde la aplicación, mantenimiento o cese de beneficios carcelarios, medidas de seguridad post-delictuales, hasta medidas cautelares como la prisión preventiva u otras, en la actualidad la mayoría de los estudios especializados se han centrado en la valoración del riesgo de violencia en diferentes escenarios (violencia de pareja, en adolescentes, etc.)⁴⁷, siendo más bien escasa su aplicación en escenarios judiciales incipientes como los relacionados con las medidas precautorias, en parte quizá por la premura con que debe resolverse la situación jurídica del investigado, lo que tornaría impracticable una pericia sobre este en tan corto tiempo.

Al margen de lo señalado, es claro que la virtualidad práctica del conocimiento que se genere desde la psicología cognitiva no depende en absoluto de su utilidad actual, por lo que la información que se obtenga con respecto a otro tipo de riesgos futuros puede brindar luz sobre su potencialidad en el área objeto de este ensayo⁴⁸.

En los últimos lustros se han afinado y estandarizado una significativa variedad de instrumentos con capacidad predictiva del comportamiento violento futuro. Por ejemplo, solamente en cuanto al riesgo de violencia contra la pareja, potencialmente

⁴⁶ LACHMAN y BUTTERFIELD han considerado que «unas pocas operaciones simbólicas relativamente básicas, tales como codificar, comparar, localizar, almacenar, etc., pueden, en último extremo, dar cuenta de la inteligencia humana y la capacidad para crear conocimientos, innovaciones y tal vez expectativas con respecto al futuro» (citado por: POZO 2006: 42). Sin embargo, debe decirse que algunos autores como RIVIÈRE consideran el concepto de psicología cognitiva como «un caso más de categoría natural o mal definida dentro de las escuelas psicológicas» (citado por: POZO 2006: 42), aunque para efectos de este trabajo es absolutamente prescindible una fina delimitación conceptual, bastando con que se tenga claro que involucra también una lectura del comportamiento humano, con miras a prever conductas futuras. Sobre algunos aspectos históricos de la disciplina *vid.* MÉNDEZ PAZ 2015.

⁴⁷ PUEYO y ECHEBURÚA destacan el HCR-20 y la PCL-R para la violencia interpersonal grave, la SARA y la EPV en el caso de la violencia contra la pareja, el SVR-20, en casos de violencia sexual, y el SAVRY, en supuestos de violencia juvenil (2010: 403).

⁴⁸ De más está recordar que la evaluación del riesgo de violencia si tiene incidencia directa en lo tocante al peligro de obstaculización, cuando la prueba en riesgo sea testimonial, pues entran en juego posibles acciones violentas para «desanimar» al testigo, provocando incluso su muerte, siendo esta la razón más usual cuando se alega este riesgo procesal, alcanzando en Costa Rica más de un 34 por 100 del total de las gestiones de prisión preventiva en el 2016, frente a un 1.9 por 100 relacionado con otro tipo de pruebas. *Vid.* III Informe Estado de la Justicia, 2020: 283.

relevante para sostener una inferencia relacionada con la causal de obstaculización de prueba testimonial, se pueden mencionar las siguientes escalas: El *Danger Assessment Tool* (DA), la *Femicide Scale*, la *Spousal Assault Risk Appraisal Guide* (SARA), la *Ontario Domestic Assault Risk Assessment* (ODARA) y la Escala de Predicción de Riesgo de Violencia contra la Pareja (EPV)⁴⁹ (ECHEBURÚA et al. 2010: 1055), mientras que en delincuencia juvenil se cuenta con el IGI-J, el *Youth Level of Service/Case Management Inventory*, el YLS/CMI, entre otros (GARRIDO y GONZÁLEZ 2018: 5)⁵⁰.

Esto ha significado una transición de la valoración de la peligrosidad del individuo hacia la evaluación del riesgo de violencia (MUÑOZ VICENTE y LÓPEZ-OSSORIO 2016: 130; PUEYO y ECHEBURÚA 2010: 403)⁵¹, aunque parece que existe cierto acuerdo en cuanto a las limitaciones predictivas de los instrumentos, principalmente en virtud de la inestabilidad, variedad y simultaneidad en la ocurrencia de los factores de riesgo, aunado a la relativa infrecuencia de la conducta violenta (ECHEBURÚA et al. 2010: 1054; PUEYO y ECHEBURÚA 2010: 403–4; ROMEO CASABONA 2018: 45).

El punto de partida en la investigación rigurosa de los componentes de predicción de la conducta se suele ubicar en el modelo integrado de conducta delictiva propuesto por Andrews y Bonta en 1994 (MUÑOZ VICENTE y LÓPEZ-OSSORIO 2016: 132), quienes enfatizan sus estudios en la existencia de distintos factores de riesgo/protección. Los factores de riesgo incorporan elementos estáticos⁵² y dinámicos⁵³, mientras que los factores de protección imbuyen aquellas condiciones que morigerarían los primeros. Precisamente de su interacción se deriva la posibilidad de predicción de la conducta⁵⁴.

⁴⁹ SARA y EPV han sido validadas para su uso en España.

⁵⁰ Sobre los factores dinámicos en el comportamiento de delincuentes juveniles puede verse: NAVARRO y PASTOR, 2017.

⁵¹ Para una perspectiva histórica *vid.* ESBEK, 2003. La diferencia radica en que para valorar el riesgo es prescindible establecer las causas de la violencia. De esta manera, se entiende que la peligrosidad es tendencial del sujeto, mientras el riesgo se concentra en la probabilidad del peligro futuro, lo que implica que el riesgo de violencia se conciba como un constructo variable y específico, mientras el peligro sería estático y genérico (PUJOL ROBINAT et al. 2014: 17). Por su parte, PUEYO y REDONDO afirman que «el desarrollo de la psicología criminológica ha mostrado que la capacidad predictiva de la peligrosidad es limitada y su uso poco eficaz para los profesionales que toman decisiones prospectivas en contextos forenses, clínicos o penitenciario» (2007: 157). MARTÍNEZ GARAY sostiene una visión crítica sobre la distinción, al estimar que «desde luego que el concepto de peligrosidad es tremendamente impreciso, pero por desgracia seguramente ocurre algo parecido con el de valoración del riesgo» (2014: 9). Cfr. ROMEO CASABONA, 2018: 40–42.

⁵² Condiciones bio-psico-socioculturales de difícil variación (cuestiones de carácter hereditario, victimización, traumatismos, etc.).

⁵³ Condiciones bio-psico-socioculturales que si son susceptibles de modificación (desempleo, valores, psicopatologías reversibles, etc.).

⁵⁴ Debe anotarse que en la actualidad se parte de un concepto de causalidad probabilística, con lo que se superan los modelos deterministas y monocausales de antaño (MUÑOZ VICENTE y LÓPEZ-OSSORIO 2016: 131).

En cuanto a la metodología utilizada para la predicción del riesgo, la psicología cognitiva utiliza en la actualidad dos modelos diversos⁵⁵, el modelo clínico y el modelo actuarial. El modelo clínico (también llamado profesional estructurado o clínico estructurado), se basa en la perspectiva ideográfica (particular) y se enfoca en los factores dinámicos del riesgo, pero su debilidad radica en poseer un «elevado margen de subjetividad al descansar el juicio pronóstico en el profesional que realiza la evaluación» (MUÑOZ VICENTE y LÓPEZ-OSSORIO 2016: 135), por lo que la fiabilidad del resultado dependerá de la competencia del experto y los posibles sesgos que presente⁵⁶, aunado a que se trata de un aspecto que no es intersubjetivamente controlable.

Se ha llegado a afirmar que las predicciones de riesgo de violencia a partir del modelo clínico estructurado mejoran significativamente los pronósticos en poblaciones penitenciarias, enfermos mentales, agresores de pareja y sexuales, asimilándose su validez a los pronósticos obtenidos en otras áreas profesionales como la meteorología (PUEYO y ECHEBURÚA 2010: 408)⁵⁷, lo que podría ser no muy esperanzador, tomando en cuenta las implicaciones del riesgo de error en el área que nos ocupa, mucho mayores que las relacionadas con llevar o no un paraguas a la oficina.

Por su parte, los modelos actuariales o estadísticos se fundan en una perspectiva nomotética (general) y se concentra en los factores estáticos. Aunque tienen la virtud de eliminar el componente subjetivo presente en los modelos clínicos, su mayor debilidad radica, conforme lo señalan MUÑOZ y LÓPEZ, en la posibilidad de «sobreestimar el nivel de científicidad del proceso decisional» (2016: 134), inobservando que los valores promedio de un grupo reflejan en muy escasa medida el posible comportamiento de un individuo concreto, debido precisamente a la heterogeneidad intragrupo, es decir, se trata de una excesiva simplificación de los procesos complejos y multifactoriales que intervienen en la determinación de la conducta de una per-

⁵⁵ Hasta los años ochenta del siglo pasado se utilizó un modelo no estructurado que se centraba en la psicopatología y se apoyaba exclusivamente en la impresión clínica del experto. Sin embargo, la deficiente capacidad predictiva y el alto nivel de subjetividad provocó su desuso en la actualidad (MUÑOZ VICENTE y LÓPEZ-OSSORIO 2016: 133). Cfr. PUEYO y ECHEBURÚA, 2010:404. ESPEC por su parte identifica 5 etapas en la valoración de la peligrosidad-riesgo: a. juicios clínicos no estructurados (hasta 1970); b. investigación empírica (durante la década de los 70 y 80); c. métodos exclusivamente actuariales (a mediados de los 90's); d. método mixto, clínico-actuarial (año 2000); y e. árbol de decisiones (en la actualidad) (2003: 52-64). MARTÍNEZ GARAY reconoce tres métodos para la predicción de la peligrosidad: a. actuarial; b. clínico no estructurado (también llamado por la doctrina alemana como método clínico idiográfico o empírico-individual; y c. clínico estructurado (2014: 15-18). No obstante, para los efectos de este ensayo, se utilizará una clasificación simplificada dual que agrupe las evaluaciones que incorporen mecanismos actuariales, sean estos *puros* o con cierta incidencia del criterio de un experto, prescindiendo del modelo clínico no estructurado, por estar este último ampliamente desacreditado (MARTÍNEZ GARAY 2014: 18).

⁵⁶ Incluyendo el sesgo de representatividad, el de intermediación y disponibilidad, el de anclaje y ajuste, el retrospectivo, el de confirmación y el llamado *miedo social al falso negativo* (MUÑOZ VICENTE y LÓPEZ-OSSORIO 2016: 135). Cfr. NIEVA FENOLL, 2010: 120-128, SUNSTEIN, 2012: 61-63.

⁵⁷ Cfr. con amplia bibliografía sobre el particular, MARTÍNEZ GARAY 2014: 27.

sona⁵⁸. Tal y como apunta MARTÍNEZ GARAY, los modelos actuariales suponen la validez de un razonamiento analógico pese a que, como resulta evidente, no ofrecen una predicción individual (2014: 16)⁵⁹.

Lo anterior significa que el riesgo de un falso negativo o un falso positivo será tanto mayor cuanto mayor sea la heterogeneidad de los sujetos dentro del grupo de referencia. Estos aspectos son coincidentes con los problemas relativos a la utilización de la estadística como *garantía* de los datos, visualizados desde el derecho. Por ejemplo, refiere DEI VECCHI, al analizar el sistema estadounidense y la aplicación del *bail denial* (2015b:152–156), que efectuar una prognosis de la conducta humana, aún disponiendo de información estadística que relacione ciertos datos con la concreción de un peligro procesal, resulta extraordinariamente complejo.

En el mismo sentido, se han señalado tres problemas concretos en cuanto a la información estadística: no se dispone de esa información; no se dispone de un mecanismo para determinar cuál es la clase de referencia (hombres, pobres, con estudios universitarios, solteros, madres, etc.), siendo las posibles combinaciones de factores infinita; y aún disponiendo de información estadística, esta no dice nada sobre el caso concreto a juzgar (FERRER BELTRÁN, 2018b: 149–150)⁶⁰.

Volviendo a la perspectiva eminentemente psicológica, el problema que se encuentra en cualquiera de los dos modelos analizados es que las correspondientes debilidades anotadas provocan un incremento significativo del número de falsos positivos⁶¹, pues aunque el valor predictivo negativo presenta márgenes altos de acierto, el valor predictivo positivo no alcanza el cincuenta por ciento, de acuerdo con los metaanálisis realizados recientemente (MUÑOZ VICENTE y LÓPEZ-OSSORIO 2016: 137–38).

⁵⁸ Un ejemplo interesante de error en la aplicación de instrumentos actuariales lo aportan MUÑOZ y LÓPEZ, quienes refieren el caso de Jeffrey DAHMER, el conocido «Canibal de Milwaukee», en donde, de haberse aplicado un modelo actuarial (el RRASOR), habría sido catalogado como de bajo riesgo de recidiva sexual (2016:135). Sobre este mismo caso también puede verse: ESBEK, 2003: 58.

⁵⁹ La autora hace un interesante análisis de los casos «Baxstrom» y «Dixon», resueltos por el Tribunal Supremo Federal de los Estados Unidos de América, así como sobre otras investigaciones realizadas en Alemania, en donde se constatan tasas altísimas de falsos positivos con la aplicación de mecanismos actuariales, que alcanzan hasta el 97% de falibilidad, lo que evidencia una notoria sobreestimación de la peligrosidad. Sin embargo, también se observa un alto grado de acierto con respecto a la conclusión de que un sujeto es poco peligroso (2014: 19–23). *Vid.* también: MARTÍNEZ GARAY 2016: 17.

⁶⁰ Se trata de la llamada falacia de la división (MARTÍNEZ GARAY 2014: 42). MARTÍNEZ GARAY destaca que el margen de error grupal (intervalo de confianza) presente en los instrumentos actuariales es completamente diverso al correspondiente a la valoración de la predicción individual (intervalo de predicción), siendo este último siempre mayor que el primero, alcanzando valores que pueden ubicarse entre 1 por 100 y 99 por 100 con lo que el instrumento «pierde toda capacidad discriminadora, en la medida en que estos márgenes de error tan amplios se solapan por completo» (2014: 41).

⁶¹ MUÑOZ y LÓPEZ refieren recientes metaanálisis que demostraron una mínima diferencia entre la capacidad predictiva de ambos modelos (2016: 135). Cfr. PUEYO y REDONDO, 2007: 169, quienes afirman que si se ha constatado una mejora en la eficacia predictiva de los métodos clínico-estructurados.

Por otra parte, en múltiples investigaciones realizadas, MARTÍNEZ GARAY ha comprobado que los pronósticos de peligrosidad sustentados en mecanismos actuariales ocultan importantes tasas de falsos positivos (2014, 2016, 2018; MARTÍNEZ GARAY y MONTES SUAY 2018), debido a diversos factores⁶². Algunos de estos son atinentes a las limitaciones propias de los métodos estadísticos y la forma en que se presentan los resultados en sede judicial, dada la altísima complejidad de los instrumentos de predicción de peligrosidad, aun para los expertos en la materia.

Desde esta perspectiva, la autora enfatiza la relevancia de diferenciar con claridad la sensibilidad (capacidad para detectar a las personas que poseen la característica buscada), la especificidad (capacidad para detectar a las personas que no poseen la característica buscada) y el valor predictivo del instrumento (determina cuántos de los sujetos identificados con cierta característica por el instrumento efectivamente la poseen), de forma que, mientras los dos primeros aspectos tienen una orientación retrospectiva, el último tiene una orientación prospectiva (MARTÍNEZ GARAY y MONTES SUAY 2018: 13–16).

Aunque el valor predictivo es el que más interesa en el ámbito jurídico (MARTÍNEZ GARAY 2014: 36), algunas investigaciones ocultan ese dato y presentan valores relativos a la sensibilidad o la especificidad, los que suelen ser diametralmente diversos al primero y, en tal sentido, pueden llevar a error al operador jurídico, en cuanto a la apreciación del grado de acierto del mecanismo⁶³.

Adicionalmente, la llamada curva ROC⁶⁴ es en la actualidad la medida estándar para establecer la capacidad predictiva de los mecanismos actuariales de evaluación de riesgo, a pesar de que la información que deriva de esta (área bajo la curva) es una medida de riesgo relativo que nada dice sobre el riesgo absoluto o sobre el valor predictivo, lo que puede llevar a severos errores en la apreciación del nivel de fiabilidad de los resultados⁶⁵. Además, no se deben inobservar las cuestiones propias de la política criminal, ya que «todo apunta a que cuando se trata de adoptar medidas sobre sujetos eventualmente peligrosos predomina el deseo de minimizar las ‘omi-

⁶² A continuación, se hará un sucinto repaso sobre algunas nociones trascendentales para entender los vicios que pueden presentarse en la aplicación de este tipo de instrumentos actuariales, para cuyo abundamiento se remite a la bibliografía ya mencionada.

⁶³ MARTÍNEZ GARAY también destaca otros elementos como parte de las limitaciones de los métodos estadísticos, por ejemplo, la tasa de prevalencia del fenómeno, ya que, de acuerdo con la fórmula de Bayes, «si la sensibilidad y la especificidad permanecen constantes, cuanto más baja es la prevalencia de un fenómeno en una población, tanto mayor será la tasa de falsos positivos que se obtenga al predecirlo» (2014: 30), los márgenes de error, la traslación de resultados grupales a individuos y la diferencia entre el riesgo absoluto y el riesgo relativo (MARTÍNEZ GARAY 2016: 16–23, 2018: 12–13).

⁶⁴ Acrónimo de *Receiver Operating Characteristics*.

⁶⁵ «puede haber instrumentos con valores altos de área bajo la curva ROC, superiores al 80%, pero cuyo valor predictivo no llegue ni siquiera en el mejor de los casos al 50 por 100 (...) Hay que tener en cuenta que la curva ROC se construye sobre los valores de sensibilidad y de especificidad, de modo que tiene -igual que ellos- una orientación fundamentalmente retrospectiva, y no prospectiva» (MARTÍNEZ GARAY y MONTES SUAY 2018: 25).

siones’, aunque ello sea a costa de asumir un importante número de ‘falsas alarmas’» (MARTÍNEZ GARAY 2014: 32).

Se suman a las limitaciones propias de las metodologías aplicadas por la psicología cognitiva algunas condiciones que podrían llamarse *estructurales*, relacionadas con el sistema de administración de justicia penal y la etapa en que se requeriría la aplicación de los instrumentos analizados. En primer lugar, debe apuntarse que para la realización de este tipo de pericias es imprescindible contar con una enorme cantidad de información, relativa al expediente judicial (incluyendo el record criminal que pudiera tener el sujeto), historial médico, entrevistas, exploración física y psicopatológica, pruebas psicométricas, entre otras (PUJOL ROBINAT, MOHÍNO JUSTES, y GÓMEZ-DURÁN 2014: 21–22), lo que sería virtualmente imposible de obtener dada la premura con que debe ser resuelta una solicitud de medidas cautelares y la saturación de los servicios de psicología forense institucionales.

Otro inconveniente de especial relevancia es que la escala, guía o formulario que utilice el profesional en psicología en la evaluación debe estar validado en el contexto específico en que se vaya a utilizar, lo que puede variar de un país a otro, pero también incluso dentro del mismo territorio, en atención a factores socioculturales y educativos que son extraordinariamente volubles⁶⁶.

Pero uno de los problemas más significativos, sin duda alguna, es el de la selección de los factores de riesgo que se consideren empíricamente válidos, ya que los mismos pueden responder a estereotipos o a la propia selectividad de los sistemas penales⁶⁷, que se dirige a «la raza distinta que tiene el privilegio de poblar los presidios y las cárceles» (FOUCAULT 2003: 262). De acuerdo con estudios realizados por MANDRACCHIA y MORGAN (2011) a partir de la revisión de 595 causas con delincuentes adultos encarcelados, intentaron identificar aquellos con una prognosis criminógena sobre la base de variables como la salud mental, factores demográficos, entre otros, logrando identificar que las personas más jóvenes, con menor educación formal, solteros, afrodescendientes e hispanos, mantenían mayores índices de *criminogenic cognitions*⁶⁸.

Lógicamente, si se parte de este tipo de indicadores, los instrumentos o guías que apliquen los expertos tendrán una cualidad de autopoiesis, que tenderá a perpetuar la desigualdad estructural de nuestras sociedades, con franca lesión de los derechos

⁶⁶ Por ejemplo, las guías SARA y EPV se encuentran validadas en España, ya que la situación cultural europea es disímil a la estadounidense con respecto a indicadores como el uso de armas y la composición familiar, elementos que pueden ser determinantes en la precisión del pronóstico (ECHEBURÚA et al. 2010: 1055).

⁶⁷ El III Informe sobre el Estado de la Justicia en Costa Rica señala que «el aumento en la pena privativa de libertad responde a entornos de exclusión y estigmatización, donde se victimiza a las poblaciones en mayor riesgo social a través de una selectividad estructurada del sistema penal» (2020: 185). Adicionalmente, sobre el llamado *self-fulfilling prophecy* puede verse: MARTÍNEZ GARAY 2014: 5.

⁶⁸ ROMEO CASABONA identifica un aspecto adicional con algunas preguntas que incorporan las guías, mismas que no guardan relación mediata o inmediata con el estilo de vida o el entorno familiar, con proclividad o que de alguna forma pudiese ser considerado como un factor criminógeno (2018: 44).

fundamentales de las poblaciones más vulnerables⁶⁹, de manera que «lo que hace relevante a la estadística no es más que el fin que se proponen quienes la invocan» (DEI VECCHI 2015a: 7).

4.2. La inteligencia artificial

En términos generales, las críticas planteadas en contra del uso de la psicología cognitiva y la estadística para predecir conductas con la fiabilidad suficiente para propiciar una respuesta penal, es trasladable a los mecanismos, en apariencia más refinados, de la inteligencia artificial, en tanto la valoración de riesgos mediante cualquier tipo de *software* se alimenta de variables estadísticas dispuestas por seres humanos⁷⁰.

Ya desde el propio ámbito de la psicología se ha advertido sobre la «tendencia a sobrevalorar desde el contexto jurídico todo aquello que lleva la etiqueta de ciencia» (MUÑOZ VICENTE y LÓPEZ-OSSORIO 2016: 130)⁷¹, lo que conlleva un enorme riesgo para que se defiera la resolución de asuntos trascendentales en programas computarizados que no generan, al menos con el desarrollo tecnológico actual, mayores garantías que la aplicación de una guía clínica estructurada como las ya analizadas, con el único aspecto positivo atinente a la eliminación de una buena parte del componente subjetivo en la aplicación del instrumento⁷².

Como se ha indicado, los sistemas de inteligencia artificial para la predicción del riesgo incorporan modelos estadísticos y matemáticos, clasificando al individuo en ciertos grupos prototípicos a los que se les asigna determinado grado de riesgo, de forma completamente despersonalizada, sin efectuar un pronóstico conductivo psicológico concreto, de manera similar a la evaluación de riesgo que se realiza en otras áreas de la actividad humana, como la industria del seguro y la industria biotecnológica (ROMEIO CASABONA 2018: 43).

⁶⁹ Sobre este aspecto se volverá en el siguiente apartado, al desarrollar los instrumentos de inteligencia artificial.

⁷⁰ «En esencia no es, por consiguiente, un método diferente, sino una herramienta tecnológica nueva que trataría de forma “inteligente” los algoritmos que sirven de base al procedimiento actuarial» (ROMEIO CASABONA 2018: 47).

⁷¹ MARTÍNEZ GARAY advierte sobre el «peligro de que la sofisticación estadística y la consideración como ‘científicas’ de las herramientas de valoración del riesgo favorezca el que se extendiera la ilusión de que su uso convierte en meras decisiones técnicas los seculares problemas valorativos que siempre han estado implicados, y lo siguen estando, en toda decisión que suponga aplicar consecuencias penales a un individuo basadas en un juicio acerca de su comportamiento en el futuro: nos referimos a la ponderación entre el respeto a los derechos fundamentales del reo, por un lado, y las necesidades de prevención por el otro» (2018: 8).

⁷² NIEVA refiere que los criterios estadísticos son aptos para componer una herramienta de inteligencia artificial, pero visualiza idénticos reparos en su utilización, relacionados con la pérdida de detalles del caso concreto y la dificultad que entraña evaluar riesgos a partir de la observación de circunstancias externas, por lo que «a una aplicación de inteligencia artificial habría que pedirle algo más» que la labor que realizan actualmente los seres humanos (2018: 63).

Un caso muy interesante, con respecto al mecanismo denominado COMPAS⁷³, permite ilustrar algunos de los aspectos que se han criticado a lo largo de este ensayo. Se trata de la sentencia del Tribunal Supremo de Wisconsin denominada *State v. Loomis*⁷⁴. En este fallo el alto tribunal evaluó la constitucionalidad del uso de algoritmos procesados informáticamente, a través de un *software* de inteligencia artificial, para sostener una sentencia condenatoria, señalando dicho órgano decisor que no resultó conculcado el debido proceso por el hecho de que no se le brindó a las partes la posibilidad de conocer cómo se realizó realmente el procesamiento de los algoritmos, dado que la empresa *Northpointe Inc.*, dueña del sistema, no estuvo de acuerdo en revelar esa información y el tribunal decidió darle primacía al derecho de propiedad intelectual sobre los derechos fundamentales del encartado (ROMEO CASABONA 2018: 48–50)⁷⁵.

En la apelación del fallo, la defensa argumentó, entre otras cosas, que la utilización de la herramienta colisionaba con el derecho del imputado a obtener una sentencia individualizada (pues, como se ha visto, el *software* se alimenta de indicadores estadísticos) así como la discriminación por género, pues uno de los factores tomados en cuenta era ese rubro; no obstante, la apelación fue rechazada por el Tribunal y, ante la posterior gestión de revisión del fallo, el Tribunal Supremo de los Estados Unidos de América denegó la admisión del recurso, el 26 de junio de 2017.

En consonancia con las críticas esgrimidas contra las guías actuariales, con respecto a COMPAS se ha denunciado que muchos de los factores que lo integran son eminentemente racistas o clasistas (NIEVA FENOLL 2018: 69)⁷⁶, por lo que, de acuerdo con lo sostenido por ROMEO CASABONA, la fiabilidad de estos mecanismos hacen que su utilización en la actualidad sea claramente prematura, máxime cuando se ha podido corroborar que existen contradicciones entre sus resultados y los obtenidos con la utilización de guías aplicadas por un ser humano; esto no excluye que ciertos mecanismos de este tipo se usen como criterios complementarios o confirmatorios, pero nunca como un aspecto definitorio de una fuente de riesgo (2018: 52–55).

⁷³ Perfiles de Gestión de Delincentes Correccionales para Sanciones Alternativas (*Correctional Offender Management Profiling for Alternative Sanctions*).

⁷⁴ Un desarrollo más amplio sobre la sentencia *State v. Loomis*, 881, N.W.2d 749, 7532 (Wis, 2016) se puede ver en: NIEVA FENOLL 2018: 67–75; SLOBOGIN 2019: 114–116; ROMEO CASABONA 2018: 48–55; MARTÍNEZ GARAY 2018.

⁷⁵ «*But rather than demanding such information so that fit, validity and fairness could be analyzed, the court lamely concluded that trial courts could continue to use the COMPAS in connection with sentencing as long as judges are cognizant of these limitations and do not make the risk score so produced determinative of whether the offender is incarcerated or receives an enhanced sentence*» (SLOBOGIN 2019: 115).

⁷⁶ Para mayor claridad, con respecto a este argumento, puede consultarse la *Practitioner's Guide to COMPAS Core* en: http://www.northpointeinc.com/downloads/compas/Practitioners-Guide-COMPAS-Core-_031915.pdf Visitada el 18 de junio de 2020. *Vid.* también: LARSON, MATTU, KIRCHNER, y ANGWIN, 2016, quienes lograron confirmar, a partir del estudio de más de 10000 casos, que «*that black defendants were far more likely than white defendants to be incorrectly judged to be at a higher risk of recidivism, while white defendants were more likely than black defendants to be incorrectly flagged as low risk*».

Con todo, siguiendo a MARTÍNEZ GARAY, en los Estados Unidos la implementación de la *evidence based sentencing* continúa extendiéndose y, a pesar de que en el caso *State v. Loomis* se estableció que este tipo de mecanismos no pueden ser empleados para decidir el encarcelamiento de una persona o para determinar la severidad de la sentencia, es decir, no pueden constituir las valoraciones de riesgo el elemento determinante en la resolución⁷⁷, es altamente probable que este tipo de *software* provoque un sesgo de anclaje en el decisor, en tanto la privación de libertad «puede hacerse con la conciencia un poco más tranquila si los jueces pueden confiar en que su intuición está corroborada por una valoración de riesgo científicamente fundada» (MARTÍNEZ GARAY 2018: 496), pese a la escasa fiabilidad que pueda tener el instrumento, lo que desaconsejaría su implementación.

4.3. ¿Cuál es el rendimiento de la estadística?

A lo largo de los capítulos precedentes se han expuesto algunas advertencias sobre la utilización en sede judicial de mecanismos predictivos conductuales. El lugar común que se ha observado es el uso de ciertas guías o formularios, que pueden aplicarse de manera *pura*, es decir, con escasa participación de seres humanos, o bien por medio de expertos en psicología cognitiva o a través de programas de inteligencia artificial. Como ocurre muchas veces en el ámbito jurídico, parece que se ha confiado de más en estos instrumentos, habida cuenta de la escasa fiabilidad que tienen sus resultados, «[n]o se olvide que al estar ocupándonos del comportamiento humano futuro, siempre deberá situarse, por su propia naturaleza, en el ámbito de lo conjetural, aunque pueda serlo en un grado mayor o menor» (ROMEO CASABONA 2018: 46).

Sin embargo, estimo que no es posible desestimar sin más el andamiaje epistémico que puede provenir de la estadística, con miras en dar sustento a las inferencias que se construyan para tener por probado un enunciado relativo a hechos prospectivos. Por otra parte, uno de los grandes problemas de la cultura jurídica latinoamericana es la inexistencia de registros empíricos sobre el funcionamiento de las instituciones, no solamente para poder abordar con solvencia reformas o mejoras que optimicen los resultados que se requieren, sino también para modular ciertas prácticas incorrectas o disfuncionalidades de estas. El caso de la medida cautelar de prisión preventiva sin duda es uno de los más llamativos, si se toman en cuenta los elevadísimos índices de hacinamiento carcelario presentes en toda la región, de cuyo total una proporción significativa corresponde a presos sin condena.

En Costa Rica desde hace algunos años se han venido incorporando en los *informes sobre el estado de la justicia* cada vez más datos relevantes que permiten tener una

⁷⁷ En realidad, la sentencia es confusa, pues establece que el *presentence investigation report* debe advertir sobre las limitaciones señaladas e informar otros aspectos sobre el funcionamiento del *software*, pese a lo cual avala la utilización de COMPAS en el caso concreto.

imagen más clara de la forma en que opera la administración de justicia en la realidad. Pero este aspecto positivo desgraciadamente no ha tenido réditos a nivel práctico, con lo que esta valiosa información es escasamente conocida y se ha convertido en *letra muerta*. A continuación, se presentarán algunos datos estadísticos interesantes, provenientes de los estudios mencionados⁷⁸, así como de una investigación privada realizada para los fines de este trabajo, que, según creo, también pudieran ser útiles para la determinación de la procedencia de una medida cautelar, siempre que su uso sea prudente y no se conviertan en parte del frecuente *fetichismo hacia las cifras*, que en ocasiones permea nuestra actividad⁷⁹.

Uno de los aspectos más preocupantes es que el índice de Estado de derecho⁸⁰ muestra que, en 2018, el sistema penitenciario fue el peor evaluado con relación a todos los indicadores asociados a la justicia penal en Costa Rica, con una tasa de prisionalización mucho más alta que la de países con situaciones de mayor violencia y desigualdad social. Los niveles de sobrepoblación carcelaria están por encima del 40 por 100 (más del doble de lo establecido por los estándares internacionales), lo que significa que el hacinamiento es crítico.

De acuerdo con los datos recabados por el Ministerio de Justicia y Paz, en junio de 2019 se llegó al nivel más alto de hacinamiento penitenciario de los últimos dos años, existiendo establecimientos carcelarios con niveles de sobrepoblación superiores al 100 por 100⁸¹, además de registrarse un importante aumento del grupo de presos sin condena, que pasó de un 20,3% en 2015, a un 29,9 por 100 en 2019, con respecto al total de la población penitenciaria costarricense.

Si bien entre 2010 y 2015 disminuyeron las personas con prisión preventiva por delitos contra la propiedad, se registró un aumento de los indiciados por infracciones a la Ley de Psicotrópicos, de forma que ambas categorías delictivas representaron el 65.7 por 100 del total de presos sin condena entre 2005 y 2015, con una característica muy llamativa: los delitos en los que es más frecuente la prisión preventiva registran menos condenas⁸².

⁷⁸ La información se ha tomado de: II Informe Estado de la Justicia 2017, III Informe Estado de la Justicia 2020. Puede accederse a esta información mediante la siguiente dirección web: <https://estadonacion.or.cr/>

⁷⁹ TARUFFO habla del síndrome de «lo que no se puede contabilizar no existe», de donde deriva la errónea idea de que lo que es susceptible de traducirse en cifras es más exacto y más racional que lo que no (2010: 248). En sentido similar, MARTÍNEZ GARAY afirma que «[l]a creencia en que algo puede ser medido con precisión favorece la tendencia a considerarlo un fundamento admisible para una consecuencia jurídica» (MARTÍNEZ GARAY y MONTES SUAY 2018: 37).

⁸⁰ De acuerdo con el *World Justice Project*.

⁸¹ El II Informe Estado de la Justicia (2017) advirtió sobre tres posibles causas del hacinamiento penitenciario: el aumento en las condenas, la creación de los tribunales de flagrancia y el uso excesivo de la prisión preventiva.

⁸² Particularmente alta es la proporción de indiciados por delitos contra la vida e infracciones a la Ley de Psicotrópicos, si se compara con las sentencias condenatorias por esas mismas causas. Por ejemplo, en las infracciones a la Ley de Psicotrópicos la prisión preventiva parece ser casi automática,

Otro de los datos realmente interesantes es que la falta de arraigo (92.5 por 100) y la expectativa de pena muy alta (57.9 por 100) son las dos causales más usuales para disponer la detención cautelar⁸³, además de la magnitud del daño (45.3 por 100), la continuidad delictiva (43.4 por 100) y la «influencia sobre peritos y testigos» (34.6 por 100) o sobre la víctima o denunciante (34 por 100).

En cuanto al perfil socioeconómico de la población privada de libertad en Costa Rica, mayoritariamente se trata de personas provenientes de estratos socioeconómicos bajos, generalmente hombres, jóvenes, costarricenses (87 por 100); el 51 por 100 se encontraba en condición de soltería, divorcio y viudez, mientras que el 49 por 100 en relaciones de unión, sea matrimonio o unión libre. También es notorio un predominio de niveles bajos de escolaridad⁸⁴.

Por otra parte, en un estudio efectuado en la zona sur del país⁸⁵, se ha podido obtener información también interesante. Por ejemplo, del total de casos ingresados al Ministerio Público, en un 23 por 100 se ordenó la libertad sin requerimiento jurisdiccional, en un 27 por 100 se requirió la aplicación de medidas cautelares no privativas de libertad, mientras que un 30 por 100 corresponde con solicitudes de prisión preventiva, una cifra que pudiera estimarse razonable, tomando en cuenta las cifras país vistas anteriormente⁸⁶. De la cifra total, el 41.7 por 100 de los casos corresponden a cuestiones relacionadas con violencia sexual o violencia de género, el 6.8 por 100 a narcotráfico y el 18.4 por 100 específicamente a gestiones de prisión preventiva por casos de violencia de género, llamando la atención que en la totalidad de los casos por tráfico de drogas el Ministerio Público solicitó prisión preventiva, siendo rechazado el pedido en un 42.8 por 100, una cifra alta si se compara con los datos generales en el país.

En cuanto al abordaje de las gestiones fiscales, el Juzgado Penal dictó prisión preventiva en un 21.3 por 100 de los casos, otro tipo de medida cautelar en el 31 por

pero la cantidad de condenas es ridículamente baja, representando un 13 por 100 del total de condenas entre 2010 y 2015. Este dato concretamente debería de utilizarse a la hora de considerar probado el *fumus commissi delicti*, ya que este riesgo no es otra cosa que un pronóstico sobre la futura condena del encartado.

⁸³ Concretamente estos datos fueron obtenidos a partir de un estudio aleatorio realizado a 159 expedientes en que se dictó prisión preventiva en 2016, en la meseta central del país.

⁸⁴ En los hombres, el 58 por 100 tan solo cuenta con primaria, completa o incompleta, un 28 por 100 con secundaria incompleta y un 5 por 100 de personas analfabetas, más del doble del promedio nacional. En las mujeres, el 58 por 100 tiene primaria completa o incompleta, un 28 por 100 secundaria incompleta y un 4 por 100 son analfabetas.

⁸⁵ Específicamente en el II Circuito Judicial de la Zona Sur. Se trata de una adscripción territorial con predominio de delincuencia ligada a la narcoactividad, por ser una zona fronteriza y costera, así como delitos relacionados con violencia de género. La muestra se ha tomado del total de casos con personas detenidas en la Fiscalía durante el año 2019.

⁸⁶ El 19.4 por 100 de las causas ingresadas fueron tramitadas con incompetencia a otra localidad, lo que puede significar un sesgo importante, ya que en estas causas perfectamente se pudo solicitar prisión preventiva u otra medida diversa.

100 de las gestiones y denegó la solicitud y ordenó la libertad en un 4.8 por 100 de los asuntos. Otro dato interesante es que de la cantidad total de personas en condición de rebeldía (no ubicables) y con orden de captura, ninguno corresponde a casos de narcotráfico, el 18.7 por 100 a delitos sexuales, el 34.3 por 100 a delitos contra la propiedad, solamente el 0.03 por 100 a delitos relacionados con violencia de género, mientras que el 59.3 por 100 corresponde a delitos con penas inferiores a los tres años de prisión, lo que permitiría al investigado obtener un beneficio de ejecución condicional de la pena y, eventualmente, no pasar un solo día en prisión.

¿Qué podemos concluir a partir de los datos que se han expuesto? En primer lugar, creo que se ha demostrado la relevancia que tiene tomar en cuenta los aspectos socioculturales, el tipo de variables, las características de la muestra, entre una pluralidad de factores que potencialmente pueden hacer que se alcancen resultados completamente disímiles, aún en una extensión territorial tan pequeña como la de Costa Rica, lo que se refleja con facilidad a partir de las divergencias de los resultados a nivel país, los que se obtuvieron a partir de una muestra aleatoria en el área metropolitana y los recabados en una zona rural. Esto implica, como se ha referido, que el carácter multifactorial y voluble de la conducta humana, además de cualquier sesgo en quien defina las variables o ejecute el muestreo estadístico, puede dar al traste con la potencialidad del instrumento, en términos de fiabilidad.

No obstante, la existencia de mecanismos refinados y validados desde la perspectiva de la psicología cognitiva, con todas las limitaciones que pudiesen tener, permiten generar algunos datos relevantes susceptibles de ser usados para sostener una inferencia en cuanto a hechos prospectivos⁸⁷. Debe recordarse que es preciso descartar como garantía de la inferencia en las medidas cautelares, principalmente cuando se trate de prisión preventiva o monitoreo electrónico, las generalizaciones de sentido común⁸⁸ y las generalizaciones espurias, habida cuenta del alto estándar de prueba exigible en estos actos.

Asimismo, de acuerdo con ROMEO CASABONA, «no debería descartarse que con un desarrollo más perfeccionado y sofisticado de sistemas de inteligencia artificial éstos fueran capaces de predecir de forma objetiva -automatizada- el comportamiento criminal de un delincuente a la vista de su personalidad» (2018: 47), pero las expectativas a corto y medio plazo aún no son promisorias⁸⁹.

⁸⁷ Piénsese que un muestreo estadístico más amplio y adecuadamente validado, que recoja datos similares a los expuestos en cuanto a la prevalencia de fugas efectivas en delitos menores, descartaría el respaldo de la máxima de experiencia que dice que «la alta penalidad puede provocar que el investigado decida fugarse».

⁸⁸ Aunque este tipo de máximas de experiencia pueden servir como *garantía* en medidas cautelares de menos impacto, como el impedimento de salida del país o la restricción de ingreso a ciertas áreas, en donde el estándar de prueba es menor.

⁸⁹ Cfr. MARTÍNEZ GARAY 2014: 52.

Finalmente, algunos resultados provenientes también de la psicología experimental y relacionados con el análisis económico del derecho y que generan algunas predicciones sobre el comportamiento general de los seres humanos podrían tener una utilidad limitada, por ejemplo, con respecto a la aversión a la pérdida, la aversión al extremismo y la *apariencia de justicia*⁹⁰, de donde podría derivarse, en cuanto a lo que aquí interesa, que una persona valore negativamente el mantenerse en la condición de prófugo, máxime cuando nos encontramos en sociedades *omnidisciplinarias*⁹¹, lo que llevaría a la mayoría a no fugarse; o bien, el hecho de que la instrucción de un proceso penal contra un individuo implica, de forma automática, que si desaparece prueba inculpatoria el primer sospechoso sea el potencial beneficiado con el acto, lo que podría influir sobre la real ejecución de actos de obstaculización por parte del encartado.

5 CONCLUSIONES

A lo largo de este trabajo se han podido delinear algunos de los aspectos que repercuten en la toma de decisiones relativas a la procedencia o no de medidas cautelares de carácter personal en un proceso penal. Se ha enfatizado que existen al menos tres tipos de variables que afectan en la práctica la adecuada fundamentación de la premisa fáctica de este tipo de actos judiciales: los factores procesales positivos, los factores de resiliencia interna del juzgador y los factores estructurales de la pretensión.

En cuanto a los factores estructurales de la pretensión, se indicó que, mientras el *fumus commissi delicti* es una hipótesis probabilística de primer orden (sobre la posibilidad de que recaiga sentencia condenatoria), el *periculum libertatis* se configura, en la mayoría de las legislaciones, incluyendo la costarricense, como una hipótesis probabilística de segundo grado (debe corroborarse un hecho base y su secuela).

De acuerdo con la postura asumida, para mantener la finalidad cautelar de la medida y permitir un verdadero razonamiento probatorio es preciso enlazar la noción de peligro con la ejecución de acciones concretas, lo que significa entender el riesgo procesal como una hipótesis probabilística de primer orden, semejante al *fumus commissi delicti*.

Con relación a la inferencia probatoria en hechos prospectivos, partiendo del esquema argumentativo de Stephen TOULMIN, se ha destacado la trascendencia de la *garantía* y su *respaldo*. No quedan claras las razones por las que, generalmente, en la disposición de medidas cautelares los decisores no explicitan esa parte del razonamiento, con lo que convierten la inferencia en un razonamiento entimemático, que sugiere además la posible automatización de la medida, es decir, que en la práctica

⁹⁰ Las personas quieren actuar con justicia y ser vistas actuando con justicia (SUNSTEIN 2012: 60).

⁹¹ Parafraseando la referencia de FOUCAULT con respecto a la prisión (2003: 238).

se esté dando un ejercicio calificadorio o definitorio y no un razonamiento probatorio en sentido estricto.

Se ha excluido la posibilidad de utilizar como *garantía* presunciones *iuris* o definiciones, por lo que es preciso echar mano a máximas de experiencia. Estas máximas pueden corresponderse con leyes científicas de carácter universal, generalizaciones no universales con alto grado de probabilidad, generalizaciones de sentido común y generalizaciones espurias.

Para la aplicación de medidas cautelares con un alto impacto en la esfera de derechos y libertades de las personas (como la prisión preventiva o el monitoreo mediante mecanismos electrónicos de seguimiento⁹²), que aparejan un alto umbral de prueba, requerirán que la *garantía* de la inferencia sea como mínimo una generalización no universal con alto grado de probabilidad, mientras que otro tipo de medidas menos invasivas podrían conformarse con generalizaciones de sentido común, habida cuenta del estándar probatorio más bajo que mantienen, pero excluyendo en cualquier caso las generalizaciones espurias.

Dentro de las herramientas que se utilizan para la predicción de conductas futuras, la psicología cognitiva, la estadística y la inteligencia artificial tienen como rasgo común que, en mayor o menor medida, parten de la base de estudios actuariales de grupo, que poco pueden decir sobre el comportamiento de un individuo concreto. Adicionalmente, estos modelos son permeables a sesgos o tienen el riesgo de ser sobrevaluados por parte del juzgador, cuando, al menos para el momento actual, la fiabilidad de sus resultados es muy limitada.

No obstante, se estima que estos mecanismos pueden tener mayor relevancia para la determinación de medidas cautelares de bajo impacto, siendo en principio insuficientes para la determinación de la prisión preventiva o el monitoreo electrónico. De cualquier manera, como propuesta de *lege ferenda* se estima imprescindible que legislativamente se incorporen en los diseños procesales, estándares de prueba claros y diferenciados para las medidas cautelares rigurosas y para las medidas de baja injerencia⁹³.

⁹² Debe hacerse hincapié en que el monitoreo electrónico representa una grave intrusión en la libertad e intimidad del usuario, de forma que, en la legislación procesal costarricense incluso se ha equiparado su utilización con la prisión preventiva, estableciendo el numeral 244 inciso j del CPP, como medida cautelar «La imposición de la medida de localización permanente con mecanismo electrónico. Para tal efecto, un día bajo localización permanente con dispositivo electrónico equivale a un día de prisión preventiva». Sobre este tema puede consultarse: MORA 2017.

⁹³ En la práctica judicial es usual que, ante el requerimiento de prisión preventiva por parte del Ministerio Público, la defensa técnica del encartado *solicite* se impongan medidas cautelares no limitativas de la libertad ambulatoria. El requerimiento de medidas coactivas por parte de la propia defensa, aunque posible con el actual diseño procesal costarricense, es, a mi juicio, una disfuncionalidad del sistema, que se potencia por la inexistencia de estándares diferenciados para la detención cautelar y para otro tipo de medidas. Si se definieran tales estándares, el Ministerio Público se vería obligado a gestionar, de

Ahora bien, tanto los mecanismos de predicción analizados como los patrones estadísticos que reflejan el funcionamiento real de los sistemas de justicia, aunque poseen un alcance limitado como aval epistémico a nivel inferencial, si mantienen enorme trascendencia en cuanto al rendimiento de las instituciones y pueden ayudar a detectar disfuncionalidades -como se ha visto en el caso del uso desmedido de la prisión preventiva- y diseñar reformas normativas que mejoren la administración de justicia⁹⁴.

Finalmente, considerando que el *periculum libertatis* debe entenderse como una hipótesis probabilística de primer orden, la capacidad predictiva de la conducta por medio de la psicología cognitiva aumenta sensiblemente su fiabilidad. Debe recordarse que todos los estudios de predictibilidad parten de una base retrospectiva, entendiéndose desde la psicología cognitiva que «[e]l futuro es un repetición del pasado, solo interesa la probabilidad que existe de que un hecho suceda en el futuro» (GARRIDO y GONZÁLEZ 2018: 11), por lo que la validez y solidez de una inferencia bajo esa estructura resulta diametralmente más fiable cuando ya se ha dado la concreción de algún acto de fuga u obstaculización de prueba.

6. BIBLIOGRAFÍA

- ATIENZA, M., 2005: *Las razones del derecho. Teorías de la argumentación jurídica*, Universidad Nacional Autónoma de México.
- CERDA LORCA, J. y VALDIVIA CABRERA, G., 2020: *Conceptos básicos de metodologías de investigación clínica y epidemiológica*, Santiago: Ediciones Universidad Católica de Chile.
- CHINCHILLA CALDERÓN, R. y GARCÍA AGUILAR R., 2003: *Disfuncionalidades en la aplicación de la prisión preventiva*, San José: Investigaciones Jurídicas S.A.
- CHRISTIE, N., 2004: *Una sensata cantidad de delito*, Buenos Aires: Editores del Puerto.
- DEI VECCHI, D., 2015a: «La peligrosidad judicial: Sobre la justificación de la premisa fáctica de la prisión preventiva en la decisión judicial y algunas críticas no tan frecuentes», *La Ley Penal* (115): 1–17.
- 2015b. *Peligrosidad judicial y encarcelamiento cautelar. una aproximación a la prisión preventiva desde la óptica de la justificación de decisiones judiciales*, Buenos Aires: Ad-Hoc.
- ECHEBURÚA, E., AMOR, P., LOINAZ, I., y DE CORRAL, P., 2010: «Escala de predicción del riesgo de violencia grave contra la pareja», *Psicothema* 22(4):1054–60.
- ESBEC, E., 2003: «Valoración de la peligrosidad criminal (riesgo-violencia) en Psicología forense. Aproximación conceptual e histórica». *Psicopatología Clínica, Legal y Forense* 3(2):45–64.
- FERNÁNDEZ LÓPEZ, M., 2004: «Presunción de inocencia y carga de la prueba en el proceso penal», Universidad de Alicante.
- FERNÁNDEZ SEPULVEDA, A., 1985: *Derecho judicial y justicia constitucional: Una aproximación al tema*, Madrid: Secretaría General Técnica. Ministerio de Justicia.

forma subsidiaria al requerimiento de prisión, otro tipo de cautelas, dado que, de rechazarse la solicitud principal, lo que procedería es la libertad del imputado, sin condicionamiento alguno.

⁹⁴ «La buena capacidad para estimar el riesgo relativo de violencia o de reincidencia que tienen los instrumentos estructurados de valoración del riesgo puede constituir una información útil para la toma de determinadas decisiones en el ámbito judicial, especialmente aquellas que tengan que ver con la asignación de recursos escasos a un grupo de personas» (MARTÍNEZ GARAY 2016: 21).

- FERRER BELTRÁN, J., 2019: *Prueba y racionalidad de las decisiones judiciales*, Hidalgo: CEJL.
- FERRUA, P., 2018: «La prova nel processo penale», *Revista Brasileira de Direito Processual Penal*, 81–128.
- FOUCAULT, M., 2003: *Vigilar y castigar. Nacimiento de la prisión*. 32a, Buenos Aires: Siglo Veintiuno Editores.
- FREISLER, R., 2017: *Derecho penal de voluntad*, Ciudad Autónoma de Buenos Aires: Ediar.
- GARRIDO, V., y GONZÁLEZ, A., 2018: «La predicción de la delincuencia», *Universitat Oberta de Catalunya*.
- GONZÁLEZ LAGIER, D., 2013: *Quaestio Facti. Ensayos sobre prueba, causalidad y acción*, Primera Ed. México D.F.: Editorial Fontamara.
- HAACK, S., 2014: *Evidence matters: Science, proof, and truth in the law*, New York: Cambridge University Press.
- HUNTER AMPUERO, I., 2015: «Las dificultades probatorias en el proceso civil. Tratamiento doctrinal y jurisprudencial, críticas y una propuesta», *Revista de Derecho Universidad Católica Del Norte*, 209–57.
- INFORME, 2017: «Segundo Informe Estado de la Justicia», en *Programa estado de la nación*.
- INFORME, 2020: «Tercer Informe Estado de la Justicia», en *Programa estado de la nación*.
- LARSON, J., MATTU, S., KIRCHNER, L., y ANGWIN, J., 2016: «How we analyzed the COMPAS recidivism algorithm», *ProPublica*.
- LLOBET, J., 2017: *Proceso penal comentado (Código procesal penal comentado)*, San José. Editorial Jurídica Continental.
- 2018: *La Corte Interamericana de Derechos Humanos y las garantías penales*, San José: Editorial Jurídica Continental.
- MAIER, J., 1996: *Derecho procesal penal. Tomo I. Fundamentos*. 2a, Buenos Aires: Editores del Puerto.
- MANDRACCHIA, J., y MORGAN, R., 2011: ««Predicting offenders” Criminogenic cognitions with status variables», *Criminal Justice and Behavior*.
- MARTÍNEZ GARAY, L., 2014: «La incertidumbre de los pronósticos de peligrosidad: consecuencias para la dogmática de las medidas de seguridad», *InDret* 0(2):1–77.
- 2016: «Errores conceptuales en la estimación de riesgo de reincidencia», *Revista Española de Investigación Criminológica* 14(3):1–31.
- 2018: «Peligrosidad, algoritmos y due process: el caso *State v Loomis*», *Revista de Derecho Penal y Criminología* 20: 485–502.
- MARTÍNEZ GARAY, L., y MONTES SUAY, F., 2018: «El uso de valoraciones del riesgo de violencia en derecho penal: algunas cautelas necesarias», *InDret* (2):1–46.
- MÉNDEZ, C., 2015: *Pensar la psicología*, Madrid: Siglo XXI.
- MIRANDA ESTRAMPES, M., 1997: *La mínima actividad probatoria en el proceso penal*, Barcelona: J.M. Bosch.
- MORA SÁNCHEZ, J., 2015: *Prisión preventiva y control de convencionalidad*, San José: Editorial Juritexto.
- 2017: *Mecanismos electrónicos de seguimiento ¿entre el panóptico y la prisión electrónica?*, San José: Investigaciones Jurídicas S.A.
- 2018: «La prueba de hechos futuros», *Revista Firma*.
- 2019: «Dificultades probatorias en la prisión preventiva», pp. 205–48, en *La prueba. Razonamiento probatorio y justificación de las decisiones judiciales y administrativas*, San José: Investigaciones Jurídicas S.A.
- MORA SÁNCHEZ, J., et. al. 2019: *La prueba. Razonamiento probatorio y justificación de las decisiones judiciales y administrativas*, San José: Investigaciones Jurídicas S.A.
- MUÑOZ VICENTE, J.M., y LÓPEZ-OSSORIO, J.J., 2016: «Valoración psicológica del riesgo de violencia: Alcance y limitaciones para su uso en el contexto forense», *Anuario de Psicología Jurídica* 26: 130–40.
- NAVARRO, J.J., y PASTOR E., 2017: «Factores dinámicos en el comportamiento de delinquentes juveniles con perfil de ajuste social. Un estudio de reincidencia», *Psychosocial Intervention* 26:19–27.
- NIÉVA FENOLL, J., 2010: *La valoración de la prueba*, Madrid: Marcial Pons.
- 2018: *Inteligencia artificial y proceso judicial*, Madrid: Marcial Pons.

- NIEVA FENOLL, J., FERRER BELTRÁN, J., y GIANNINI L., 2019: *Contra la carga de la prueba*, Madrid: Marcial Pons.
- POZO, J.I., 2006: *Teorías cognitivas del aprendizaje*, Madrid: Ediciones Morata.
- PUEYO, A.A., y ECHEBURÚA E., 2010: «Valoración del riesgo de violencia: Instrumentos disponibles e indicaciones de aplicación», *Psicothema* 22(3):403–9.
- PUEYO, A.A., y REDONDO S., 2007: «Predicción de la violencia: Entre la peligrosidad y la valoración del riesgo de violencia», *Papeles Del Psicólogo* 28(3):157–73.
- PUJOL ROBINAT, A., MOHÍNO JUSTES, S., y GÓMEZ-DURÁN, E., 2014: «Valoración forense del riesgo de violencia», *Medicina Clínica* 142(2):16–23.
- ROMEO CASABONA, C., 2018: «Riesgo, procedimientos actuariales basados en inteligencia artificial y medidas de seguridad», *R.E.D.S.* (13):39–55.
- SCHAUER, F., 2013: *Pensar como un abogado*, Madrid: Marcial Pons.
- SILVA, D., MERCURIO, E., y LÓPEZ F., 2008: *Imputabilidad penal y neurociencias*, Buenos Aires: Ad-Hoc.
- SLOBOGIN, C., 2019: «A defence of modern risk-based sentencing», en *Predictive sentencing. Normative and empirical perspectives*, Oxford: Hart.
- SUNSTEIN, C., 2012: «Análisis conductual del derecho», *THÉMIS-Revista de Derecho* 0(62): 53–64.
- TARUFFO, M., 2010: *Simplemente la verdad*, Madrid: Marcial Pons.
- TOULMIN, S., 2007: *Los usos de la argumentación*, Barcelona: Ediciones Península.
- VALENZUELA, J., 2018: «Hacia un estándar de prueba cautelara en materia penal: Algunos apuntes para el caso de la prisión preventiva», *Política Criminal. Vol. 13, N 26*, 836–57.
- VEGA MONGE, A., 2015: *Populismo punitivo y medios de comunicación*, San José: Editorial Jurídica Continental.

EVIDENCE ASSESSMENT AND STANDARDS OF PROOF: A MESSY ISSUE*

Giovanni Tuzet

Bocconi University

giovanni.tuzet@unibocconi.it

ABSTRACT: The article distinguishes evidence assessment criteria from standards of proof and addresses three main questions. First: why do some scholars and decision-makers take assessment criteria as standards of proof and vice versa? The answer comes from the fact that some legal systems are more concerned with assessment criteria and others with standards; therefore, jurists educated in different contexts tend to emphasize what they are more familiar with, and to assimilate to it what they are less familiar with. Second: why do systems differ in those respects? Here the answer stems from the historical, institutional and procedural differences that explain why some systems are more concerned with assessment criteria and others with standards of proof. And third, assuming that both criteria and standards are necessary to legal decision-making about facts: how can a system work if it neglects one of these things? Here the article argues that there is a *functional connection* between criteria and standards. The functional connection account is distinguished from a functional equivalence account, and some systems and jurisdictions are referred to in greater detail to support the functional connection claim.

* Earlier versions of this work have been presented in Trento (“GTR 15. Trento Days on Rhetoric”, University of Trento, Italy, June 2015), Washington D.C. (“XXVII World Congress of the International Association for the Philosophy of Law and Social Philosophy”, Georgetown Law Center, July 2015), London (“ICAIL 2017 Workshop”, King’s College, June 2017) and Santiago de Chile (“Seminario 3”, Universidad Alberto Hurtado, June 2017). I thank all the participants in those events who commented on my work; in particular, I wish to thank Daniela Accatino, Ron Allen, Damiano Canale, Flavia Carbonell, Rodrigo Coloma, Christian Dahlman, Marcello di Bello, Paolo Heritier, Mike Pardo and Michele Taruffo. An abridged and earlier version has appeared as Tuzet 2020. Last but not least, I thank the reviewers of this journal for their comments on the submitted draft.

KEYWORDS: assessment of evidence, fact-finding, functional connection, legal proof, standards of proof

SUMMARY: 1. INTRODUCTION.— 2. ANALYTICAL STARTING POINTS.— 3. COMPARATIVE CONSIDERATIONS: 3.1. Common Law Countries. 3.2. Civil Law Countries.— 4. THE FUNCTIONAL CONNECTION ACCOUNT.— 5. CONCLUSION.— 6. BIBLIOGRAPHY

RECOMMENDED CITATION: TUZET G., 2020: «Evidence Assessment And Standards Of Proof: a Messy Issue», in *Quaestio facti*, 2: 87-113. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22480

1. INTRODUCTION

There is a persistent confusion in legal literature. It concerns evidence assessment criteria (or evidence evaluation methods) and standards of proof, which are frequently mixed up and treated as if they were the same thing.

The present work is motivated by the wish to clarify this issue providing an adequate conceptual distinction between assessment criteria and standards of proof. In a nutshell, evidence must be assessed in order to check whether it satisfies a relevant standard of proof; the assessment is operated with some criteria, and both criteria and standards are necessary to fact-finding. This is not as trivial as it may appear because, as I will show, some legal scholars, practitioners and decision-makers seem to miss the difference between criteria and standards.

In addition to this conceptual point, the article addresses three main questions. First: why do some scholars and decision-makers persistently take assessment criteria as standards of proof and vice versa? The answer to this question, simply put, comes from the fact that legal systems are different because, inter alia, some are more concerned with assessment criteria and others with standards; therefore, jurists educated in different contexts tend to emphasize what they are more familiar with, and to assimilate to it what they are less familiar with. Second question: why do systems differ in those respects? Here the answer stems from the development of different perspectives over time: there are historical, institutional and procedural differences that explain why some systems are more concerned with assessment criteria and others with standards of proof. And third, assuming that both criteria and standards are necessary to legal decision-making about facts: how can a system work if it neglects one of these things? The answer to this is the most theoretically ambitious part of the article, since I argue that there is a *functional connection* between criteria and standards; the functional connection account is distinguished from a functional equivalence account, and some systems and jurisdictions are referred to in greater detail to support the functional connection claim.

This article relies on, and is indebted to, a rich literature on evidence assessment and standards of proof. Indeed, there has been a lot of work, over the last years, on

legal standards of proof. Philosophers, legal argumentation theorists, evidence scholars, civil and criminal procedure scholars, among others, have extensively dealt with this issue. Some works make an analytical effort to clarify the idea of a probative standard; others carry out a descriptive study of how standards actually operate; others advance normative claims in the hope of suggesting better standards, or at least better-defined ones; and some do more than one of these things (by defining the issues, describing the existing standards and making normative proposals)¹.

That complex work on standards has been mainly, but not exclusively, carried out in the common law world. In the civil law world, there is less attention to that issue, but a greater concern for the criteria of evidence assessment. A very old Continental debate addresses the question whether it is better to have evidence assessed according to predetermined and fixed criteria (the idea of a «legal proof» system), or according to more relaxed ones accommodated by fact-finders themselves (the idea of a «free assessment» of evidence)². A way of stressing this difference consists in saying that fixed criteria are objective and relaxed ones are subjective, in that the former constrain the fact-finders in a way that the latter do not. However, it is also possible to understand the latter as objective: their objectivity would not consist in the fact that they are legally mandated in virtue of some authoritative source, but in the fact that they put some epistemological constraints on legal decision-making³.

Now, looking at these issues and trying to conceptualize them, it becomes clear that *an evidence assessment criterion is one thing, and a proof standard is quite another*. And both are necessary to legal fact-finding, because evidence assessment is not sufficient to make a decision and because, in principle, a proof standard can be met with different criteria. Consider the «beyond a reasonable doubt» standard: in principle it can be met both with fixed and relaxed criteria, provided that in the first scenario the competent legal authority determines *ex ante* the fixed value of the relevant types of evidentiary items, whereas in the second scenario fact-finders make appeal to non-fixed criteria (being epistemic, practical, or other) to assess the evidence at disposal and judge whether it satisfies the relevant standard.

Still, as previously mentioned, both in the common law and in the civil law world there is some tendency to overlap those issues and mix them up. For instance, the French *intime conviction* is often taken as a subjective criterion of assessment, while others take it as a subjective standard of proof (more on this below). A dispute that occurred some years ago in a top journal still maintains its significance in this respect⁴.

¹ See, among others, REDMAYNE, 1999, LILLQUIST, 2002, PICINALI, 2013. See also VÁZQUEZ, 2013.

² On the criminal side of the debate see FERRAJOLI, 1989: 112 ff. See TARUFFO, 1992: 361 ff on civil aspects too.

³ See TARUFFO, 1992: 370 ff and FERRER 2007: 147 ff.

⁴ See CLERMONT and SHERWIN, 2002 and TARUFFO, 2003a.

The structure of the present article is as follows: § 2 makes some analytical starting points to clear the discussion ground; § 3 performs some comparative work considering different legal systems and the ways in which they deal with the issues discussed here; § 4 tries to convey a sharper analysis and argues that criteria and standards have a functional connection; § 5 concludes by recalling the questions addressed and summing up the answers provided.

2. ANALYTICAL STARTING POINTS

Evidence presented at trial enables fact-finders to reach accurate verdicts on litigated facts. Evidence is necessary to support the factual claims made by the parties and the findings of fact made by the decision-makers. But evidence per se does not yield verdicts. To this purpose there are at least five requirements on juridical evidence: 1) evidence must be admissible according to the rules of the relevant legal system; 2) it must be presented to fact-finders through some ostensive act; 3) it must be «inferentialized» by the parties and the fact-finders, since evidence doesn't speak for itself and the participants in a dispute have to construct evidentiary arguments based on the items presented; 4) it must be assessed to determine its probative value—or, better, the evidentiary inferences and arguments at stake must be assessed to determine the evidentiary support, or warrant, or justification provided by the premises to the conclusions; 5) fact-finders need to consider whether the evidence meets the relevant standard of proof, or whether the relevant burden of proof has been satisfied⁵.

The points that concern us here are the fourth and the fifth. It is one thing to assess, or to evaluate, the evidence presented. And quite another to find whether it satisfies the relevant standard of proof. The two operations are conceptually different and the latter requires the former as a necessary condition⁶. No one can determine whether the evidence admitted and presented meets the relevant standard, unless the evidence is assessed. To be more precise, once the evidence has been «inferentialized», namely translated into evidentiary arguments, the fact-finders have to assess such arguments in order to see whether the relevant standard is met or not. On the other hand, assessment per se doesn't determine a decision either, because a definite outcome can be justified only if a standard of proof is considered.

So, both assessment criteria and standards of proof are necessary conditions of legal decision-making about facts, and none of them is a sufficient condition of it.

⁵ Many of the points discussed in this article under the heading of «standards of proof» can be also discussed under the heading of «burdens of proof» (see ALLEN, 2014). To make things simpler I won't use the latter terminology and the corresponding conceptual apparatus. (However, to my understanding, the burdens terminology is more focused on procedural and argumentative features, while the standards terminology emphasizes probative and epistemic issues.)

⁶ Notice these different but related points: assessment is necessary for the judgment on standard satisfaction, and both are necessary for a justified decision-making.

This appears quite obvious from a conceptual and philosophical point of view. But, surprisingly enough, some jurists and courts seem to miss the point. Some scholarly literature overlooks the issue, and some decision-makers seem to mix things up, which is more concerning. One significant example is provided by the European Court of Human Rights (ECtHR). In the 2010 case *Gäfgen v. Germany*, for instance, the Court says that «in assessing the evidence on which to base a decision [...] the Court adopts the standard of proof “beyond reasonable doubt”» (§ 92). Apparently in this⁷ and similar decisions⁸ the standard of proof is taken to be an evidence assessment criterion, instead of being a threshold for decision which comes into play once the evidence has been assessed. The same confusion is not infrequent in scholarly work, when evidence assessment criteria are equated with standards of proof or vice versa⁹. One scholar, for instance, asks this question: «Is the standard of proof in the Continental systems, sometimes called *intime conviction* according to its French variant or *freie Überzeugung* in German, really lower than the standard “beyond reasonable doubt” of which common lawyers tend to be so proud?»¹⁰. This question reveals the confusion I am addressing.

Now, why is that? Why do educated scholars and skilled decision-makers fall into this confusion about evidence and proof? Why do they fail to see the difference between assessment criteria and standards of proof? Before advancing my hypothesis on this somewhat surprising phenomenon, let us take a closer look at the rules of some contemporary systems.

⁷ The interest of *Gäfgen* comes also from the fact that it deals with a puzzling «fruit of the poisonous tree» case: conclusive evidence of guilt was obtained as a result of a confession that was allegedly coerced in that the applicant had been threatened by the police for the purpose of finding a boy he had kidnapped. On this case see JACKSON and SUMMERS, 2012: 192-193.

⁸ The understanding of «beyond reasonable doubt» as an assessment criterion appears as a recurrent theme in the decisions of the ECtHR; see also, e.g., *Ireland v. UK* (1978), § 161; *Labita v. Italy* (2000), § 121; *Fedorov v. Russia* (2011), § 57; *Rizvanov v. Azerbaijan* (2012), § 45; *Najafli v. Azerbaijan* (2012), § 36. On the other hand, in these decisions the Court usually adds that such proof may follow from sufficiently «strong, clear and concordant inferences» or «similar unrebutted presumptions of fact», which in fact amount to ways of assessing the evidence (epistemically as to «strong, clear and concordant inferences» and legally as to presumptions). Instead, the Inter-American Court of Human Rights does not apply an explicit standard but according to some commentators (e.g. BOVINO, 2005 and PAUL, 2012) it generally uses the preponderance of the evidence (civil) standard, for its purpose is not the punishment of human rights violations but the protection of victims and the «reparation of damages».

⁹ On the common law side, see, e.g., CLERMONT and SHERWIN, 2002 (treating *intime conviction* as a standard); on the civil law side, see, e.g., IACOVIELLO, 2006 and CAPRIOLI, 2009 (treating the «beyond a reasonable doubt» standard as an evidence evaluation method because of the judicial obligation to deliver written reasoned opinions). Of course, there are also scholars who see well the difference (especially in the Spanish-speaking world: Daniela Accatino, Rodrigo Coloma, Jordi Ferrer, Marina Gascón in particular).

¹⁰ BOHLANDER, 2012: 6.

3. COMPARATIVE CONSIDERATIONS

I sketch in this section an overview of the rules of some actual legal systems about evidence evaluation and proof standards. Some systems have rules that clearly consist of standards of proof, whereas others have rules that clearly consist of evidence assessment criteria, while there are also rules which are difficult to categorize one way or the other. Given the theoretical purpose and the conceptual approach of the present article, I will overlook many positive details and make abstraction from many systemic complexities. What is important for the purposes of this work is the grasping of the ways in which different systems handle the messy issue we are dealing with.

I shall start by making some remarks on the common law world and then move to some civil law countries.

3.1. Common Law Countries

In the common law world there are three basic standards of proof. Two of them apply to civil cases, one to criminal cases. In the following I refer especially to the systems of the US and the UK (England and Wales in particular).

In civil cases the usual standard is the *preponderance of the evidence* (as they call it in the US), or the *balance of probabilities* (as they call it in the UK)¹¹. This means that, according to the received view, the claim of the burdened party (usually the plaintiff) must be more probable than not in light of the evidence presented. If, given the evidence, the factual claim of the party with the burden of proof is considered to be more probable than not, that party is entitled to a verdict in its favor. If it is not considered to be so, the party has no such entitlement.

But in some civil cases the US system adopts a second standard, the *clear and convincing evidence* standard. It is used in «serious cases», namely those in which what is at stake is more important than in the rest. Examples of such «seriousness» are punitive damages cases—for instance for a fraud charge—and immigration cases where what is at stake is the right to be in a certain country¹². This second standard is more demanding than the first. In order to have a verdict in its favor, the burdened

¹¹ By the way, it is possible to speculate about some differences: one of these might be that the UK formula explicitly suggests a probabilistic account of evidence, which the US formula does not; similarly for the «more probable than not» version of the standard, which is even more committed to some probability account, namely to an account of probability that assumes the complementation rule according to which, if the probability of H is p, then the probability of not-H is $1 - p$ (which is not necessarily the case with the «balance of probabilities» version of the standard, and even less so with the «preponderance of the evidence» version if probabilities are understood epistemically rather than mathematically; see COHEN, 1977 and HAACK, 2014: 47 ff).

¹² See REDMAYNE, 1999: 189 (suggesting the adoption of this standard in the UK to face such serious cases).

party must make a greater evidentiary effort, presenting evidence which is stronger, or which has a higher probative value, than that which suffices in the rest of cases.

In criminal matters the well-known standard of guilt, especially entrenched in the US¹³, is the proof *beyond a reasonable doubt* (BARD). This standard requires an amount of evidence that only leaves room for unreasonable doubts about the defendant's guilt. So, if the evidence presented makes it reasonable to believe that the defendant is guilty and makes it unreasonable to doubt it, then the prosecution is entitled to a verdict in its favor and the triers of facts are committed to decide against the defendant. Of course this criminal standard is even more demanding than the clear and convincing evidence standard used in serious civil cases: the evidence presented has to be even stronger than this, leaving room for unreasonable doubts only, which is to say doubts that lack specific epistemic reasons, as spectacularly exemplified by merely skeptical doubts (such as Cartesian doubts and «brains in a vat» doubts¹⁴) or motivated by essentially philosophical concerns cast in probabilistic terms (such as the idea that we cannot be 100% sure of anything¹⁵).

There have been some attempts to translate those qualitative standards into quantitative thresholds. This is done with some probability theory and some assignment of a numerical probability value to each of the standards. It is not a difficult task if we consider the preponderance of evidence, which is easily translated into something such as $>.5$, while the clear and convincing evidence is usually considered to be around $.7$ (or $.75$) and the proof beyond a reasonable doubt is usually located around $.9$ (or $.95$)¹⁶.

The rationale for having a higher standard in serious civil cases, compared to non-serious ones, is the protection of some more important value which is at stake in them. And the rationale for having an even higher standard in criminal cases, compared to civil ones, is the protection of a fundamental value such as individual liberty (and even life in those systems where capital punishment is still in place). This is fairly obvious and there are quite sophisticated pieces of literature that work out the details of this plain remark¹⁷. What is less obvious is the capacity of those standards to do the job they are designed for, due to the lack of empirical information about

¹³ It is so by virtue of some Supreme Court's decisions such as *Miles v. United States*, 103 U.S. 304 (1880) and, in particular, *In re Winship*. 397 U.S. 358 (1970).

¹⁴ PUTNAM, 1995: 17 and PUTNAM, 1981: 1 ff.

¹⁵ A traditional phrase used in various contexts to qualify the certainty required for a guilt verdict is «moral certainty», which falls short of «absolute certainty» and sounds like a standard rather than a criterion.

¹⁶ Of course, the $.9$ value is worrisome for criminal defendants, but the higher the standard is set (to avoid false convictions) the higher is the expected rate of false acquittals. Consider also that, when a 5/6 jury majority is required for conviction (as it was according to the French 1791 system), the corresponding probability is even lower, namely $.83$ (see PADOA-SCHIOPPA, 1987: 92, 142). I cannot expand on these topics here.

¹⁷ See especially BELL, 1987. Cf. STEIN, 2005: 118 ff and LAUDAN, 2006: 63 ff.

particular cases or to conceptual confusion about some aspects of the standards. There is in fact an increasing amount of critical literature (on the criminal standard in particular¹⁸) motivated by the idea that such standards are not really helpful in legal decision-making, insofar as they do not provide genuine epistemic reasons to decide a case one way or the other. This is not the topic of the present article, though. What I want to focus upon is the ultimate nature, so to say, of the standard of proof rules: they are *decision rules*.

The standard of proof rules set the qualitative or quantitative thresholds that must be reached to have a decision in favor of the burdened party. This party bears the risk of there being an amount of evidence which is insufficient with respect to the standard. If the evidence is insufficient, the burdened party has no entitlement to a decision in its favor. If it is sufficient, the party has such an entitlement. So standard of proof rules are in the end decision rules. They justify decisions in favor of one party or the other, based on the factual claims and the arguments given the evidence.

Now, all of that presupposes an assessment of the evidence. An evaluation of the evidence which was admitted, presented and discussed is necessary to the judgment about standard satisfaction. Standards of proof, pace some scholars and judges, are not criteria of evidence assessment. Instead, they presuppose and require some criteria that determine the probative force, or probative value, of the evidence in play, so as to see whether it meets the relevant threshold.

To make a simple exemplification, it is not enough to assign the plaintiff's claim a .65 probability of being correct: in order to make a decision we need to know whether the standard is the simple preponderance of the evidence or the clear and convincing evidence standard. On the other hand, the standard itself does not tell you if the evidence presented satisfies it: you need to assess the evidence, either quantitatively or qualitatively. To take a qualitative example, the well-known decision in *O. J. Simpson's* criminal case can be understood like this: the assessment led to the «he did it» hypothesis as the best explanation of the evidence but it also led to an acquittal given that, according to the jury, the evidence didn't match the BARD standard.

If this reconstruction is correct, where are the evidence assessment criteria in the common law world? Basically, they are in the scholarly disputes about evidentiary value, or probative force, or the weight of (combined) evidence¹⁹. They are not cod-

¹⁸ LAUDAN, 2006: 29 ff. Note that the UK has been recently abandoning BARD, or at least its wording (out of troubles in making it understood): now, following the Judicial College (formerly the Judicial Studies Board), UK fact-finders are to convict only if they are *sure* of guilt (see *R. v. Majid* [2009] EWCA Crim 2563; see also ROBERTS and ZUCKERMAN 2010: 253 ff). However, on BARD's interesting origins in England and Continental Europe, see WHITMAN, 2008 (it was supposed to protect the juror's religious conscience facing the risk of wrongful convictions, it was not originally designed to protect the defendant).

¹⁹ See, among others, ANDERSON, SCHUM & TWINING, 2005: 224 ff; BARZUN, 2008; HAACK, 2014: 208 ff.

ified, as far as I know, in statutory rules nor in any other legislative materials. Truly, there are some documents that give general guidance. Consider for instance the Crown Court Compendium (Part I, on «Jury and Trial Management and Summing Up») by the English Judicial College, where assessment criteria for the jury are given with reference to inferences from expert evidence, bad character evidence, hearsay, and the defendant's silence at trial, to mention some of the issues. For the rest, the criteria are advanced in scholarly controversies and reflected in jury instructions and judicial opinions insofar as judges and courts take a position on such controversies. Consider the literature about probability theories and their application to legal matters and cases; consider the controversy between the supporters of some subjective (Bayesian) conception of probability and the supporters of objective probabilities, or the dispute between Pascalian and Baconian probabilities, or the passionate controversy between the supporters of the relative plausibility account and the supporters of the mathematical probability account, that is, between an ordinal assessment of factual claims and a cardinal one²⁰; consider in addition the debate between atomism and holism in evidence assessment²¹; and consider also the ways in which judges explicate to juries, with the so-called «jury instructions», how to evaluate whether a standard is met. All of this is about the criteria that fact-finders need to use to check whether the evidence presented in a case satisfies the relevant standard of proof.

3.2. Civil Law Countries

Civil law countries usually have in their respective codes some basic rules about evidence and proof. Most of the time these rules set evidence assessment criteria instead of standards of proof. And nowadays these criteria are rather relaxed, leaving room to judicial discretion. This has a definite historical explanation: in many Continental countries the codification effort pointed at getting rid of, inter alia, the prior complexities in the assessment of evidence. Those complexities characterized Roman Canon law and amounted to the so-called system of «legal proof», which lived from the end of the XIII to the end of the XVIII century up to the French Revolution, and which dominated Europe in the XVI and XVII century in particular²². In that system every sensible kind of evidence was assigned a predetermined and fixed probative value, on the basis of some legal authority or doctrinal opinion; then, once some evidence was presented at trial, the judge's task was basically that of calculating the value of each factual claim given the evidence at disposal. In other words, once the *types* of evidence had received a fixed value, the judges were to consider the *tokens*

²⁰ See, e.g., TRIBE, 1971, COHEN, 1977, TILLERS and GREEN, 1988, ALLEN, 1991, GOLDMAN, 2002, REDMAYNE, 2003, ALLEN and PARDO, 2007, KAPLOW, 2012, ALLEN and STEIN, 2013, VERHEIJ et al., 2016, ALLEN and PARDO, 2019.

²¹ See DAMAŠKA, 1997a: 34-37, HAACK, 2014: 235-238 and SCHWEIZER, 2014.

²² See among others DAMAŠKA, 1986: 29 ff, PADOA-SCHIOPPA, 2003: 280-292, JACKSON and SUMMERS, 2012: 57 ff.

in the case at hand and were to calculate the outcome accordingly, with no or little discretion on their part²³. The best-known example of that is the traditional two-witness rule according to which two independent witnesses testifying the same amount-ed to «full proof» of the fact so testified, leaving no room to the judicial appreciation of the witnesses' credibility, of the circumstances of testimony, etc. And at the same time, according to that rule, a single testimony was not sufficient to prove a claim, however credible and accurate the testimony might have been. Then, to cut a long story short, Enlightenment legal thinkers and French Revolutionary authorities designed a system that, on the contrary, trusted fact-finders conferring some discretion upon them but, at the same time, stating guidelines of evidence assessment²⁴. Eventually the spirit of the Revolution spread throughout Europe and civil law countries adopted the new framework, which is still in place.

Let me start with Italy. The actual code of civil procedure states that fact-finders must evaluate the evidence according to their «prudent assessment» (*prudente apprezzamento*, Art. 116), unless a criterion of «legal proof» (*prova legale*) is mandated. There are in fact some provisions attributing legal value to some kinds of evidence such as official documents (Art. 2700). But these are the exceptions to the general rule of «prudent assessment». This rule provides a general criterion of evidence assessment. Then the question is: what is the standard of proof in the Italian system of civil procedure? It is hard to say, because there is no explicit rule neither in the code nor in statutory law. In the last years—against the view that the standard is something like «moral certainty» or even full conviction—some courts have felt the need to specify that the standard of proof in matters like medical malpractice is the «more probable than not» idea²⁵. So, according to these judges, it is basically the standard used in common law countries. On the other hand, it is unclear whether the standard is conceived to apply only to such tort matters as medical malpractice or more generally to any matter of civil law and procedure.

The situation is different if we take into consideration the Italian criminal procedure code: fact-finders shall convict only if the BARD standard is met (Art. 533, as modified in 2006)²⁶. The standard was introduced by Parliament some years ago

²³ But see DAMAŠKA, 2003: 129: «It is true that these lawyers developed a maze of rules about the quantity and quality of evidence needed for fact determinations. But it is a mistake to believe that these rules turned adjudicators into automatons, who made factual determination on the authority of rules, independently of their beliefs». On the other hand, DAMAŠKA, 1986: 55 points out that «even today the Continental “free evaluation of evidence” is not really free: as befits the hierarchical process, trial judges are required to justify their findings of fact, and the cogency of their reasoning is scrutinized by appellate courts». Cf. DAMAŠKA, 2018.

²⁴ A key event was the adversary reform and the introduction of the criminal jury in France by the law of September 16-29, 1791 (statute of September 16, decree of September 29).

²⁵ See, e.g., Cass. sez. III civile, decision n. 10285/2009, n. 10741/2009, n. 15991/2011, n. 23933/2013, n. 18392/2017.

²⁶ A nice aspect of the Italian formula is that it reads *oltre ogni ragionevole dubbio* («beyond all reasonable doubt»). There isn't any logical difference with the «beyond a reasonable doubt» formula (for

with the purpose of making criminal convictions harder. And it was supported by the belief that Americans better protect criminal defendants. Some commentators objected, saying that the idea was already implicit in the system, and that therefore that legislative intervention was superfluous²⁷. But this is not our concern here. The interesting question is, given the now explicit standard, whether the code mandates any assessment criterion: in fact, as a general matter, there is just one laconic provision according to which the judge assesses the evidence giving an account of the criteria adopted and the results obtained (Art. 192 c. I). Literally construed, this provision makes any assessment criterion acceptable, provided that the judge specifies the method used and the results arrived at. Anything goes, apparently. But this would be an inappropriate construal of the text; for the system purports to be rational and to have evidence rationally assessed²⁸. To take a somewhat extreme example, the judge could not say something like this: «I have consulted a clairvoyant and reached the conclusion that the defendant is guilty». Consulting a clairvoyant would be the method used, and the defendant's guilt would be the result arrived at. So, literally speaking, the judge would comply with the rule that requires making method and results explicit. But that would run counter to the goal of rationality. Additionally, there are more definite rules concerning things like circumstantial evidence (that must be «serious, precise and consistent» to prove a fact, according to Art. 192, c. II) and the declarations of the co-defendant (that must be assessed considering the evidence that corroborates their reliability, according to Art. 192, c. III). So one can point out that Italy has a criminal procedure system with an explicit standard of proof but, as a general matter, no explicit assessment criterion, whereas the civil procedure system has it the opposite way, for it has an explicit assessment criterion and no explicit standard of proof.

Let's consider France now. According to the French civil system there is «free proof» of legal facts (see Arts. 9-10, 179, 198, 213 of the code of civil procedure)²⁹ and «legal proof» of (certain) legal acts (see, e.g., Arts. 1359, 1371 of the civil code). No explicit standard is given, and the situation is similar to the Italian civil one.

The most interesting thing of the French system is the highly controversial notion of *intime conviction* used in criminal trials. According to the criminal procedure code, judges must decide on the basis of the evidence according to their *intime conviction* (Art. 427); the same holds for mixed courts, made of judges and jurors who

here «a reasonable doubt» means «any of the doubts that are reasonable»); but the rhetorical emphasis of the «all» formula appears to be heavier. Similarly for the Chilean formula (see below in the text) «más allá de toda duda razonable».

²⁷ See the reconstruction of the debate in DELLA TORRE, 2014: 9-11.

²⁸ The fundamental right to have evidence examined and the other «fair trial» (*giusto processo*) requirements militate for this conclusion. See Arts. 24 and 111 of the Italian Constitution, Art. 6 of the European Convention on Human Rights, and Art. 47 of the EU Charter of Fundamental Rights.

²⁹ See also Art. 1381 of the French civil code on the probative value of testimony, to be determined by the judge.

are not bound by legislative rules of assessment and must ask themselves «in silence and meditation» (*dans le silence et le recueillement*) if, given the evidence, they have an *intime conviction* about the defendant's guilt (Art. 353; see also Art. 304). There has been a lot of talk about this idea. Several scholars have criticized it assuming it is nothing but an appeal to the fact-finder's psychology³⁰. If it were so, it would not provide any epistemic ground for the verdict and it would open the door to any form of bias, suggestion and emotive decision-making. Now, frankly, this is an uncharitable reading of the rule. The relevant texts underscore the need to fix that state of mind *on the basis of the evidence*; it is not evidence-free, it does not boil down to the purely subjective and psychological feeling about the defendant's guilt. It is true that the appeal to «silence and meditation» strikes an Anglo-American scholar as utterly different from the idea of heated disputes between jurors, with arguments and counterarguments being debated until a decision is made. It strikes even a civilian like me in that it appears as the «monk way» to the discovery of truth, or, philosophically speaking, the «Cartesian way» to legal fact-finding (as is well known the dimension of solitary meditation is a peculiar feature of Descartes' philosophy). But *intime conviction* should not be taken as the legitimation of any kind of subjectively formed or emotionally driven decision³¹. Originally it was adopted with the idea of transplanting into the French system the decision-making way of English juries³². Given the Enlightenment trust in the judgment of lay people (and a lesser trust in the opinion of professionals like judges) the French system adopted the jury method for criminal trials and set *intime conviction* as the check on the defendant's guilt. Eventually the idea of *intime conviction* was extended to mixed courts and to professional fact-finders such as judges.

So, is *intime conviction* a standard of proof? Is it rather an evidence assessment criterion? Is it both? Anglo-American commentators tend to think it is a standard³³. Ironically, French jurists and civil lawyers tend to think it is not a threshold, being instead a way to assess the evidence and make a decision³⁴. And even reputed scholars and evidence theorists of the civil law world disagree on the understanding of it; for instance, Jordi Ferrer thinks it is an «entirely subjective» standard of proof,

³⁰ See, e.g., TARUFFO, 2003b: 81-82, LAUDAN, 2005: 98-99, FERRER, 2007: 144-145, BAYÓN, 2008: 17-18. An echo of this criticism can be found in some judicial opinions in different countries: see, e.g., Tribunale di L'Aquila, sez. penale, decision n. 380/2012, § 4.1 (which is a decision on the criminal charges relating to the L'Aquila earthquake of 2009).

³¹ HANS and GERMAIN, 2011: 755: «the decision based on *intime conviction* is not best viewed as the expression of a feeling, but rather as a considered opinion based on the charges, evidence, and defenses presented by the parties».

³² See PADOA-SCHIOPPA, 1987 and 1994.

³³ See, e.g., CLERMONT and SHERWIN, 2002: 256. Cf. HANS and GERMAIN, 2011: 754.

³⁴ See, e.g., BREDIN, 1996: 23, DELMAS-MARTY, 1996: 59-60, BOULOC, 2004: 54, CARTIER, 2004: 70, BERGEAUD, 2010: 160; see also TARUFFO, 2003a: 666, CASTILLO DE LA TORRE, 2009: 522 and UBERTIS, 2013: 332.

whereas Daniel González Lagier says it is a «system of evidence evaluation»³⁵. To my understanding the reading of it as a standard is in fact supported by the presence, in the same code, of a provision according to which judges are to assess freely every evidentiary item (Art. 428). This is an evaluation criterion and if *intime conviction* were a criterion itself the system would be either redundant or misleading.

Let us turn to Spain now. According to the civil procedure code, evidence must be assessed according to the «rules of sound criticism» (*reglas de la sana crítica*, e.g., in Art. 348 on expert testimony, and in Art. 376 on witnesses). The situation is similar to that of other civil law countries in that «legal proof» is the exception, not the rule. Here the rule is assessment according to «sound criticism». Some commentators appreciate this criterion insofar as it is more epistemologically inclined than the mere idea of a «free assessment»³⁶. Many Spanish-speaking countries in Latin America adopt the same criterion. Such an epistemological flavor is missing instead from the Spanish code of criminal procedure, according to which judges decide assessing the evidence with their «conscience» (Art. 741.1). The appeal to conscience seems to be in line with the French appeal to *intime conviction*. Other Spanish-speaking countries follow the same path, except Chile where both «sound criticism» and «beyond a reasonable doubt» are used in criminal cases (see Art. 369-bis of the criminal code and Art. 340 of the criminal procedure code)³⁷. The case of Chile is interesting because it provides a confirmation of the conceptual intuition that assessment criteria and standards of proof are different things. Chile adopts «sound criticism» as the evidence assessment criterion to be used in criminal cases, and «beyond a reasonable doubt» as the standard of proof to be used there. This double discipline is not redundant, for criteria and standards are different³⁸.

Finally, a few words on Germany. According to the civil procedure code, the court is to assess the evidence freely (*Freie Beweiswürdigung*): it is supposed to decide «at its discretion and conviction» and taking into account «the entire content of the hearings and the results obtained by evidence being taken» (Sec. 286.1); the court «shall be bound to statutory rules of evidence only in the cases designated in the present Code» (Sec. 286.2). The emphasis on judicial discretion confirms the trend of civil law systems in civil matters³⁹. Coming to the German code of criminal

³⁵ FERRER, 2007: 145 and GONZÁLEZ LAGIER, 2013: 52.

³⁶ COLOMA and AGÜERO, 2014; LUCH, 2012: 194-196; TARUFFO, 2003b: 82.

³⁷ See also Art. 297 of the criminal procedure code, that establishes the free assessment of evidence in compliance with logic, experience and science. This provision does not mention «sound criticism» but is usually read as a reformulation of this traditional idea; the previous code of criminal procedure (substituted by the present one in 2000) mandated an evaluation of the evidence according to the «rules of sound criticism» (Art. 562).

³⁸ See ACCATINO, 2011 and COLOMA, 2012. See GONZÁLEZ, 2006 on civil aspects too.

³⁹ Just before the German reunification a scholarly dispute occurred on the alleged «German advantage in civil procedure»: see LANGBEIN, 1985, ALLEN et al., 1988, LANGBEIN, 1988 and ALLEN, 1988. The dispute is still interesting for a cost-benefit analysis of party-dominated and judge-dominated systems of procedure and fact-gathering in particular. See also SCHWEIZER, 2016 on the civil

procedure, the court shall decide «on the result of the evidence according to its free conviction gained from the hearing as a whole» (Sec. 261). Again, the emphasis is on free assessment and free conviction (*Freie Überzeugung*)⁴⁰. But still on the basis of the evidence and the hearing. So, both in the civil and criminal context German law emphasizes the role of judicial discretion in evidence assessment. No standard of proof is made explicit in positive law, unless one reads «free conviction» as a standard, which is puzzling as much as the idea of having a «free standard». But the attentive reader would have noted that those provisions make reference to decision, not only to evidence assessment. So, such evidence assessment criteria are supposed to play also the role of decision rules.

To summarize, contemporary civil law countries are similar to one another in that, with respect to civil matters, «legal proof» is the exception and «free assessment» of evidence is the rule. According to some scholars there are degrees of such freedom, for the French way is somewhat extreme in giving discretion to fact-finders, while the Spanish way puts more epistemological constraints on their judgment, and the Italian way of «prudent assessment» appears to be in the middle. Michele Taruffo has claimed that «free assessment» is not to be taken as freedom from the rules of logic, the requirements of rationality, and the constraints that science and epistemology put to the triers of fact⁴¹. It is simply freedom from predetermined and fixed rules of evidential weight; freedom from the old system of «legal proof» (with the few exceptions that still remain). It is the idea of having evidence assessed case by case by sensible fact-finders, by working minds that are able to appreciate the particularities of concrete cases and to make responsible factual determinations. Contemporary civil law countries display a lower degree of similarity with respect to criminal matters, since some of them (like Chile and Italy) have explicitly adopted the BARD standard of proof, while others just stick to the Continental idea of a «free» assessment of evidence⁴².

standard of proof in Switzerland and Germany (claiming that it is lower than what doctrine and case law proclaim, namely «full conviction»).

⁴⁰ See, e.g., BOHLANDER, 2012: 8 ff, 32.

⁴¹ TARUFFO, 2009: 160 ff, 219. See also GASCÓN, 2010: 32.

⁴² There are of course further interesting things I cannot examine here: for instance, Sweden has a unique code of procedure (concerning both civil and criminal matters) which mandates the free evaluation of evidence (Chap. 35, Sec. 1) and is quite generous as to evidence admission (see DAHLMAN, 2015 on the use of prior convictions as evidence of guilt); in the Netherlands Sec. 338 of the criminal procedure code mandates free evaluation through the hearing, but Sec. 342.2 states that the court may not find there is proof that the defendant committed the offence as charged in the indictment exclusively on the basis of the statement of one witness (see MACKOR, 2016 on the ways in which the Dutch Supreme Court and some scholars have tried to complement this *unus testis* rule); in Slovenia no explicit standard is mandated, but in criminal matters the Supreme Court has introduced BARD with the holist assessment criterion of interconnected circumstantial evidence (see NOVAK, 2016); in Poland Sec. 233.1 of the code of civil proceedings gives discretion in the evaluation of evidence but requires it to be holistic (I've been told so by M. Araszkiwicz, whom I thank).

The following tables summarize the above findings (the considered countries are in the columns⁴³; in the rows «AC» stands for «assessment criterion» and «SP» for «standard of proof»; in the cells, «?» stands for a missing or unclear aspect, «POTE» stands for the «preponderance of the evidence», «BARD» stands as usual for «beyond a reasonable doubt», and «IC» stands for *intime conviction* or a similar mental state):

Tab. 1. Civil Cases

	US	UK	I	F	S	C	G
AC	?	?	Prudent	Free	Sound Criticism	Sound Criticism	Free
SP	POTE	POTE	?	?	?	?	?

Tab. 2. Criminal Cases

	US	UK	I	F	S	C	G
AC	?	?	Free	IC?	Con-science	Sound Criticism	Free
SP	BARD	Certainty	BARD	IC?	?	BARD	IC?

A few systems among those considered in these tables are fully explicit (just Italy and Chile on criminal procedure); the majority are half-explicit, for some of them (common law ones) only have explicit standards of proof, and the others (civil law ones) only have explicit assessment criteria.

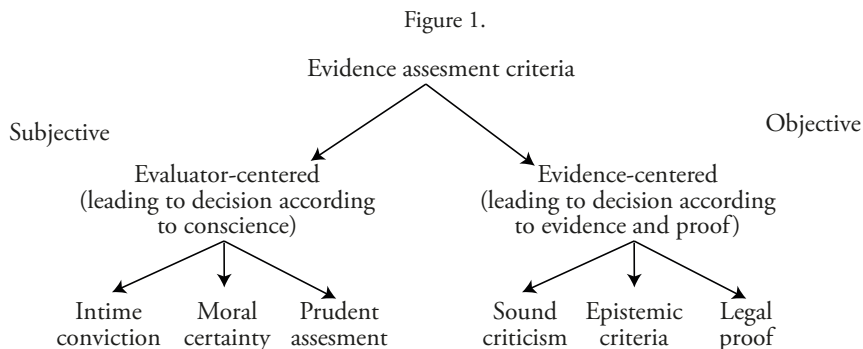
Overall it is not an equilibrium situation, because the general discretion of evidence assessment and the lack of explicit standards in some contexts generate not only additional litigation and controversy but also confusion as to what is missing in, and prescribed by, a given legal system (like the civil procedure one in Italy, for instance).

As to assessment criteria, in a comparative scheme they can be ordered according to the degree of freedom they confer upon the fact-finders, stipulating that a criterion is as «objective» as it decreases such freedom and is as «subjective» as it increases it. So it is not a categorical distinction, it is a matter of degree. Maximal objectivity minimizes judicial discretion, and subjective criteria maximize it⁴⁴. On one side we can locate the fixed criteria of legal proof; on the opposite side we can locate *intime conviction* (if it's read as a subjective criterion); next to legal proof we can locate epis-

⁴³ To avoid any misunderstanding: «US» stands for the United States, «UK» for the United Kingdom, «I» for Italy, «F» for France, «S» for Spain, «C» for Chile, and «G» for Germany.

⁴⁴ Even though it may happen that in practice there are «subjective patterns» or intersubjective ones.

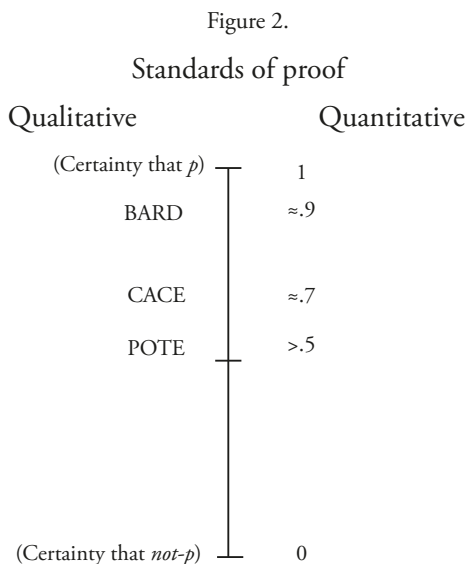
temic criteria, and moral certainty next to *intime conviction*; in the middle, sound criticism is closer to objectivity and prudent assessment is closer to subjectivity (see Figure 1). Probabilistic accounts can be located too in the scheme, according to the probability theory they endorse (subjective v. objective, Pascalian v. Baconian, etc.).



In other words, objective criteria are evidence-centered since they focus on the evidence at disposal, while subjective criteria are evaluator-centered since they focus on the states of mind and attitudes of the fact-finder (on prudence, conscience, conviction, etc.). This is also in tune with the subject-matter of the old controversy over decision according to conscience (*secundum conscientiam* in Latin) or according instead to evidence and proof (*secundum alligata et probata*)⁴⁵.

As to standards of proof, in a different comparative scheme they can be ordered according to the degree of belief on the claim at stake (qualitative reading) or according to the probability value that the claim is attributed (quantitative reading). So, if the claim is p , at one extreme of a vertical segment we have the certainty that p (or probability 1), at the other extreme the certainty that *not-p* (or probability 0), and in the middle the different degrees of belief about p and the three basic standards, where «BARD» stands for «beyond a reasonable doubt», «CACE» for «clear and convincing evidence» and «POTE» for «preponderance of the evidence» (see Figure 2). (Note that this assumes the conventional, or Pascalian, view on probability; a Baconian view would need a different scheme.)

⁴⁵ For instance, in a medieval society it was not unlikely that the same person acted as a confessor and as an ecclesiastical judge; then a problem was posed by the information the judge received as a confessor, especially by a crime confession: were such judges permitted to use it in their judicial capacity? So, were the judges to decide according to conscience or to evidence and proof? See PADOA-SCHIOPPA, 2003: 251 ff.



As an additional point, note some practical consequences of the above distinctions. Note that the scope of legal argumentation is significantly reduced when fixed criteria of evidence assessment are mandated, or as far as these criteria are objective. This was pretty obvious in the old system of «legal proof», where any discussion about the witnesses' credibility or the circumstances of testimony was beside the point once the fact-finders were bound by some measure like the two-witness rule. The more discretion fact-finders have, the more prominent is the role of legal argumentation. Parties are supposed to provide persuasive arguments about evidence and fact-finders are supposed to provide arguments that justify their decisions. This is the case in particular when judges have to provide written opinions that (supposedly) justify the decisions they made. *Intime conviction* is frequently ridiculed as epistemologically unsound, but consider how frustrating it is to be bound by fixed criteria that impede any appreciation of the particular circumstances of the case and block any argumentative effort to make sense of the evidentiary items at disposal.

Interestingly enough, inquisitorial systems are historically tied to legal proof criteria⁴⁶; instead, *intime conviction* was introduced in France in the adversarial spirit of English trials, and if criteria are subjective and there is room for argumentation, parties have a fair chance to make their points (the worst scenario for a criminal defendant is an inquisitorial system with subjective criteria of evidence assessment). Remember also that, according to certain surveys⁴⁷, lay triers of fact judging on *inti-*

⁴⁶ See FERRAJOLI, 1989: 112-113.

⁴⁷ HANS and GERMAIN, 2011: 746.

me conviction are more prone to acquittal than professional decision-makers (against repeated cries that subjective criteria are very bad). Hence something like *intime conviction* is not that bad if it is: i) understood in connection with evidence; ii) encapsulated in an adversary system with argumentation coming from both parties; iii) used by lay fact-finders.

Before moving on, let me also stress that fixed criteria are not to be equated with explicit ones: the fixed criteria of the «legal proof» system were of course explicitly mandated by authoritative legal sources, but these can also mandate relaxed or subjective criteria that leave discretion to fact-finders. Whether a criterion is fixed or not is a substantive issue. Whether it is explicit or not is simply a matter of positive law.

4. THE FUNCTIONAL CONNECTION ACCOUNT

A comparative analysis of the foregoing remarks focuses on the relevant systemic differences. Some legal systems have explicit standards of proof (as decision rules) and lack explicit assessment criteria. Some other systems have explicit assessment criteria and lack explicit standards of proof (again, understood as decision rules). In addition, it is hard to say whether the non-explicit components are implicit or undetermined. For instance, it is *prima facie* hard to say whether in the Italian civil procedure system there is an implicit standard of proof or none.

If all of that is correct two further questions arise, after the first we posed above (in § 2) regarding the confusion between criteria and standards in scholarly work and judicial opinions. The second question is: why do systems differ? That is, why do some of them have assessment criteria and lack standards of proof, while others have it the other way round? And the third question is: how can systems work without one of such elements? That is, if assessment criteria and standards of proof are necessary components of legal fact-finding, how can it be the case that a system that lacks one of them can work nevertheless?

The second of the three questions mainly finds its answer in history. The above-mentioned differences between systems (and the relative confusion in scholarly works and judicial opinions) are generated by different historical, institutional and procedural concerns. The present article is not a piece in legal history, so I won't expand on this point. Suffice it to say that the common law world shows a prominent concern for standards of proof and this finds an explanation, I guess, in the features of jury trial. The fact that the triers of fact are lay people, who need instructions from judges but don't provide written reasons for the decisions they make, militates in favor of a simple system where fact-finders have to consider a standard of proof and can be instructed by judges about the meaning of the standard and the ways to evaluate whether the standard is met. This facilitates the task of the jury, and arguably it does so without dictating how jurors are meant to evaluate the evidence.

On the other hand, the civil law world shows a prominent concern for evidence assessment and this finds an explanation in the features of trials with professional fact-finders like judges, who once were supposed to apply the complex system of «legal proof» and eventually, in contrast with the previous binding system, were asked to freely assess the evidence. Notice that these judges are committed to justify their decisions, namely to give ex post written reasons for the final decisions they make both on factual and legal claims; to this purpose it is not sufficient to say something like «We have been persuaded that the burden of proof was satisfied»: it is essential to justify the way the evidence was evaluated, and to understand how a certain degree of conviction was determined, because the justification of the decision depends not only on the satisfaction of the proof standard but also on the correctness of the process. This focus on the assessment process has likely contributed to the neglect of the standards issue in the civil law world.

The puzzling nature of *intime conviction* is less puzzling if we consider it in the perspective of that historical process. The *intime conviction* idea was introduced in France to emulate the English way to decide criminal cases with jurors and eventually it was extended to judges freed from the complexities of the «legal proof» system; so it is Janus-faced in that it reflects different legal cultures and traditions and because it has absorbed, so to say, the different concerns of these cultures and traditions.

This also explains some conceptual confusion about evidence and proof. Civilians often take standards as assessment criteria because they are more familiar with the latter (consider again the ECtHR decision quoted above). And vice versa common lawyers often take assessment criteria as standards because they are more familiar with the latter (think again of their understanding of the French *intime conviction*). But consider that those views are not entirely unjustified as far as there are some functional connections between criteria and standards: some standards incorporate assessment criteria and vice versa. Consider again BARD: it is arguable that the reasonable doubt standard incorporates the idea of a reasonable assessment of the evidence (which is not very informative, but still a form of guidance). Similarly, for *intime conviction*: if it is an assessment criterion, it incorporates the threshold of the fact-finder's conviction as a decision rule.

These last remarks guide us to the answer to the third question asked above. *There is some connection* between criteria and standards, notwithstanding their conceptual difference. What kind of connection? There is a *functional connection* between them which must be distinguished from a sort of functional identity or equivalence⁴⁸. The idea of a *functional equivalence* would consist in the fact that standards and criteria perform the same procedural and systemic function. This idea has some intuitive interest, for it would explain away the puzzle of systems working fairly well without

⁴⁸ For the equivalence idea see LLUCH, 2012: 192 ff. On functional explanation and functional analysis in different areas (biology and philosophy of mind) see, e.g., WRIGHT, 1973, CUMMINS, 1975, FODOR, 1985.

one of the apparently necessary components of factual decision-making. Remember what the question was: how can systems work without one of such elements, if we assume that assessment criteria and standards of proof are necessary components of legal fact-finding? The functional equivalence idea would explain the puzzle away because it would claim that criteria and standards perform the same function. So there would be no mystery in the fact that systems which have one but not the other component work fairly well. It would be unnecessary to have both, it would be redundant. Because they perform the same function, namely the function of guiding factual decision-making.

The functional equivalence idea is interesting but has some drawbacks. One is the (perhaps annoying) generality of the function appealed to: «guiding factual decision-making» is remarkable for its lack of specificity; but it's true that functional accounts always abstract from many details, insofar as they consider that such details are not relevant to the performance of the function envisaged (if the function is writing you don't bother about the differences between pens and pencils, which you bother about if you need devices which draw signs that cannot be canceled)⁴⁹.

A further and more significant drawback is the failure of the functional equivalence idea to make sense of some interesting cases. What about the Italian system of civil procedure? As stated, it has an explicit assessment criterion but no explicit standard. And there is an increasing debate about the standard or the standards that judges should use. If criteria and standards were functionally the same, this debate would be meaningless because the problem would not subsist. On the contrary there is a debate because it is unclear whether there is a standard and what this standard might be, whereas no one has doubts about the existence of an assessment criterion (Art. 116 of the civil procedure code) however vague it may be. The existence of this assessment criterion doesn't solve by itself the standard of proof problem. So there is no equivalence. Moreover, what about the Chilean system of criminal procedure? As mentioned, in this system you find both a standard (BARD) and an assessment criterion («sound criticism»). In this respect, assuming the functional equivalence approach, the system would be redundant. This conclusion would follow from the idea that the two perform the same function. It would be redundant to have a standard once you have a criterion, or vice versa. Now, this understanding of the Chilean system seems uncharitable, if not unreasonable, and I consider this to militate against the functional equivalence approach.

Indeed, Chile is not the only system that has both things explicit: the criminal procedure code of Colombia mandates BARD, since according to Art. 7 and Art. 381 guilt must be proven «más allá de toda duda» and according to Art. 372 evidence will serve to generate judicial knowledge of the relevant facts «más allá de duda razonable»⁵⁰. At

⁴⁹ On functional accounts of law see GREEN, 1998 and EHRENBURG, 2016. See also the sociological literature, e.g., PARSONS, 1951 and FRIEDMAN, 1975.

⁵⁰ See also Art. 91 and Art. 101 of the same code.

the same time the code mandates (i) a general holistic assessment criterion (according to Art. 380 evidentiary items will be assessed «en conjunto»)⁵¹ and (ii) more specific assessment criteria for specific kinds of evidence (for instance, Art. 404 states that lay testimony will be evaluated considering scientific and technical knowledge about perception and memory, checking in particular the nature of the thing perceived, the functioning of the witness' perceptual system, the circumstances of the perception, the memory processes and the demeanor of the witness at trial)⁵². All of this would be superfluous and redundant were standards and criteria functionally equivalent; on the contrary, such a system is highly interesting for the level of detail of the regulation concerned. The French criminal procedure system is not substantially different if *intime conviction* is taken as a standard (Art. 353 and Art. 427 of the criminal procedure code) and free assessment is the relevant criterion of assessment (Art. 428). And again, this would be redundant if criteria and standards were functionally equivalent.

A more promising idea is that of a functional connection between criteria and standards: *criteria indicate how to assess evidence and standards indicate how to make a decision once the evidence at disposal has been assessed*. A system like the Chilean code of criminal procedure makes everything explicit in requiring an assessment according to the principle of «sound criticism» and a decision according to the BARD standard. Similarly, for the even more detailed system of Colombia. The systems which are less explicit have just one component stated in positive law; then there are two possibilities: 1) that the other component is somehow incorporated in the explicit one; 2) that the other component is undetermined.

According to the first scenario, there can be explicit standards of proof and implicit assessment criteria, or vice versa explicit criteria and implicit standards. The incorporation idea consists in the fact that some system implicitly conveys a certain component (criterion or standard) incorporated in the explicit one or in the rest of the normative context (its rules, principles, etc.). I would locate the Italian criminal procedure system in this category, since it is characterized by an explicit standard (the BARD one) and an implicit assessment criterion consisting in a free but rational assessment of the evidence. Similarly, common law systems that adopt the BARD standard appear to convey at the same time the requirement of a reasonable assessment of the evidence; judicial instructions to juries perfectly fit this picture. Similarly, «legal proof» criteria usually incorporated standards of proof in that they set evidence thresholds and decision rules. Consider again the traditional two-witness rule, according to which two independent testimonies of the same amounted to «full

⁵¹ Similar holistic criteria can be found in European systems about circumstantial evidence: see, e.g., Art. 1382 of the French civil code and Art. 2729 of the Italian civil code on presumptions; cf. Art. 192 c. II of the Italian criminal procedure code.

⁵² In the same Colombian code see also, in particular, Art. 420 on scientific testimony and Art. 432 on documentary evidence. See also Art. 176 of the Colombian *Código general del proceso*: evidence shall be assessed «en conjunto» and complying with the «reglas de la sana crítica».

proof» of the fact: it was a rule about the value of testimonial evidence and, at the same time, it set a threshold for decision.

According to the second scenario, there can be systems where one of the necessary components of legal fact-finding is made explicit and the other is undetermined. What happens then in such systems? The analytical gap is filled by legal practice, with the possible support of legal theory and scholarly opinions. Legal practitioners and participants in legal disputes take sides about the undetermined component; arguments are developed and decisions are made; and in the long run, if no legislation intervenes, some component entrenched in legal practice will emerge as dominant through widely accepted arguments, judicial precedents, etc. The gap will be filled by legal practice. And in this kind of context the importance of legal argumentation is greater than elsewhere in this respect because legal practitioners and scholars need to construct and develop persuasive arguments to fill the gap. As an example, I would locate the Italian civil procedure system in this category, since it is characterized by an explicit assessment criterion and the lack of a positive standard of proof. This system differs from the criminal procedure system of Italy because in the latter, according to most commentators and practitioners, there is an implicit component (the implicit assessment criterion); as to the former, on the contrary, no general agreement exists about the component which is not explicit (the standard of proof); therefore, it is arguable that this component is undetermined and it is up to legal practice to fill in the gap. Recent judicial opinions in Italy confirm this reading.

But one may astutely object, saying the following: how could it be the case that Italian civil judges have been deciding disputes over the years if the standard of proof, being a necessary component of fact-finding, were completely undetermined? Decisions would have been impossible. So, there must have been a standard. This is true. However, this sort of *reductio ad absurdum* of the argument I provided is fallacious because it does not take into account the gap-filling role of practitioners, attorneys and judges in particular: these social actors contribute to that gap-filling even if what they do is not transparent to them. Even before they used the Anglo-American vocabulary and conceptual apparatus of standards of proof, for sure Italian legal practitioners had in their mind some idea about the evidence threshold that justified a civil decision, for it is plausible to think that they had some intuition about such sufficiency threshold. (Of course, the point applies to other jurisdictions.) When is the evidence sufficient to justify a certain belief about the litigated facts? When is it sufficient for the acceptance of a certain version of the litigated facts? You need not have the vocabulary of the «standards of proof» to grasp this idea and the problem. Then solutions come along with legal practice and argumentation.

So, in some way or other, both assessment criteria and standards of proof are needed. However, one may think that current legal theory is «forcing» the practice of civil law systems. It is putting pressure on legal practitioners, so the arguments would go, to compel them to use the standards vocabulary. To address this remark, recall some of the facts. Why don't civilians usually discuss standards? Because they do criteria. Why do common lawyers neglect assessment criteria? Because they focus

on standards. To have an explicit standard is, most of the time, to have an implicit criterion. And to have an explicit criterion is, most of the time, to have an implicit standard. Therefore, theory does not force practice if it makes explicit what is implicit in the system because of some functional connection. And theory does not force practice if it tries to give an accurate account of its gap-filling role when a necessary component of fact-finding is undetermined.

In sum, the functional connection account successfully explains how systems work without one of such analytically necessary components of factual decision-making (or, better, without one such component being made explicit in positive law). Systems work nevertheless either because there is some incorporation of the lacking element in the explicit one, or because legal practice fills the gap when one such component is undetermined.

5. CONCLUSION

After distinguishing assessment criteria from standards of proof, the first question we asked (in § 2) was why some jurists and courts do not see their difference. The second question (in § 4) was why some systems have criteria and others have standards. And the third question (again in § 4) was how systems work without one of such necessary components of legal fact-finding.

The answer to the first question came with the answer to the second. Some historical, institutional and procedural differences explain why some systems have assessment criteria and others have standards. This is basically the answer to the second question. And this answer suggests that jurists belonging to one context are prone to see also elsewhere what they are more familiar with. This means that common lawyers have some tendency to take assessment criteria as standards and, vice versa, civil lawyers have some tendency to see standards as assessment criteria. The less familiar thing is understood in terms of the more familiar one. More importantly, when different systems come to overlap to a certain extent (as is the case in the jurisdiction of courts like the ECtHR) parties and judges who come from different contexts and backgrounds run the risk of mixing things up. Some terms, concepts and institutions that belong to one context are used in combination with terms, concepts and institutions of a different context. Some merging of the issues (assessment and standards) comes from the work of supranational courts such as the ECtHR, insofar as they use some traditional common law terminology and conceptual apparatus to deal with civil law countries issues, or vice versa. Not surprisingly, the result can be confused and confusing to a certain extent. This is one of the prices of a more integrated or even globalized legal world, where different concepts and institutions play together and interact in ways that are not always foreseeable⁵³.

⁵³ For a similar point see HANS & GERMAIN, 2011: 738, 761-763. Cf. DAMAŠKA, 1997b.

On the other hand, there are international jurisdictions that set things more clearly: for instance, the Rome Statute of the International Criminal Court (ICC) mandates BARD as the conviction standard of proof (Art. 66.3), while the related Rules of Procedure and Evidence establish the free assessment of evidence (Rule 63.2) for relevance and admissibility purposes (in accordance with the discretion conferred upon the decision-makers by Art. 64.9, Art. 69.4 and Art. 74.2 of the Rome Statute)⁵⁴. Here standards and criteria are clearly distinguished. Also, the International Court of Justice has a clear view; for instance, in the well-known 2007 decision on *Bosnia and Herzegovina v. Serbia and Montenegro*, the Court is careful to distinguish three «matters»: the burden of proof (§§ 204-207), the standard of proof (§§ 208-210) and the methods of proof (§ 211ff) as assessment of evidentiary weight (§ 213).

The answer to the third question asked above is more theoretically intriguing. It consists in the claim that there is a functional connection between assessment criteria and standards of proof. Because of that connection, one element can incorporate the other. And because of that connection, where one of those things is neither made explicit in positive law nor incorporated in the explicit component, legal practice and scholarly literature make an effort to fill in the relevant gap, providing answers and solutions to the questions and problems that actual cases pose to legal practitioners and scholars.

6. BIBLIOGRAPHY

- ACCATINO, D., 2011: «Certezas, dudas y propuestas en torno al estándar de la prueba penal», *Revista de Derecho de la Pontificia Universidad Católica de Valparaíso*, 37(2): 483-511.
- ALLEN, R.J., 1988: «Idealization and Caricature in Comparative Scholarship», *Northwestern University Law Review*, 82(3): 785-807.
- 1991: «The Nature of Juridical Proof», *Cardozo Law Review*, 13: 373-422.
- 2014: «Burdens of Proof», *Law, Probability and Risk*, 13: 195-219.
- ALLEN, R.J. et al., 1988: «The German Advantage in Civil Procedure: A Plea for More Details and Fewer Generalities in Comparative Scholarship», *Northwestern University Law Review*, 82(3): 705-762.
- ALLEN, R.J. and PARDO, M.S., 2007: «The Problematic Value of Mathematical Models of Evidence», *The Journal of Legal Studies*, 36(1): 107-140.
- 2019: «Relative Plausibility and Its Critics», *The International Journal of Evidence and Proof*, 23(1-2): 5-59.
- ALLEN, R.J. and STEIN, A., 2013: «Evidence, Probability, and the Burden of Proof», *Arizona Law Review*, 55: 557-602.
- ANDERSON, T., SCHUM, D. & TWINING, W., 2005: *Analysis of Evidence*, 2nd ed., Cambridge: Cambridge University Press.
- BARZUN, C., 2008: «Rules of Weight», *Notre Dame Law Review*, 83(5): 1957-2017.

⁵⁴ See CUMIZ, 2016 for a discussion on BARD and a reference to the other standards of proof used by the ICC for different procedural purposes. But see also, for a critical account of international criminal trials, COMBS, 2010. Cf. JACKSON & SUMMERS, 2012: 108 ff and GASKINS, 2016.

- Bayón, J.C., 2008: «Epistemología, moral y prueba de los hechos: hacia un enfoque no benthamiano», *Analisi e diritto*, 2008: 15-34.
- BERGEAUD, A., 2010: *Le droit à la preuve*, Paris: L.G.D.J.
- BELL, 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: 557-585.
- BOHLANDER, M., 2012: *Principles of German Criminal Procedure*, Oxford: Hart Publishing.
- BREDIN, J.-D., 1996: «Le doute et l'intime conviction», *Droits*, 23: 21-29.
- BOULOC, B., 2004: «La preuve en matière pénale», in PUIGELIER, C. (ed.), *La preuve*, Paris: Economica: 43-56.
- BOVINO, A., 2005: «Evidential Issues before the Inter-American Court of Human Rights», *Sur. International Journal on Human Rights*, 2: 56-79.
- CAPRIOLI, F., 2009: «L'accertamento della responsabilità penale "oltre ogni ragionevole dubbio"», *Rivista italiana di diritto e procedura penale*, 52: 51-92.
- CARTIER, M.-E., 2004: «Brèves remarques sur la preuve devant la Cour pénale internationale», in PUIGELIER, C. (ed.), *La preuve*, Paris: Economica: 57-72.
- CASTILLO DE LA TORRE, F., 2009: «Evidence, Proof and Judicial Review in Cartel Cases», *World Competition*, 32(4): 505-578.
- CLERMONT, K. & SHERWIN, E., 2002: «A Comparative View of Standards of Proof», *The American Journal of Comparative Law*, 50: 243-275.
- COHEN, L.J., 1977: *The Probable and the Provable*, Oxford: Clarendon Press.
- COLOMA, R., 2012: «¿Realmente importa la sana crítica?», *Revista Chilena de Derecho*, 39(3): 753-781.
- COLOMA, R. and AGÜERO, C., 2014: «Lógica, ciencia y experiencia en la valoración de la prueba», *Revista Chilena de Derecho*, 41: 673-703.
- COMBS, N.A., 2010: *Fact-Finding Without Facts. The Uncertain Evidentiary Foundations of International Criminal Convictions*, Cambridge: Cambridge University Press.
- CUMIZ, J.A., 2016: «El estándar de prueba "más allá de toda duda razonable" en la jurisprudencia de la Corte Penal Internacional», *Revista de Derecho Penal y Procesal Penal*, 2016/1: 5-21.
- CUMMINS, R., 1975: «Functional Analysis», *The Journal of Philosophy*, 72: 741-765.
- DAHLMAN, C., 2015: «The Felony Fallacy», *Law, Probability and Risk*, 14: 229-241.
- DAMAŠKA, M., 1986: *The Faces of Justice and State Authority. A Comparative Approach to the Legal Process*, New Haven and Chicago: Yale University Press.
- 1997a: *Evidence Law Adrift*, New Haven & London: Yale University Press.
- 1997b: «The Uncertain Fate of Evidentiary Transplants: Anglo-American and Continental Experiments», *The American Journal of Comparative Law*, 45: 839-852.
- 2003: «Epistemology and Legal Regulation of Proof», *Law, Probability and Risk*, 2: 117-130.
- 2018: *Evaluation of Evidence. Pre-Modern and Modern Approaches*, Cambridge: Cambridge University Press.
- DELLA TORRE, J., 2014: «Il lungo cammino della giurisprudenza italiana sull'«oltre ogni ragionevole dubbio"», *Diritto Penale Contemporaneo* (<https://www.penalecontemporaneo.it>), June 20, 2014.
- DELMAS-MARTY, M., 1996: «La preuve pénale», *Droits*, 23: 53-65.
- EHRENBERG, K.M., 2016: *The Functions of Law*, Oxford: Oxford University Press.
- FERRAJOLI, L., 1989: *Diritto e ragione. Teoria del garantismo penale*, Roma-Bari: Laterza.
- FERRER, J., 2007: *La valoración racional de la prueba*, Madrid: Marcial Pons.
- FODOR, J., 1985: «Fodor's Guide to Mental Representation», *Mind*, 94: 76-100.
- FRIEDMAN, L.M., 1975: *The Legal System. A Social Science Perspective*, New York: Russell Sage.
- GASCÓN, M., 2010: *Los hechos en el derecho. Bases argumentales de la prueba*, 3rd ed., Madrid: Marcial Pons.
- GASKINS, R., 2016: «The Legal Characterization of Facts at the International Criminal Court», in FETERIS, E. et al. (eds.), *Legal Argumentation and the Rule of Law*, The Hague: Eleven International Publishing: 59-68.
- GOLDMAN, A., 2002: «Quasi-Objective Bayesianism and Legal Evidence», *Jurimetrics*, 42: 237-260.

- GONZÁLEZ, J., 2006: «La fundamentación de las sentencias y la sana crítica», *Revista Chilena de Derecho*, 33(1): 93-107.
- GONZÁLEZ LAGIER, D., 2013: *Quaestio facti: Ensayos sobre prueba, causalidad y acción*, Mexico: Fontamara.
- GREEN, L., 1998: «The Functions of Law», *Cogito*, 12(2): 117-124.
- HAACK, S., 2014: *Evidence Matters. Science, Proof, and Truth in the Law*, Cambridge: Cambridge University Press.
- HANS, V. & GERMAIN, C., 2011: «The French Jury at a Crossroads», *Chicago-Kent Law Review*, 86(2): 737-768.
- IACOVIELLO, F.M., 2006: «Lo standard probatorio dell'al di là di ogni ragionevole dubbio e il suo controllo in Cassazione», *Cassazione penale*, 46: 3869-3884.
- JACKSON, J.D. & SUMMERS, S.J., 2012: *The Internationalisation of Criminal Evidence. Beyond the Common Law and Civil Law Traditions*, Cambridge: Cambridge University Press.
- KAPLOW, L., 2012: «Burden of Proof», *The Yale Law Journal*, 121: 738-859.
- LANGBEIN, J.H., 1985: «The German Advantage in Civil Procedure», *The University of Chicago Law Review*, 52(4): 823-866.
- LANGBEIN, J.H., 1988: «Trashing “The German Advantage”», *Northwestern University Law Review*, 82(3): 763-784.
- LAUDAN, L., 2005: «Por qué un estándar de prueba subjetivo y ambiguo no es un estándar», *Doxa*, 28: 95-113.
- 2006: *Truth, Error, and Criminal Law. An Essay in Legal Epistemology*, Cambridge: Cambridge University Press.
- LILLQUIST, E., 2002: «Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability», *U. C. Davis Law Review*, 36: 85-197.
- LLUCH, X.A., 2012: «La dosis de la prueba: entre el *common law* y el *civil law*», *Doxa*, 35: 173-200.
- MACKOR, A.R., 2016: «Improving Judicial Argumentation about Evidence in Dutch Unus Testis Cases», in FETERIS, E. et al. (eds.), *Legal Argumentation and the Rule of Law*, The Hague: Eleven International Publishing: 111-122.
- NOVAK, M., 2016: «Arguing for Certainty in Criminal Decision-Making and the Rule of Law», in FETERIS, E. et al. (eds.), *Legal Argumentation and the Rule of Law*, The Hague: Eleven International Publishing: 147-157.
- PADOA-SCHIOPPA, A. (ed.), 1987: *The Trial Jury in England, France, Germany 1700-1900*, Berlin: Duncker & Humblot.
- 1994, *La giuria penale in Francia*, Milano: LED.
- 2003, *Italia ed Europa nella storia del diritto*, Bologna: il Mulino.
- PARSONS, T., 1951: *The Social System*, Glencoe: The Free Press.
- PAÚL, Á., 2012: «In Search of the Standards of Proof Applied by the Inter-American Court of Human Rights», *Revista Instituto Interamericano de Derechos Humanos*, 55: 57-102.
- PICINALI, F., 2013: «Two Meanings of “Reasonableness”: Dispelling the “Floating” Reasonable Doubt», *The Modern Law Review*, 76(5): 845-875.
- PUTNAM, H., 1981: *Reason, Truth and History*, Cambridge: Cambridge University Press.
- 1995: «Are Moral and Legal Values Made or Discovered?», *Legal Theory*, 1: 5-19.
- REDMAYNE, M., 1999: «Standards of Proof in Civil Litigation», *The Modern Law Review*, 62: 167-195.
- 2003: «Objective Probability and the Assessment of Evidence», *Law, Probability and Risk*, 2: 275-294.
- ROBERTS, P. & ZUCKERMAN, A., 2010: *Criminal Evidence*, 2nd ed., Oxford: Oxford University Press.
- SCHWEIZER, M., 2014: «Comparing Holistic and Atomistic Evaluation of Evidence», *Law, Probability and Risk*, 13: 65-89.
- 2016: «The Civil Standard of Proof—What Is It, Actually?», *The International Journal of Evidence and Proof*, 20(3): 217-234.
- STEIN, A., 2005: *Foundations of Evidence Law*, Oxford: Oxford University Press.
- TARUFFO, M., 1992: *La prova dei fatti giuridici*, Milano: Giuffrè.

- 2003a: «Rethinking the Standards of Proof», *The American Journal of Comparative Law*, 51: 659-677.
- 2003b: «Algunos comentarios sobre la valoración de la prueba», *Discusiones*, 3: 81-97.
- 2009: *La semplice verità. Il giudice e la costruzione dei fatti*, Roma-Bari: Laterza.
- TILLERS, P. & GREEN, E.D. (eds.), 1988: *Probability and Inference in the Law of Evidence*, Dordrecht: Kluwer Academic Publishers.
- TRIBE, L.H., 1971: «Trial by Mathematics. Precision and Ritual in the Legal Process», *Harvard Law Review*, 84(6): 1329-1393.
- TUZET, G., 2020: «Assessment Criteria or Standards of Proof? An Effort in Clarification», *Artificial Intelligence and Law*, 28(1): 91-109.
- UBERTIS, G., 2013: «Razionalismo processuale e verità», *Il giusto processo civile*, 1/13: 325-333.
- VÁZQUEZ, C. (ed.), 2013: *Estándares de prueba y prueba científica: ensayos de epistemología jurídica*, Madrid: Marcial Pons.
- VERHEIJ, B. et al., 2016: «Arguments, Scenarios and Probabilities: Connections Between Three Normative Frameworks for Evidential Reasoning», *Law, Probability and Risk*, 15: 35-70.
- WHITMAN, J., 2008: *The Origins of Reasonable Doubt. Theological Roots of the Criminal Trial*, New Haven: Yale University Press.
- WRIGHT, L., 1973: «Functions», *The Philosophical Review*, 82: 139-168.

REPENSANDO LAS MÁXIMAS DE EXPERIENCIA*

Alan Limardo

Universidad de Buenos Aires

alanlimardo@derecho.uba.ar

RESUMEN: El trabajo se propone realizar un análisis del concepto subyacente al término «máximas de experiencia». En primer lugar, se intenta establecer el estado actual de la comprensión del término por parte de la doctrina. A partir de ello, se da cuenta de la diversidad de conceptos que se manejan y se propone evaluar la estructura lógica de los enunciados a los que se suele catalogar como máximas de experiencia y el rol que cumplen en la valoración de la prueba. A continuación, el texto busca explorar el reemplazo del término por el concepto de generalizaciones y escoge una clasificación de estas. Finalmente, se advierte acerca de algunos peligros en su utilización y se proponen ciertos criterios para su empleo en los razonamientos sobre hechos.

PALABRAS CLAVE: máximas de experiencia; generalizaciones; razonamiento sobre hechos; valoración de la prueba; sana crítica racional.

* Agradezco a Daniel González Lagier las observaciones y comentarios que me realizara a las sucesivas versiones de este trabajo, y a Edgar Aguilera García por las críticas y recomendaciones efectuadas al integrar el tribunal —junto con González Lagier— que evaluó una versión previa de este artículo presentada como Trabajo Final del Máster en Razonamiento Probatorio de la Universidad de Girona. También deseo expresar mi gratitud, por la lectura del texto y las sugerencias que me propusieran —que enriquecieron estas líneas—, hacia Julieta Beutler, Pablo Larsen, Damián Radiminski, Pablo Rovatti, Agustín Varela, Diana Veleda y dos revisores/as anónimos/as que evaluaron la versión final. Por último, quisiera dedicarle este trabajo a Romeo.

RETHINKING «MAXIMS OF EXPERIENCE»

ABSTRACT: This study aims to analyze the concept beyond the term «maxims of experience». First, it attempts to settle how the term is currently understood by legal scholars. Secondly, it finds that a wide variety of concepts are being used and, consequently, examines the logical structure of the statements that are usually classified as maxims of experience and their role in evidence assessment. Afterwards, the essay seeks to explore if the term could be replaced by the concept of generalizations and choose a classification of these. Finally, the work warns about some risks in their use and some criteria are proposed for their application to legal fact reasoning.

KEYWORDS: generalizations; legal fact reasoning; evidence assessment; experience; stock of knowledge.

SUMARIO: 1. INTRODUCCIÓN.— 2. LAS MÁXIMAS DE EXPERIENCIA: ENTRE LA SANA CRÍTICA, LOS CONOCIMIENTOS CIENTÍFICOS Y LA EXPERIENCIA DEL JUEZ O LAS COMUNIDADES.— 3. LA ESTRUCTURA LÓGICA DE LOS ENUNCIADOS Y SU ROL EN LA VALORACIÓN DE LA PRUEBA. EL CONCEPTO DE GENERALIZACIONES EMPÍRICAS: 3.1. ¿Cuál es la estructura lógica de los enunciados que, bajo diversas clasificaciones, llamamos máximas de experiencia? — 3.2. ¿Cuál es su rol en la valoración de la prueba? — 3.3. El concepto de generalizaciones empíricas.— 3.4. El aparato conceptual de Schauer sobre las generalizaciones empíricas.— 4. Los peligros de las generalizaciones: invalidez de la generalización y peligros en el uso de las generalizaciones: 4.1. Invalidez de la generalización.— 4.2. Peligros en el uso de las generalizaciones: 4.2.1. La solidez de las generalizaciones.— 4.2.2. La selección de las generalizaciones: 4.2.2.1. Las generalizaciones específicas del caso y de contexto vs las generalizaciones «ad-hoc».— 4.2.2.2. Las fuentes de las generalizaciones.— 5. ALGUNOS CRITERIOS PARA LA UTILIZACIÓN DE LAS GENERALIZACIONES.— 6. CONCLUSIONES.

RECOMMENDED CITATION: LIMARDO A., 2020: «Repensando las máximas de experiencia», in *Quaestio facti*, 2: 115-153. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22464

“Querer lograr la absoluta exclusión de la duda, es algo que se prohíbe a sí mismo toda persona que haya seguido la historia del saber humano”
Friedrich STEIN (1893: 36)

1. INTRODUCCIÓN

Este trabajo busca comenzar a efectuar un análisis del concepto subyacente al término «máximas de experiencia».

Si existiese un concepto claro, preciso, bien definido y útil, toda esta empresa no tendría mucho sentido, ya sea teórico o práctico. Sin embargo, considero que, si bien se suele utilizar y repetir con frecuencia el término en los razonamientos sobre hechos, podría integrar ese conjunto oscuro de nociones respecto de las cuales no se sabe muy bien a qué se hace referencia con exactitud o, al menos, de los que no se repara muy rigurosamente en sus características y notas definitorias¹.

Por esa razón, aquí se intentará arrojar algo de luz sobre el concepto para iniciar un proceso de delimitación más precisa y, a su vez, una utilización más fructífera en el campo del razonamiento probatorio, lo que podría llevar, al final del día, a su reemplazo.

Creo que la importancia de la tarea puede apreciarse si se atiende a la trascendencia que posee la utilización de las «máximas de experiencia» en las decisiones sobre hechos de los procesos judiciales. Si, como se verá en profundidad más adelante, nos valemos de estas máximas para formular inferencias fácticas a partir de ciertas pruebas y pretendemos hacer valer sus conclusiones como coincidentes con la realidad², parece que su selección, formulación y utilización constituyen pasos importantes en el marco de un procedimiento preocupado por aproximarse en la mayor medida posible a la verdad de los hechos.

Piénsese en el siguiente caso. La policía es alertada por la detonación de un arma de fuego en una casa. Al llegar al lugar, los oficiales ven a un hombre salir del domicilio y comenzar a correr. Esta persona es detenida e identificada como la pareja de la víctima, quien se encuentra muerta en el interior producto de un disparo. Supóngase que a partir de esta limitada información (el sujeto fue detenido cuando salía corriendo del lugar del asesinato) debe formularse una inferencia sobre la responsabilidad, o no, de la persona detenida.

La fiscalía podría recurrir, para sostener la hipótesis de culpabilidad de la persona, a un enunciado que indique que «las personas que huyen de la escena del crimen suelen ser sus autores». La defensa, en sentido contrario, tal vez alegue que su cliente no es el autor, pues «las personas que cometen un asesinato, y son detenidas mientras huyen, habitualmente tienen en su poder el arma homicida». La primera inferencia se orienta a la culpabilidad del acusado, mientras que la segunda a su inocencia. La diferencia, *ceteris paribus*, radica en la «máxima de experiencia» empleada.

¹ Podría compartir este conjunto oscuro de nociones con, por ejemplo y sin ánimo de exhaustividad, «más allá de toda duda razonable», «reglas de la sana crítica», y las categorías «acusatorio» o «inquisitivo» como grandes etiquetas.

² Este trabajo, conviene adelantar desde ahora, toma como punto de partida las tesis centrales de la *tradición o concepción racionalista de la prueba*. Sobre el contenido de esas tesis, en especial sobre la prueba jurídica y los presupuestos filosóficos y epistemológicos, puede verse una excelente presentación en ACCATINO, 2019. También puede consultarse un resumen, crítico de algunos presupuestos expresados por FERRER BELTRÁN —en particular sobre estándares de suficiencia probatoria—, en la obra de AGUILERA, 2016.

Ahora bien, ¿de dónde surge el enunciado que indica que las personas que huyen de la escena del crimen suelen ser sus autores? ¿De qué modo podríamos justificar su formulación en un proceso judicial? Es decir, ¿Con qué criterios o parámetros contamos para diferenciar las «máximas de experiencia» de otro tipo de enunciados? ¿Son acaso aquellos mencionados en el párrafo anterior verdaderas máximas de experiencia? A su vez, ¿cómo podríamos defender racionalmente nuestra preferencia por la primera o por la segunda de ellas? Y como si estos interrogantes no fueran suficientes, ¿en qué grado esos enunciados apoyan la conclusión a la que queremos arribar?

Todo esto demuestra que poner el foco de atención en aquello que se ha denominado máximas de experiencia es, en gran medida, preocuparnos por una valoración racional de la prueba y una reconstrucción lo más precisa posible de los hechos. En términos más concretos, una o varias máximas de experiencia pueden llegar a constituir la diferencia entre una persona culpable que es dejada en libertad o una inocente que es condenada erróneamente.

Si bien un análisis exhaustivo del asunto llevaría a abordar una gran cantidad de aspectos que guardan vinculación con el tema escogido —directa o indirectamente—, solo concentraré mi atención en la función que las «máximas de experiencia» poseen en el momento de la valoración de los medios de prueba, algo que podríamos denominar su rol *justificativo*³.

El trabajo se estructura de la siguiente manera. En primer lugar, examinaré el modo en que la doctrina de la tradición continental-europea ha entendido a las máximas de experiencia, para demostrar las profundas diferencias en la comprensión del término y, por consiguiente, la imprecisión conceptual que existe (2). A continuación, a partir de un análisis de la estructura lógica de los enunciados a los cuales típicamente se ha calificado como máximas de experiencia y de su rol en la valoración de la prueba, propondré como fructífero su reemplazo por el concepto de *generalizaciones empíricas* (3). Luego, ingresaré en la tarea de señalar algunos criterios para determinar cuándo una generalización puede ser empleada en un razonamiento sobre hechos, con el objetivo de evitar, reducir o ser conscientes al menos de ciertos peligros inherentes a las generalizaciones y otros vinculados con su utilización (4). Finalmente, intentaré enunciar algunos criterios para afrontar esos problemas (5).

³ Como dije al inicio, se trata de comenzar a explorar un camino y trazar algunas líneas de trabajo. Quedarán por fuera, al menos de un análisis detallado, otros puntos respecto de los cuales se podrían alegar interrelaciones concretas, tales como las presunciones, la denominada prueba «por indicios», los hechos notorios, el rol de las narrativas o historias (*story-telling*), el razonamiento abductivo, la función heurística de las generalizaciones, otras herramientas (por ejemplo, definiciones o conceptos) para formular inferencias, la prueba de la relación de causalidad o el control de la valoración en etapas recursivas.

2. LAS MÁXIMAS DE EXPERIENCIA: ENTRE LA SANA CRÍTICA, LOS CONOCIMIENTOS CIENTÍFICOS Y LA EXPERIENCIA DEL JUEZ O LAS COMUNIDADES

El origen del término «máximas de experiencia» suele atribuirse a Friedrich STEIN (1893)⁴, quien a finales del siglo XIX ya destacó su falta de atención teórica⁵. En aquel entonces, STEIN afirmó que las máximas de experiencia⁶:

Son definiciones o juicios hipotéticos de contenido general, desligados de los hechos concretos que se juzgan en el proceso, procedentes de la experiencia, pero independientes de los casos particulares de cuya observación se han inducido y que, por encima de esos casos, pretenden tener validez para otros nuevos (1893: 27).

A lo largo de los años subsiguientes, la doctrina procesal de la tradición europeo-continental levantó el guante dejado por STEIN⁷ y comenzó a emplear la denominación que él acuñó⁸. Sin embargo, bajo ese término se han ofrecido definiciones de lo más disímiles. Algunas de ellas se señalarán a continuación para intentar demostrar la imprecisión conceptual que afecta al término hoy en día.

En la doctrina española se ha considerado, en el marco de una concepción silogística-deductiva del razonamiento judicial —tal como la sostenida por STEIN⁹—, que las máximas de experiencia constituyen su premisa mayor en materia de hechos y permiten valorar los distintos medios de prueba para llegar a conclusiones fácticas:

La valoración de la prueba es una operación mental que se resuelve en un silogismo en el que:
1) La premisa menor es la fuente-medio de prueba (el testigo y su declaración, el documento y

⁴ Así, entre otros, TARUFFO (2002: 270 y 2009: 439), NIEVA FENOLL (2010: 28) y UBERTIS (2017: 75).

⁵ El autor afirma en el prólogo de su obra: «Las cuestiones que van a ser tratadas en las páginas siguientes pertenecen en su mayor parte a la vida jurídica de todos los días: tal vez sea esa precisamente la causa por la que la mayor parte de ellas nunca hasta ahora han sido objeto de trabajo científico» (1893: 1).

⁶ El término empleado no le resultaba del todo convincente: «Nos quedaremos pues, por ahora, con este último nombre. No está libre de objeciones: lo sé y estoy dispuesto a sustituirlo por otro mejor que me sea propuesto» (STEIN, 1893: 19).

⁷ STEIN no profundizó particularmente sobre los rasgos definitorios y términos de la definición ofrecida, de manera que ello tal vez incidió en el desarrollo histórico de diversas interpretaciones conceptuales. Ciertamente, de su noción de máximas de experiencia se desprende con claridad que constituyen la premisa mayor de los razonamientos sobre hechos en una concepción silogística del juicio de hecho—de manera que sirven para valorar los medios de prueba—, poseen carácter fáctico, no son declaraciones sobre acontecimientos individuales sino el producto de la observación de numerosos casos, y que no existiría mayor diferencia entre los conocimientos científicos —que ingresan al proceso mediante los peritos— y las máximas de experiencia (1893: 15, 23-24, 34, 41-44, 59 y 67).

⁸ Tal vez por la influencia de la *intime conviction* en el modo de concebir la valoración de la prueba, TARUFFO señala que no se encuentran rasgos del término en la cultura jurídica francesa (2009: 439).

⁹ La concepción silogística-deductiva del juicio de hecho sostenida por STEIN es también destacada por TARUFFO (2009: 439).

su presentación), 2) La premisa mayor es una máxima de la experiencia, y 3) La conclusión es la afirmación de la existencia o de la inexistencia del hecho que se pretendía probar (MONTERO AROCA, 2008: 437).

En este contexto, se afirmó que las máximas de experiencia son el *equivalente* a las reglas de la sana crítica y se obtienen de la experiencia vital de cada juez:

En el sistema de prueba libre las máximas de la experiencia se aplican por el juez, se extraen por este de su experiencia de la vida, y se deben aplicar a cada uno de los medios de prueba. No se trata de que el juez decida en conciencia sino de [que] aplique las [...] reglas de la sana crítica y estas reglas son simplemente máximas de experiencia (MONTERO AROCA, 2008: 438)¹⁰.

La visión de las máximas de experiencia como premisa mayor de un razonamiento silogístico-deductivo que permite avanzar a la conclusión sobre los hechos no es aislada. Así se ha expresado que:

conforme al modelo silogista, la «máxima de experiencia» vendría a ser la premisa mayor —a menudo implícita— del susodicho silogismo; el «dato probatorio» disponible funcionaría como la premisa menor; y la conclusión constituiría el «hecho probado» (IGARTUA SALAVERRÍA, 2004: 116).

A pesar de esa coincidencia parcial¹¹ sobre su ubicación sistemática, existe una discrepancia respecto a su origen. Pues mientras anteriormente se indicó que encuentran su génesis en la experiencia vital del juez, aquí se señala que «[l]as máximas de experiencia son *reglas* que se extraen observando *numerosos y similares* casos reales» (IGARTUA SALAVERRÍA, 2004: 117).

Otros autores de la doctrina española, si bien sitúan a las máximas de experiencia como parte de las «reglas del criterio racional» para valorar la prueba, i) no las consideran completamente *equivalentes*, sino como un componente más de aquellas y, a su vez, ii) efectúan una ulterior distinción entre las máximas «comunes» y aquellas «especializadas», que ingresarían al proceso a través de los peritos:

Libre valoración de la prueba no significa discrecionalidad o arbitrariedad, sino que la valoración de la prueba se ha efectuar, [...], «según las reglas del criterio racional», es decir, según las normas de la lógica, las máximas de experiencia común o de las especializadas que proponen los peritos (DE LA OLIVA SANTOS ET. AL., 2007: 486).

Por su parte, GONZÁLEZ LAGIER (2013: 43-44) —sobre cuya posición se volverá— realiza una clasificación más amplia y distingue entre aquellas de carácter «científico o especializado» (como las que aportan los peritos), «jurídico» (derivadas del ejercicio profesional del juez) y «privado» (provenientes de la experiencia del juez al margen de su profesión).

También es posible advertir cómo algún autor coincide con la circunstancia de que la premisa mayor del razonamiento judicial sobre hechos se trata del resultado de

¹⁰ LLUCH (2012: 415) destaca que también sostiene esa posición SERRA DOMÍNGUEZ.

¹¹ La coincidencia es parcial pues el último autor citado, si bien considera que las máximas de experiencia integran la premisa mayor del razonamiento sobre hechos, no postula que sean equivalentes a las reglas de la sana crítica.

una inducción —algo deslizado ya por STEIN e insinuado por IGARTUA SALAVERRÍA—, producto de la «cultura» del juez y de su «contacto» con la realidad:

En el silogismo clásico, es decir, el silogismo deductivo, la premisa mayor presupone un razonamiento inductivo basado en la experiencia. Cuando el juez, por las reacciones emocionales del testigo infiere que este no es veraz, [...] es evidente que sobre el cauce formal silogístico se halla aplicando anteriores criterios empíricos adquiridos merced a su cultura y su continuado contacto con la realidad. De ahí que afirmemos que al lado de la estructura lógica existe asimismo un fundamento experimental (MUÑOZ SABATÉ, 1967: 174-175).

Y respecto de su función, es claro el autor citado en punto a que posibilitan avanzar desde la información aportada por los medios de prueba hacia la conclusión del juez, ya que hacen factible que este último subsuma el supuesto de hecho con el que se enfrenta en una categoría mayor, más general, a partir de las experiencias previas:

Representa en el campo más reducido del derecho probatorio el magno problema de los universales, de la generalización la experiencia, del paso de lo particular a lo universal. Aunque no consistan en una norma abstracta aplicable, sin más, al caso concreto, las máximas de experiencia contribuyen de un modo eficaz a la formación de la persuasión judicial en tanto en cuanto el juez se aperciba de que aquel supuesto concreto entra en la categoría de acontecimientos que han dado origen a la máxima aludida (MUÑOZ SABATÉ, 1967: 175).

Asimismo, en la doctrina italiana se ha considerado que se tratan de:

enunciaciones de tipo general obtenidas a partir de la observación de sucesos pasados, susceptibles de ser formuladas por «cualquier persona de buen entendimiento y mediana cultura» (UBERTIS, 2017: 75, con cita de CHIOVENDA, 1977: 575).

En la medida en que pueden ser formuladas por cualquier persona, se diferenciarían de las leyes científicas pues estas últimas poseen:

la peculiaridad [...] de hallarse inscritas (y justificadas por) un sistema de asertos generalmente aceptado, a diferencia de las máximas de experiencia que se consideran existentes o no cada una con independencia de las demás (UBERTIS, 2017: 82).

Por su parte, TARUFFO (2002: 219) señala que son nociones derivadas de la experiencia común que representan la base de conocimientos generales para la valoración de la prueba, pero que «esos conocimientos expresan nociones de sentido común que tienen como único fundamento el hecho de formar parte de la cultura del hombre medio en un cierto lugar y en un cierto momento». A continuación, agrega:

unas veces, en efecto, son vulgarizaciones de leyes lógicas o naturales, de modo que existe —al menos en teoría— la posibilidad de fundar la valoración de las pruebas sobre esas leyes. Más a menudo, en cambio, estas bases no existen y las máximas de experiencia expresan únicamente toscas generalizaciones, tendencias genéricas, opiniones o prejuicios difundidos, en cuya base está la cultura del sentido común sin convalidación o confirmación alguna de tipo científico (2002: 219).

En otro trabajo, puntualiza que aquello prototípico de las máximas de experiencia es que consisten en enunciados que expresan una regularidad sobre acontecimientos —de modo similar a IGARTUA SALAVERRÍA, MUÑOZ SABATÉ y UBERTIS— y que se manifiestan en forma de reglas, pues se refieren a una:

pluralidad de hechos o de comportamientos cuyo conocimiento se supone derivado de la experiencia que se tiene a partir de estos hechos o comportamientos, y enuncian la que parece ser una *regularidad* en la ocurrencia de los mismos (2009: 440).

Finalmente, retomando la primera de las definiciones, precisa que se tratan de conocimientos generales, derivados de la experiencia «común» y que su función consiste en posibilitar la valoración de los elementos de prueba:

conocimientos generales vinculados con el hecho a probar y útiles, de alguna manera, a los efectos de su determinación; es más, sin estos conocimientos la valoración de la prueba sería normalmente imposible. Se trata de las nociones derivadas de la experiencia común que encuentran su formulación en las denominadas máximas de experiencia (las *Erfahrungssätze* de la doctrina alemana) y que desarrollan un papel relevante en la valoración de las pruebas. Estas nociones representan, pues, la base de conocimientos «generales» que sirven para la valoración de las pruebas y determinan en gran medida su resultado (2012: 219).

En latitudes latinoamericanas, es posible encontrar que, en la doctrina colombiana, las máximas o reglas de la experiencia se tratarían de «simples normas de criterio para el entendimiento de los hechos» (DEVIS ECHANDÍA, 1981: 176), o bien que se utilizarían por los jueces para valorar la prueba y consistirían, simplemente, en la *experiencia que todo hombre posee*:

eso que aprendió y que acumuló para ser empleado en nuevas situaciones. En otras palabras, lo que llamamos en el mundo del proceso reglas de la experiencia, no es más que una aplicación en concreto de la experiencia que todo hombre posee (PARRA QUIJANO, 2007: 80).

Por su parte, en la doctrina rioplatense se pueden observar también conceptos discordantes.

Algunas personas autoras han considerado que las máximas de experiencia son «conocimientos pacíficamente incorporados al patrimonio cultural de un círculo social determinado, y se hallan por ello al margen del objeto probatorio» —en otras palabras, no deberían ser probadas en el marco del proceso— y que «*entrañan principios generales* extraídos de la observación de los fenómenos físicos o del corriente comportamiento de los hombres», como también «juicios adquiridos a raíz de conocimientos técnicos especiales, en cuyo supuesto aquellas son transmitidas al juez por los peritos», y se utilizan al momento de valorar la prueba (PALACIO, 2000: 21). Aquí vemos como, una vez más, implícitamente se mantiene la distinción entre máximas «comunes» y «especializadas», se las considera el resultado de una inducción y se las vincula, a su vez, a un círculo o grupo social particular.

Otros han destacado que la valoración de las pruebas debe realizarse de acuerdo a las reglas de la sana crítica, que estarían integradas por la lógica, la psicología y la experiencia común (VELEZ MARICONDE, 1982: 363), o a «las reglas que orientan el recto pensamiento humano», «reglas que permiten discernir lo verdadero de lo falso», que resultarían ser la lógica, los principios de las ciencias y (una vez más) la experiencia común (CAFFERATA NORES, 1998:7), y que esta última se encontraría constituida por «conocimientos vulgares indiscutibles por su raíz científica; v.gr., inercia, grave-

dad» (CAFFERATA NORES, 1998:46), es decir, estaría —una vez más— la experiencia en una relación de género-especie con los principios científicos.

Por otra parte, también se ha considerado que las máximas de experiencia son: «normas de valor general, independientes del caso específico, pero que, extraídas de cuanto ocurre generalmente en múltiples casos, pueden aplicarse en todos los otros casos de la misma especie» (COUTURE, 1958: 229-230), utilizadas junto con los principios de la lógica para la valoración de la prueba, pues:

La sana crítica es, además de lógica, la correcta apreciación de ciertas proposiciones de experiencia de que todo hombre se sirve en la vida. Esas conclusiones no tienen la estrictez de los principios lógicos tradicionales, sino que son contingentes y variables con relación al tiempo y al lugar. El progreso de la ciencia está hecho de una serie de máximas de experiencia derogadas por otras más exactas (COUTURE, 1958: 272).

En la doctrina chilena, HORVITZ LENNON y LÓPEZ MASLE entienden que integran la sana crítica racional junto con los principios de la lógica, pero que a diferencia de ellos, las máximas de experiencia serían variables en función del tiempo y el lugar (2004: 149) y que «por definición, expresan nociones de sentido común cuyo único fundamento es el hecho de formar parte de la cultura del hombre medio en un cierto lugar y en un cierto momento» (2004: 336), aunque en la nota al pie aclaran, siguiendo a TARUFFO, que en algunos casos pueden constituir la vulgarización de leyes lógicas o naturales.

Se podría continuar con la enumeración de definiciones o menciones realizadas por diversos doctrinarios y doctrinarias acerca de las máximas de experiencia, pero por cuestiones de extensión —y para evitar el tedio del lector—, se tomará lo dicho hasta aquí como un panorama legítimo sobre el estado de la cuestión¹².

Pues bien, a partir del relevamiento efectuado de la doctrina de la tradición jurídica europeo-continental es posible afirmar que el panorama es desalentador para tratar de comprender los contornos y componentes del concepto que subyace al término «máximas de experiencia» o, lo que es igual, contar con parámetros útiles para calificar válidamente a un enunciado como una de ellas.

¹² Vale aclarar que se omitió conscientemente un relevamiento acerca de aquello que la jurisprudencia ha venido sosteniendo sobre el punto. Esto se debe a un doble orden de razones. Por un lado, un criterio de autoridad impide efectuar una reseña que pueda presentarse como parámetro válido: usualmente los máximos tribunales nacionales o las altas cortes internacionales —aquellos que podrían ser utilizados como *referencias de peso*, tanto por sus consideraciones como por su influencia jerárquica—, debido a factores vinculados a la estrechez o especificidad de sus recursos, no suelen abordar temas vinculados a la valoración de la prueba en los procesos. En segundo término, y en relación con lo anterior, la alternativa de analizar casos judiciales de diversos países resulta, por un lado, inabarcable para los límites de esta investigación y, a su vez, la utilización de algunas sentencias como criterio de referencia podía generar una falta de homogeneidad profunda de la muestra escogida. Al respecto, podría resultar un ejercicio interesante —para ajustar o profundizar las consideraciones de este trabajo— analizar alguna jurisdicción en particular. Sin embargo, creo que el argumento que pretendo poner en discusión puede sostenerse de manera genuina a partir del estado de la cuestión en la doctrina.

Muchos de los interrogantes formulados al inicio del trabajo encuentran respuestas confusas en este marco. De STEIN para aquí, los intentos doctrinales no han podido ofrecer una definición lo suficientemente clara y completa para mejorar la delimitación del concepto —y así conseguir una utilización más fructífera—; por el contrario, se ha configurado una confusión preocupante sobre lo que las máximas de experiencia son, al punto de encontrar conceptos abiertamente contradictorios. El aspecto inmutable de coincidencia es que se utilizarían, de alguna manera imprecisa, para valorar la prueba. De allí en adelante, los caminos se bifurcan.

Para algunas personas autoras representan el equivalente de las reglas de la sana crítica, para otros tan solo son una parte integrante de aquellas —es decir, un *subconjunto* dentro de ellas—, pasando por aquellos que las consideran como «simples normas de criterio» o ciertos juicios generales, que pueden —o no— implicar conocimientos técnico-científicos especiales. A su vez, podemos hallar definiciones que las equiparan, sin matices, con los conocimientos científicos y, a la par, otras que consideran que existe entre ambos una relación género-especie.

También es posible apreciar que ciertos doctrinarios y doctrinarias ponen el foco en su volatilidad, es decir, en la idea de que no son inmutables ni necesariamente evidentes y que, por el contrario, varían constantemente en función del tiempo y en espacio, debido al nivel cultural y las características de las personas en cuyo ámbito nos encontremos.

En paralelo, otras personas autoras sostienen que debe eximirse la prueba de aquellas, es decir, que no podrían ser objeto de actividad probatoria; esta mirada parece aceptar una transversalidad, una notoriedad —al menos parcial— de cierto núcleo de máximas que integrarían algo así como la experiencia o la cultura de *toda* persona.

No es difícil imaginar las consecuencias que puede generar un panorama como el expuesto. Tómese como punto de partida los siguientes enunciados y piénsese si pueden ser considerados máximas de experiencia, de acuerdo con las definiciones señaladas más arriba:

- 1) «Los testigos capaces y no sospechosos dicen la verdad» (CARNELUTTI, 1982: 63).
- 2) «Quien roba una cosa, ordinariamente, la conserva en su poder» (CAFFERATA NORES, 1998: 193).
- 3) «Las personas no trabajan gratuitamente» (con una formulación ligeramente distinta es mencionada por UBERTIS, 2017: 82, con cita de NOBILI, 1969: 144).
- 4) «Estómago e intestino de un recién nacido se llenan de aire después de aproximadamente seis horas» (ROXIN, 2000: 238).
- 5) «Los actos se inspiran en los móviles que se reflejan en el mundo exterior» (COUTURE, 1958: 232).

Es claro que no habría una coincidencia unánime. Mientras algunas personas autoras solo reconocerían aquellas que poseen una raíz científica (en principio, únicamente el enunciado 4), otros podrían sostener que aquellas que provienen de la experiencia profesional del juez también lo son (1 y 2, por ejemplo), en tanto al-

gunos considerarían las que cuenten con un cierto consenso o transversalidad en la experiencia común (asumamos los enunciados 3 y 5), pero también podría limitar su campo de aplicación en razón del tiempo o el lugar (así, el enunciado 3 podría no ser considerado como tal en alguna comunidad).

A su vez, podrían imaginarse otros enunciados que suelen mencionarse como «máximas», tanto en la vida cotidiana como en los procesos judiciales, en cuyo caso las dificultades parecen ser aun mayores para evaluar su procedencia como partes integrantes de un razonamiento probatorio: «el que calla otorga» (en su versión adaptada al proceso penal, podría ser «las personas que guardan silencio frente a una imputación suelen ser culpables»), «el que se excusa, se acusa» (que es posible traducir judicialmente como «las personas que dan demasiadas explicaciones no son creíbles») o «el hábito no hace al monje» (es decir, algo así como «el aspecto de las personas no es un indicador sobre su confiabilidad»).

Entonces, parece llevar razón TARUFFO cuando considera que:

Cualquier análisis sintético de todo lo que circunda a la noción de «máximas de la experiencia», demuestra que en ella se producen las cosas más diversas y extrañas, frecuentemente privadas de cualquier credibilidad, ya sea por errores, prejuicios, generalizaciones infundadas o simplificaciones indebidas (2012: 250).

Así las cosas, comenzar a recorrer el camino para conseguir una mayor precisión sobre el concepto subyacente al término «máximas de experiencia» parece ser una tarea necesaria si se tiene una genuina preocupación epistémica acerca de los procesos judiciales.

A continuación, se propondrá una manera de iniciar el abordaje de esa tarea.

3. LA ESTRUCTURA LÓGICA DE LOS ENUNCIADOS Y SU ROL EN LA VALORACIÓN DE LA PRUEBA. EL CONCEPTO DE GENERALIZACIONES EMPÍRICAS.

3.1. ¿Cuál es la estructura lógica de los enunciados que, bajo diversas clasificaciones, llamamos máximas de experiencia?

Para responder a la primera pregunta partiré de tres ejemplos. Uno de ellos es aquel presentado en la introducción: *i)* «Las personas que huyen de la escena del crimen suelen ser sus autores». El segundo será: *ii)* «Los testigos que declaran bajo juramento suelen decir la verdad». El último: *iii)* «Los seres humanos no pueden estar al mismo tiempo en dos lugares».

Si se observa con atención, esta clase de enunciados sujetan a la presencia de una condición, una determinada propiedad o característica. Es decir, si se constata la condición *co1)* «persona que huye de la escena del crimen», *co2)* «testigo que declara bajo juramento» o *co3)* «ser humano» o «humano en lugar *X*», se atribuye o

correlaciona la propiedad o característica ca1) “autoría del crimen”, ca2) “veracidad en el relato” y ca3) “imposibilidad de estar en dos lugares al mismo tiempo” o “imposibilidad de estar en un lugar distinto de X”.

De este modo, el antecedente del enunciado estaría constituido por una condición («P») y el consecuente por una propiedad vinculada causalmente a la constatación de ese antecedente («Q»). En otras palabras, podríamos expresar su estructura lógica como «Si P, entonces Q». Los enunciados que presentan esta forma han sido denominados *enunciados condicionales* (COPI y COHEN, 2013: 379; HEMPEL, 1966: 39).

A su vez, antecedente y consecuente del enunciado poseen un referente empírico constatable, y se expresa entre ellos una implicación causal que determina que la conexión entre ellos deba (o pueda) descubrirse empíricamente (COPI y COHEN, 2013: 380)¹³. Es decir, se encuentran integrados por datos o afirmaciones acerca del mundo, que buscan describirlo y explicarlo, y que es posible conocer y verificar. De ese modo, podemos hablar, con mayor precisión, de enunciados fácticos condicionales. Esta circunstancia determina que este tipo de enunciados pueda ser susceptible de verdad o falsedad, en la medida en que se correspondan, o no, con la realidad.

La correlación expresada en un enunciado de esta clase nunca puede ser absoluta, es decir, infalible o exenta de toda excepción, aun a futuro. Esto, en definitiva, es algo inherente al conocimiento humano. Ni siquiera en el caso de la ley de gravedad, que usualmente es mencionada y empleada como una regla absoluta, podemos asegurar que en el futuro nunca nos encontraremos con un caso que la desafíe, por lo que se trata de un enunciado condicional derrotable, en el sentido de hallarse sujeto a alguna excepción. En esa medida, la correlación tendrá siempre un grado de regularidad en la asociación (o, lo que es igual, un «grado de derrotabilidad»), aunque muchas veces implícito por cuestiones de economía del lenguaje, que puede verse expresado tanto en el antecedente como en el consecuente.

De esta manera, la verdad o falsedad de estos enunciados dependerá de la verificación, en primer lugar, de la existencia de la correlación y, a su vez, de la correspondencia entre el grado de intensidad de esa asociación expresado y aquel constatable en la realidad. Así, en el primer ejemplo, la verdad del enunciado «las personas que huyen de la escena del crimen suelen ser sus autores» dependerá de: a) la constatación de un vínculo o correlación entre la huida de la escena del crimen y la autoría del delito, b) la verificación del grado de intensidad de la correlación expresado en el enunciado (en este caso, aquel equivalente a «suelen»).

La omisión de explicitar el grado de la asociación entre antecedente y consecuente en esta clase de enunciados podría explicar por qué razón, en diversas ocasiones, son

¹³ En este trabajo, expresiones como «implicación causal», «correlación causal», etc., deben tomarse en un sentido amplio, es decir, tanto como relaciones de causalidad *stricto sensu* (causa-efecto) como, también, asociaciones regulares que no puedan ser calificadas como tales. Se volverá sobre los tipos de asociaciones en el apartado 4.

expresados como «reglas» pretendidamente capaces de conformar un razonamiento deductivo, es decir, sin posibilidad de ser derrotadas.

Así, en nuestra vida cotidiana, por motivos prácticos (por ejemplo, comunicarnos más rápido con nuestro interlocutor) no aclaramos que «[la mayoría de las/muchas de las/algunas de las] personas que viven en Argentina comen asado» y, en su lugar, directamente afirmamos que «los argentinos comen asado».

Por otra parte, existen casos en los que utilizamos enunciados cuyo tan alto grado de asociación entre antecedente y consecuente nos lleva a omitir precisar que podría existir alguna clase de excepción. Un ejemplo de ello es el tercer enunciado expresado más arriba. Si bien sabemos que ningún conocimiento es absoluto o infalible, no afirmamos diariamente que «los seres humanos [por ahora/hasta ahora no se ha refutado que] no pueden estar al mismo tiempo en dos lugares», lo que implica ocultar su carácter derrotable.

Por último, en los enunciados objeto de análisis, el antecedente siempre se encuentra configurado por una condición *general*, es decir, que representa a una clase (de hechos, de personas, etc.) y no particulariza sobre un integrante puntual de aquella. En esta medida, un patrón común de lo que habitualmente se califica como máximas de experiencia es que la correlación expresada busca explicar y brindar información acerca de un *conjunto* y una propiedad a él presuntamente asociada, es decir, un grupo de casos en abstracto (casos o situaciones tipo), y no consisten, por consiguiente, en observaciones singulares o particularizadas.

Ese conjunto que conforma el antecedente podrá ser más grande o más pequeño («los seres humanos» o «los argentinos», «los automóviles» o «los automóviles automáticos Ford»), pero siempre, como es obvio, conformado por más de un integrante. De esta manera, nuestras intuiciones nos guían a excluir como una máxima o regla a la asignación de una propiedad o característica determinada a la verificación de una condición individual.

De este modo, el enunciado «Juan no es veraz» no es más que una descripción de una característica presente en un sujeto *en particular*, una observación o enunciado singular que, en todo caso, como veremos, podrá ser la conclusión de un razonamiento fáctico, pero nunca podrá integrarlo para *avanzar* en la valoración de la prueba. Sin embargo, sí tendemos a aceptar como una máxima un enunciado del tipo «los familiares del acusado no son veraces», esto es, un enunciado *general* (HERNÁNDEZ MARÍN, 2013: 14) y, *a priori*, nos inclinamos a pensar que tendría algún tipo de utilidad en la tarea de analizar circunstancias concretas de un caso.

A partir de lo dicho hasta aquí, es posible comenzar ya a separar la paja del trigo. Una primera distinción se hace necesaria para alejar una parte de la confusión conceptual. No pueden asimilarse estos enunciados con lo que se denomina —también, en general, de modo difuso— como «reglas de la sana crítica».

Esto requeriría mayores explicaciones, pero bastará con señalar aquí que estas últimas no se vinculan con cierto grado de conocimiento empírico acerca de hechos,

eventos o asociaciones, sino que se trata de reglas que *prescriben* cómo debe valorarse racionalmente la prueba, es decir, qué clase de razonamientos son correctos para tal fin (TARUFFO, 2009: 441) y podrían considerarse como criterios de racionalidad epistemológica, que configuran enunciados de carácter *normativo* (GONZÁLEZ LAGIER, 2018). Por lo tanto, su fundamentación no es empírica —lo que daría lugar a enunciados de carácter descriptivo susceptibles de verdad o falsedad— como es el caso de los enunciados condicionales anteriormente señalados.

Veamos. El enunciado condicional que establece que «los aficionados de Boca Juniors se retiran del estadio diez minutos antes de que termine el encuentro cuando su equipo va perdiendo» es un enunciado condicional (asocia la condición «hincha de Boca Juniors» con la característica «retirarse antes del encuentro cuando su equipo va perdiendo») y puede ser contrastado empíricamente: bastaría detenerse en las salidas del estadio para evaluar si, efectivamente, eso sucede en el mundo. De este modo, puede ser calificado como verdadero o falso. En cambio, el enunciado normativo que determina «cuantos más elementos de juicio a favor de una hipótesis, esta se encuentra corroborada en un grado más alto» no puede ser objeto de ese análisis (GONZÁLEZ LAGIER, 2018)¹⁴.

Similares consideraciones pueden efectuarse para distinguir estos enunciados del conjunto de las reglas o principios de la lógica. En esta dirección, COLOMA CORREA y AGÜERO SAN JUAN señalan que ese conjunto no suministra información sobre el mundo o sobre la realidad, sino que instituye los límites del ejercicio del razonamiento y permite conocer «de antemano, qué movimientos o jugadas argumentativas pueden ser calificadas como correctas, posibles, incorrectas o imposibles» (2014: 682).

Una conclusión parcial con lo dicho hasta aquí es, entonces, que aquello que usualmente se ha denominado como «máximas de experiencia» son enunciados fácticos generales y condicionales.

3.2. ¿Cuál es su rol en la valoración de la prueba?

Ahora bien, uno de los pocos puntos de coincidencia en la reseña efectuada en el apartado 2 es que las «máximas de experiencia» servirían para, de algún modo,

¹⁴ De cualquier manera, podría discutirse si el fundamento último de aquellos enunciados no es de carácter empírico, esto es, del resultado de una inducción integrada por determinados procedimientos que han resultado exitosos para conocer el mundo y que, en consecuencia, ese éxito haya determinado que les otorguemos este carácter. A su vez, podría evaluarse si la actividad epistémica de recolección y/o constatación de los datos empíricos que nos permiten otorgar ese carácter normativo es una tarea posible o realizable. Por último, habría que preguntarse cómo sortear el problema de la circularidad (GONZÁLEZ LAGIER, 2018) y cómo, entonces, sería posible asignarles una calidad normativa. La discusión excede los límites de este trabajo y creo que bastará aquí para los objetivos propuestos con lo señalado más arriba. Agradezco a Daniel GONZÁLEZ LAGIER la observación que me efectuara sobre el punto. En un reciente trabajo, GONZÁLEZ LAGIER (2020) comienza a abordar algunos de los interrogantes mencionados anteriormente, a la luz de la epistemología *fundherentista* de Susan HAACK.

valorar la prueba, es decir, para la tarea de avanzar a conclusiones fácticas a partir de cierta evidencia con la que contamos dentro de un proceso judicial.

Una aclaración importante, en función de la diferencia de estos enunciados con las *reglas de la sana crítica* o *criterios de racionalidad epistemológica* efectuada en el apartado anterior, es que la utilidad que pueden tener en la actividad de valoración de la prueba no puede comprenderse como una de tipo *normativa* o *prescriptiva*; su valor no reside en acercarnos los mejores criterios para aproximarnos a la realidad de los hechos, sino que constituye un medio o un instrumento —un enunciado general descriptivo, con proyección empírica— que debemos emplear de acuerdo con esos criterios epistémicos generales¹⁵.

Para tratar de obtener un poco más de precisión sobre cuál es su rol *específico* en ese complejo entramado de instrumentos que entran en juego en la valoración de la prueba, se partirá de un esquema clásico sobre la estructura de un argumento acerca de hechos, como el propuesto por TOULMIN (2007: 132 y ss.)¹⁶.

De modo simplificado, el autor propone que es posible advertir que en todo razonamiento fáctico se contará, al menos, con tres elementos: 1) un dato o afirmación [“D”], 2) una conclusión a la que pretendemos arribar [“C”] y 3) una *garantía* [“G”], que permitirá legitimar el paso entre «D» y «C».

Resulta sencillo identificar los datos o afirmaciones con la información que surge de los medios de prueba con los que se cuenta en un proceso judicial y a la conclusión con el enunciado fáctico, al que se refiere ese medio de prueba, que pretendemos establecer como verdadero.

Ahora bien, siguiendo a TOULMIN, la garantía de los argumentos consiste en enunciados hipotéticos (o condicionales) de carácter general del tipo «Si D, entonces C» (2007: 134). Si bien la propuesta del autor incluye también *modalizadores* [“M”] y *condiciones de excepción o de refutación* [“E”] de las garantías, como elementos «dis-

¹⁵ Agradezco a un/a revisor/a anónimo/a por permitirme notar la importancia de esta aclaración. Cuando en la doctrina se afirma que esas máximas de experiencia se utilizan para valorar la prueba, debiera comprenderse, por lo visto hasta ahora, en el sentido más instrumental posible: como una herramienta o un medio para alcanzar el fin de esa actividad (aproximarse a la verdad de los hechos) y no como una suerte de guía para su concreción. ¿Cómo debe realizarse esa actividad y, de modo más específico, cómo deben utilizarse esos enunciados? La respuesta la encontraremos en la propuesta prescriptiva que nos brindan los criterios de racionalidad epistemológica o reglas de la sana crítica. Considerar que esta clase de enunciados posee una capacidad normativa sería equivalente a intentar conocer cuál es la mejor manera de hacer un agujero con un taladro a partir de la compra de esa clase de herramienta.

¹⁶ Soy consciente de que el esquema de TOULMIN podría no reflejar con precisión todas las etapas del razonamiento probatorio en un proceso judicial. Sin embargo, como enuncié al comienzo, me interesa concentrarme únicamente en el momento de la valoración de la prueba. Una crítica al potencial del esquema aquí escogido puede verse en AGUILERA (2020: 31-32), quien recurre al modelo —por cierto, sumamente interesante— de TUZET (2014) para distinguir dos fases del razonamiento probatorio y destacar, a su vez, la importancia del razonamiento abductivo. Sin dudas, este aspecto excede los límites del presente trabajo como se señaló en la introducción, pero es un punto que debe estudiarse para comprender de manera global el objeto de esta investigación.

tintos» de estas últimas (2007: 137) que integran el argumento sobre hechos, por ahora será suficiente aceptar como punto de partida la propuesta básica del argumento integrado por D, G y C.

Asimismo, TOULMIN (2007: 141) señala que toda garantía posee un *respaldo* [«R»], esto es, razones a partir de las cuales aceptamos como válida una G en particular, y que consisten en *enunciados categóricos acerca de hechos* —de la misma clase que los datos o afirmaciones—. Sin embargo, aclara de seguido que la garantía «es algo más que la repetición de los hechos aducidos: es una *moraleja* o consecuencia general de carácter práctico sobre la manera en que se puede argumentar con seguridad a la vista de esos hechos» (2007: 143).

Sobre ese marco, si lo que se suele llamar máxima de experiencia busca servir como una herramienta —empírica— la valoración de la prueba y, como se vio, posee la estructura de los enunciados que prototípicamente cumplen la función de garantía en un argumento sobre hechos —enunciados generales condicionales—, y explícitamente descartamos que la garantía sea un enunciado particular sobre un hecho (una observación singular), podemos afirmar que su utilización en el razonamiento probatorio busca *legitimar el paso* entre un dato (que proviene de un medio de prueba) y una conclusión (o enunciado fáctico a probar) al cumplir el rol de *garantía* en un argumento sobre hechos.

La utilización del esquema propuesto no es novedosa y la ubicación de esta clase de enunciados dentro de él tampoco. Por ejemplo, GONZÁLEZ LAGIER sigue el esquema de TOULMIN sobre modelos de argumentos y propone una estructura de inferencia probatoria análoga: «Los hechos probatorios constituirían las razones del argumento; los hechos a probar, la pretensión o hipótesis del caso; la garantía estaría constituida por las máximas de experiencia, las presunciones y otro tipo de enunciados generales que correlacionan el tipo de hechos señalados en las razones con el tipo de hechos señalados en la pretensión» (2013: 41)¹⁷.

Para diferenciar los distintos tipos de garantías que podemos encontrar en una inferencia probatoria, GONZÁLEZ LAGIER señala que si bien tanto las máximas de experiencia, las presunciones y las definiciones son enunciados generales, lo que cambia es su concreta estructura: en el caso de las máximas de experiencia, «Si X, entonces probable Y»; en el caso de las presunciones: «Si X, entonces probado Y»; y en el caso de las definiciones, «X cuenta como Y», y explica que la diferencia entre presunciones y máximas de experiencia consiste en que las primeras son enunciados revestidos de autoridad (2013: 44).

Aquí, como se dijo en un comienzo, no se abordarán otras clases de garantías como podrían ser las presunciones o definiciones. Sin embargo, el párrafo anterior

¹⁷ En este punto, GONZÁLEZ LAGIER (2013: 59) precisa que la garantía consiste en «una regla (en el sentido de enunciado que expresa una regularidad) que correlaciona las razones (los hechos probatorios) con la pretensión (la hipótesis)».

demuestra un punto importante. En la medida en que el razonamiento probatorio es prototípicamente probabilístico (FERRER BELTRÁN, 2007: 91-92), excepto en el caso de las presunciones —que cuentan con fuerza normativa¹⁸— y las definiciones —que se trataría de un supuesto de subsunción o calificación¹⁹—, las conclusiones a las que permitirán llegar los enunciados condicionales con referente empírico serán únicamente probables, en mayor o menor medida, por el carácter también eminentemente probabilístico del conocimiento humano. Ello descarta la posibilidad de considerar como válida una estructura sobre la inferencia probatoria como un silogismo deductivo, en los términos expuestos —por ejemplo y entre otros— por STEIN.

Asimismo, al menos en el marco de un sistema de libre valoración de la prueba, la utilización de garantías que podrían catalogarse como *normativas*²⁰, como las presunciones, suele ser bastante limitada y, por el contrario, *la mayoría de los argumentos sobre hechos se construyen con base en garantías con fundamento o respaldo empírico, como los enunciados que se vienen considerando*: enunciados fácticos generales y condicionales.

Esto también deja en evidencia la importancia de avanzar en una mayor precisión del asunto.

3.3. El concepto de generalizaciones empíricas

En la doctrina de la tradición angloamericana, se ha afirmado de modo pacífico que toda inferencia probatoria empírica²¹ depende de una *generalización* (ANDERSON, 1999: 455, ANDERSON, SCHUM y TWINING, 2015: 139, STEIN, A., 2005: 65, ROBERTS y ZUCKERMAN, 2004: 111 y 115). De hecho, se ha replicado en diversas

¹⁸ Cuando se habla de fuerza normativa, lo que se quiere significar, en definitiva, es que para los enunciados presuntivos que cumplen la función de garantía, el respaldo (el enunciado categórico o, en otros términos, la *razón* que sostiene la garantía) proviene de la autoridad de la que se encuentran revestidos a partir del Derecho (GONZÁLEZ LAGIER, 2013: 62). De este modo, no se trata aquí de un respaldo vinculado a enunciados o razones empíricas sino, en verdad, su apoyo lo confieren razones normativas (la existencia de una norma jurídica de presunción). Por eso sería posible hablar de garantías de tipo normativas, a diferencia de las garantías empíricas, en virtud de la clase de razones que subyacen a cada una de ellas. De cualquier manera, ello no impide afirmar que las normas de presunción no posean, a su vez, razones empíricas que las sustenten en muchas ocasiones, a las que correspondería añadir, también en algunas oportunidades, razones no necesariamente epistémicas (*i.e.*, pragmáticas). Agradezco a un/a revisor/a anónimo/a por advertirme acerca de la necesidad de precisar este punto.

¹⁹ Esto es advertido por GONZÁLEZ LAGIER (2013: 61): «cuando la unión entre los hechos probatorios y el hecho a probar viene dada por una teoría o una definición (esto es, cuando el vínculo es conceptual), no nos encontramos en sentido estricto ante un caso de prueba, sino de interpretación o calificación de los hechos».

²⁰ Véase *supra* nota 18.

²¹ El concepto de inferencia probatoria empírica se utiliza aquí en los términos de GONZÁLEZ LAGIER (2018): aquella cuya garantía consiste en un enunciado fáctico general y condicional, y su respaldo, por tanto, es de corte empírico.

oportunidades una metáfora de SCHUM que afirma que «las generalizaciones son el pegamento que mantiene a nuestros argumentos unidos» (1984: 82)²².

En el plano continental europeo, FERRAJOLI ha señalado que la inferencia probatoria es típicamente inductiva, y que sus premisas la constituyen la descripción del hecho, las pruebas practicadas y «generalizaciones habitualmente sobreentendidas (entimemáticas) sobre la fiabilidad de experiencias análogas», para así arribar a una conclusión que consiste en el hecho que se acepta como probado (1995: 53). También GONZÁLEZ LAGIER explica que, en ocasiones, una generalización es lo que asocia los hechos que queremos probar con aquellos que surgen de las pruebas con las que contamos (2013: 76), y FERRER BELTRÁN refiere que son la garantía de una inferencia que va de un hecho a otro (2007: 133).

En síntesis, diversos autores y autoras han identificado, aunque no necesariamente de manera explícita, a las generalizaciones con las garantías de los argumentos sobre hechos. De manera tal que podría resultar fructífero explorar ese concepto para aclarar el panorama de las máximas de experiencia, tal como parece insinuar TARUFFO al vincular el fundamento cognoscitivo de las máximas de experiencia a la *validez y admisibilidad* «de las generalizaciones que expresan» (2009: 441).

SCHAUER es quizás quien mejor se ha ocupado de las generalizaciones. Su obra *Profiles, probabilities and stereotypes* (2003) comienza con la descripción del «pensamiento actuarialista» de las compañías aseguradoras: agrupar a las personas en grandes categorías y asignar a la totalidad de cada una de esas categorías ciertas características probabilísticamente presentes en la mayoría de sus miembros, más allá de que alguno en particular no la posea (2003: 4). A partir de allí, su reflexión es la siguiente: *todos operamos de ese modo en diferentes aspectos de nuestra vida*, en mayor o menor medida. Por ejemplo, al elegir compañías aéreas que consideramos confiables o puntuales, o asociar ciertas características a una determinada nacionalidad, por lo cual esa actividad, esto es, la actividad de *generalizar* no puede ser considerada irracional (aún cuando sepamos que puede producir errores en casos particulares).

En este sentido, SCHAUER explica que al enfrentar la realidad si bien nos encontramos con individuos, los identificamos como parte de agrupamientos mayores. De esta manera, no los vemos como «individuos» a secas, sino como *individuos con una cierta característica x por su pertenencia a esas clases o grupos*, motivo por el cual son, en verdad, *instancias* de esa categoría mayor (2004: 76).

A su vez, un punto importante de su tesis es la *simultaneidad* de las categorías, esto es, que un individuo pertenece, al mismo tiempo, a más de una. Estas no son excluyentes ni, muchas veces, de tan nítida diferenciación, pero lo central aquí es que: 1) elegimos categorizar al individuo y dejar de verlo como un singular y, 2) en paralelo, también optamos por una de las diversas clases o categorías en la cual subsumirlo.

²² La referencia a SCHUM puede verse, entre otros, en BEX, 2011: 36; ANDERSON, SCHUM y TWING, 2015: 139; STEIN, A., 2005: 65-66; TARUFFO, 2009: 446.

De este modo, SCHAUER define a la actividad de identificar a un individuo como parte de una categoría como *generalizar* y a su producto como una *generalización* (2004: 77). Esta actividad es siempre *gradual* y *direccionada*: la selección de una categoría más amplia o restringida, integrada parcialmente por miembros en común, decantará en una generalización más o menos general, y demostrará también la preferencia por la asignación de un grupo de pertenencia respecto de otro (2004: 79).

El resultado de la actividad de generalizar descripta por SCHAUER consiste, básicamente, en formular enunciados condicionales generales de la clase que se analizaron más arriba. Esos enunciados se denominan *generalizaciones empíricas*, en virtud de que son producto del ejercicio de generalizar y, a su vez, buscan describir e informar acerca del mundo explicándolo de manera adecuada. Cuando afirmamos que «los brasileños son buenos jugadores de fútbol» estamos asociando una característica a la categoría «brasileño» —a pesar de ser conscientes de que existen personas de Brasil que no juegan bien al fútbol— y pretendemos describir adecuadamente la realidad de ese conjunto.

A su vez, a partir de otro ejemplo podemos ver la cuestión vinculada al grado de generalidad y la dirección de la generalización, y cómo varían en función del modo en que realizamos la actividad de generalizar.

Uno podría generalizar en una dirección en particular: «las mascotas son buenas compañeras», pero también podría decidir hacerlo en otra: «los seres sensibles son buenos compañeros». Aquí puede verse cómo, de acuerdo con la variación en la dirección escogida, algunas clases de individuos integraran un grupo, pero estarán fuera del otro, o bien podría estar en los dos (los «perros» probablemente estén en los dos, ¿pero diríamos lo mismo de los «zorros»?). Por otra parte, «los gatos son buenos compañeros» es un grupo con un mayor grado de generalidad, y uno más reducido sería «los gatos siameses son buenos compañeros». Las generalizaciones empíricas que se obtienen en cada caso buscan describir e informar sobre realidades diversas.

A partir de lo dicho hasta aquí, podría evaluarse reemplazar aquel concepto oscuro de «máximas de experiencia» por el de *generalización* o *generalizaciones empíricas*, y evaluar si el aparato teórico asociado a este concepto nos permite una implementación fructífera.

3.4. El aparato conceptual de Schauer sobre las generalizaciones empíricas

La cuestión central se trata, entonces, de analizar cuándo una generalización podría ser considerada válida, o no, para integrar el razonamiento probatorio como garantía de nuestras inferencias, si tenemos en miras una finalidad epistémica de su utilización, esto es, aproximarnos de la mejor manera posible a la verdad de los hechos.

Se tomará como punto de partida la propuesta analítica de SCHAUER, quien efectúa una primera distinción entre generalizaciones *espurias* (*spurious generalizations*) y *no espurias* (*nonspurious generalizations*): el primer grupo no presenta una base empírica sólida, mientras que el segundo sí lo hace (2003: 7).

A su vez, dentro de la última categoría, diferencia entre aquellas generalizaciones *universales*, esto es, en las que la característica asociada al conjunto define a *todos* sus integrantes y se encuentra *siempre* presente (2003: 8)²³ y las *no universales*, es decir, aquellas en las cuales la característica asociada al conjunto no es definitoria y no se aplica a *todos* los miembros (2003: 9)²⁴, pero cuya invalidez para algún caso particular no nos limita a dejar de utilizar la generalización²⁵. De hecho, destaca que muchas veces aplicamos generalizaciones cuando ni siquiera la mayoría de los integrantes del conjunto de *x* poseen la característica *y* (2003: 10)²⁶.

De este modo, SCHAUER considera que una generalización contará con apoyo empírico sólido cuando: a) se trate de una de las caracterizadas como *universales*; b) describa con precisión las características de una mayoría de los integrantes de una clase; o c) retrate a los miembros de la clase *x* que tienen un rasgo o característica con mayor prevalencia respecto del conjunto *z* (conjunto mayor del que también forma parte *x*), aunque el rasgo o característica aparezca en menos que la mayoría de los integrantes de ambos grupos (2003: 10).

Sobre este último caso de generalizaciones *no universales*, SCHAUER precisa que aquello que torna sólida a una generalización de este tipo (en su ejemplo: «los bulldogs tienen problemas de caderas») es que conocer el dato acerca del conjunto *x* resulta *relevante* a los fines de la característica *y*, esto es, saber quienes conforman *x*, *otorga más información* sobre la presencia de la característica *y* que no poseer el dato relativo al conjunto *x*.

²³ Su estructura sería «todos los *x* son *y*». A pesar de utilizar el término *definen*, el propio SCHAUER destaca que existen casos en los que las generalizaciones universales no se deben a razones de definición de un término sino, en verdad, a cuestiones empíricas: mientras que «los solteros no están casados» se vincula con el primer grupo, «no hay seres humanos mayores a 2,74 metros» [el autor usa aquí «9 feet tall»] es una generalización con sustento empírico, pues nada de la definición de ser humano indica que deban tener una altura máxima determinada. El primer supuesto podría tratarse de una generalización *conceptual*, mientras que el segundo ejemplo se trata de una generalización *empírica*.

²⁴ En este caso, la estructura de estas generalizaciones sería: (*Muchos, la mayoría, un número desproporcionado* de *x* es *y*).

²⁵ Así, el autor explica que la generalización «los quesos suizos tienen agujeros» *se utiliza a pesar de saber que no todos los quesos suizos poseen esa característica*, y aún con ese conocimiento no antepone «la mayoría», «muchos» u otra expresión similar. De este modo, señala que mientras nos encontremos satisfechos con ella, por cuestiones de celeridad y comunicación, es parte aceptada de cualquier conversación, a sabiendas de su inexactitud (2003: 9).

²⁶ El siguiente ejemplo es utilizado en el texto: la generalización «los bulldogs tienen problemas de caderas» no significa que la *mayoría* de los bulldogs tenga ese problema, sino que presentan ese problema *en una cantidad mayor que toda la categoría (más amplia) «perros» a la que pertenecen*, y también en un porcentaje mayor que otras razas.

Por el contrario, SCHAUER considera que se tratará de una generalización *espuria* cuando no exista evidencia relevante de conexión entre la categoría *x* y la característica *y*, lo que –señala– en términos de derecho probatorio se denomina *irrelevancia* (2003: 13): esto sucederá en cada oportunidad en que la generalización no cumpla con alguno de los tres criterios mencionados anteriormente para determinar su apoyo empírico.

Las generalizaciones espurias, explica, se suelen denominar *prejuicios* (*prejudices*) cuando se vinculan con grupos de personas. Sin embargo, también pone de manifiesto SCHAUER que se suele emplear el término «prejuicio» cuando se realiza una generalización no universal con base sólida (2003: 16), lo que beneficia la confusión conceptual entre estas dos categorías. A su vez, advierte que similares consideraciones pueden realizarse respecto del término *estereotipo* (2003: 17), pues la ambigüedad aquí también se encuentra presente: se puede referir a una generalización sin ningún tipo de respaldo empírico, o bien a la utilización de generalizaciones con base sólida para la toma de decisiones respecto de toda la clase «estereotipada», incluyendo los miembros a los cuales no se les aplica correctamente la generalización.

A continuación, se pondrá a prueba este aparato conceptual vinculado a las generalizaciones empíricas desarrollado por SCHAUER y se evaluará concretamente su utilidad para tratar los problemas más habituales vinculados a las garantías de los razonamientos sobre hechos, de modo comparativo con algunas otras propuestas analíticas y clasificatorias.

4. LOS PELIGROS DE LAS GENERALIZACIONES: INVALIDEZ DE LA GENERALIZACIÓN Y PELIGROS EN EL USO DE LAS GENERALIZACIONES

4.1. Invalidez de la generalización

El primer paso de quien pretenda utilizar una generalización como garantía de una inferencia, o analizar su corrección, será evaluar si se trata de una generalización *válida*²⁷. Es decir, de acuerdo con la propuesta de SCHAUER, una que cuenta con base empírica sólida o si, caso contrario, se intenta utilizar una generalización espuria. Se trata aquí de detectar (y descartar) un posible caso de invalidez *de la generalización*: una generalización será inválida en la medida en que no cuente con apoyo empírico, de alguno de los modos expresados más arriba.

Este punto de partida es *independiente de la inferencia concreta en la que se la pretenda utilizar o se la haya utilizado*, pues una generalización será válida (no espuria)

²⁷ De aquí en adelante, generalización «válida» debe entenderse como sinónimo de generalización *no espuria*, es decir, que cuente con base empírica sólida o, más sencillamente, apoyo empírico, en los términos explicados.

o inválida (espuria), en los términos de apoyo empírico, con prescindencia del razonamiento en el cual se pretenda insertar o se la haya incluido, y *con independencia del sujeto que la formule*.

Sin perjuicio de ello, parece sensato que, si se pretende efectuar este análisis en el marco de una argumentación en un proceso judicial, el primer paso sea identificar con precisión cuál es la generalización que se busca analizar. Muchas veces este no es un ejercicio sencillo, en la medida en que nuestros razonamientos –en la vida cotidiana y, también, en el derecho– suelen ser entimemáticos (ALCHOURRÓN, 2000: 12; FERRAJOLI, 1995: 53; TOULMIN; 2007: 136, BONORINO; 2015: 42), es decir, alguna de sus partes (premisas o conclusión) no se encuentra explícitamente formulada.

Si bien el modo de enfrentarnos a los entimemas requeriría un análisis más profundo, se enunciará alguna propuesta en el siguiente apartado. Por el momento, bastará decir que si la parte omitida es la garantía del argumento, lo que resulta sumamente frecuente, se deberá *completar* –obviemos por ahora cómo llevar adelante esta tarea– con la generalización ausente para poder analizar, entonces, si aquella es válida o espuria.

Pues bien, una vez identificada la generalización, para cumplir ese objetivo debe analizarse si cuenta o no con apoyo empírico o base empírica sólida. Como se mencionó más arriba, las expresiones «apoyo empírico» o «base empírica sólida» reflejan tanto: 1) la verificación de vínculos que expresan *asociaciones universales* (por ejemplo, relaciones causales en sentido estricto como la ley de gravitación o características presentes en todos los integrantes de una clase, como una altura menor a x en las personas humanas); 2) la constatación de una característica en, al menos, más de la mitad de los miembros de una clase (lo que podríamos llamar una *asociación fuerte*); 3) casos de un alto grado de representatividad de una clase menor respecto de una mayor en punto a la presencia de una característica (que se podría denominar como *asociación por representatividad*).

La validez de una generalización implica que la pertenencia a la clase que conforma su antecedente hace más probable poseer la característica prevista en su consecuente²⁸. Por lo tanto, la clase resulta *relevante* –como expresa SCHAUER, en términos similares a lo que llamamos *relevancia* en el derecho probatorio– para la característica asociada, toda vez que aumenta la probabilidad de que esta se encuentre presente. Así como, en el razonamiento probatorio, la existencia de una prueba relevante hace más probable una conclusión sobre un determinado hecho²⁹.

²⁸ DAHLMAN (2016: 84), de un modo similar, considera que una generalización conecta dos «clases»: la clase «de referencia» (*reference class*) y la clase «apuntada» (*target class*); y que la pertenencia a la clase de referencia aumenta las posibilidades de integrar la clase apuntada.

²⁹ Si las generalizaciones permiten mantener nuestros argumentos unidos y dar el paso datos-conclusión, podría pensarse que el criterio general de relevancia (FERRER BELTRÁN, 2007: 42-43) no es más que una instancia de análisis de, al menos, la validez de las generalizaciones que se pretenden utilizar. Las pruebas que se intentan hacer valer como relevantes, en definitiva, lo serán *en función de*

Por lo dicho hasta aquí, es más fácil comprender por qué SCHAUER plantea ese *piso mínimo para la validez de una generalización*: si la generalización no nos proporcionara información correcta sobre el mundo, no lo describiera de una manera adecuada, en términos de aproximación a la realidad del grupo de casos que describe, no tendría ninguna capacidad o utilidad epistémica y por esa razón se trataría de una generalización inválida, ya que no se correspondería con la realidad. Por ese motivo, también SCHAUER considera que los casos de asociación por representatividad (el ejemplo que propone sobre los bulldogs) constituyen generalizaciones válidas, en la medida en que la constatación del antecedente posee utilidad epistémica, al incrementar las posibilidades de que se encuentre presente una característica determinada, al menos en términos relativos respecto de cualquier otra clase mayor, más allá de que, en términos absolutos, no pueda predicarse una asociación fuerte³⁰.

Las ventajas de este aparato conceptual para funcionar como un primer filtro de exclusión de generalizaciones inválidas-espurias podrían ser las siguientes.

En primer lugar, permite tomar distancia de un factor que, si bien muchos doctrinarios y doctrinarias han destacado, quizás no resultaría tan útil, al menos para

una generalización que asocie una característica (que se encontrará en la conclusión a la que queremos llegar) con una clase (a la que pertenecerá la prueba, esto es, la información que de allí se derive). Algo similar parecería sostener GONZÁLEZ LAGIER (2013: 58) cuando afirma que: «determinar qué hechos son pertinentes para confirmar la hipótesis depende de las máximas de experiencia y presunciones que constituyan la garantía del argumento. Un hecho no será pertinente cuando no está correlacionado con la hipótesis ni por presunciones ni por máximas de experiencia adecuadas y bien fundadas, por lo que este requisito remite a la corrección de la garantía».

³⁰ Esto también permite observar otro aspecto importante. Muchas veces se considera que el filtro de relevancia es una cuestión *de todo o nada*: se dice que una prueba es o no es relevante, sin matices. Sin embargo, si la relevancia se trata de la verificación de una instancia o un caso particular que pertenece a una generalización, podría ser más difícil sostener esa posición. Esto, centralmente, en la medida en que, como se señaló más arriba, la actividad de generalizar, y por lo tanto su producto –la generalización–, admite una graduación: puede ser más o menos general y eso afectará, casi con seguridad, el grado de la asociación, y en consecuencia, puede suceder que, incluso, una generalización válida para una clase resulte inválida para otra mayor a la que pertenece –no necesariamente–, como también que una generalización inválida para una clase más amplia puede ser correcta para un antecedente más reducido –no necesariamente–. O bien que, entre dos generalizaciones válidas (clase x y clase x1), existan distintos grados de solidez de la asociación. Esto podrá verse mejor con algunos ejemplos (asumiendo que la calificación de validez o invalidez es una conclusión verdadera sobre su grado de apoyo empírico):

Ampliación del antecedente y modificación de su relevancia hasta la invalidez

Generalización válida: Los estudiantes universitarios de Girona suelen trabajar cuatro horas por día.

Generalización válida: Los estudiantes universitarios de Catalunya suelen trabajar cuatro horas por día.

Generalización inválida: Los estudiantes universitarios de España suelen trabajar cuatro horas por día.

Reducción del antecedente y modificación de su relevancia hasta la validez

Generalización inválida: Los argentinos perciben un sueldo de \$60.000.

Generalización inválida: Los habitantes de la ciudad de Buenos Aires perciben un sueldo de \$60.000.

Generalización válida: Los profesionales de la ciudad de Buenos Aires perciben un sueldo de \$60.000.

este primer paso vinculado al análisis de validez: la clasificación según el origen o la fuente de la generalización³¹.

En efecto, si bien se podría intentar construir una prelación abstracta de fuentes o ámbitos de generalizaciones en función de su validez habitual (conocimiento científico altamente afianzado, conocimiento de las comunidades expertas, conocimiento «común», etc.), o establecer orígenes de generalizaciones («científico o especializado», «jurídico» y «privado», como se vio que propone GONZÁLEZ LAGIER), en principio parecería no tener mucha utilidad, toda vez que, más allá de cuál fuera su fuente u origen —circunstancia que, por lo demás, en ocasiones es difícil de identificar³²— nunca se encontrará eximida la generalización en concreto del análisis mencionado. Es decir, aun cuando *a priori* pueda afirmarse que una generalización tiene su origen en un conocimiento científico afianzado, o en el «sentido común», o bien que proviene de la experiencia profesional o personal del juez, será necesario detenerse a evaluar su apoyo empírico³³.

En segundo término, también hace posible otorgarles un justo lugar a los factores *tiempo y espacio* que, como se pudo ver, eran otro de los aspectos mencionados por diversos doctrinarios y doctrinarias cuando trataban las «máximas de experiencia». Así, tal como en las ciencias se exige para la formulación de hipótesis que estas se funden en los conocimientos existentes en un tiempo y lugar determinado (FERRER BELTRÁN, 2007: 130, con cita de BUNGE, 1967: 200 y 237-242), a la epistemología de su tiempo, también este requisito se reflejará en la validez de las generalizaciones: una generalización que pretenda sustentarse a contramano de aquel conocimiento no estará reflejando una correspondencia con la descripción de la realidad que se considera verdadera con base en los métodos de obtención de conocimiento vigentes.

³¹ Esta tendencia hacia el establecimiento de fuentes *con mayor validez, con menor validez y con nula validez*, de modo abstracto, puede verse, por ejemplo, en el eje de confiabilidad propuesto por ANDERSON, SCHUM y TWINING, 2015: 325, aunque allí también se efectúan consideraciones respecto del grado de solidez de la generalización. También parecen insinuarlo GASCÓN ABELLÁN, 2010 [1999]: 160 y BEX, 2011: 19.

³² Piénsese en la tan repetida generalización acerca de la huida como indicador de culpabilidad: ¿se trata de una generalización que proviene del sentido común o, en cambio, se vincula con conocimientos de física o biología (capacidad de desplazamiento de los seres humanos o la huida como un rasgo evolutivo del propio ser humano), de psicología (huida como reacción frente a una situación de *stress*)?

³³ Vemos con un ejemplo por qué este modo de diferenciar las generalizaciones, en abstracto, no sería tan útil. La generalización «los hinchas de Boca Juniors se retiran del estadio antes de que finalice el partido cuando su equipo se encuentra perdiendo» podría ser una generalización válida o inválida: esto dependerá de su correspondencia con la realidad. Sin embargo, la clasificación que se critica parecería indicar que es una generalización espuria *per se*, o *tendencialmente* espuria, al no poder identificar una fuente «científica» o de conocimientos afianzados, con independencia de su grado de correspondencia con la realidad. Si la clasificación pretendiera funcionar como una suerte de indicador de «alerta» para analizar con mayor detenimiento una clase de generalizaciones, las que se alejan del polo «científico», podría correrse un riesgo muy alto frente a su pretendida utilidad: el riesgo de aceptar acríticamente —es decir, sin un análisis riguroso de su validez— aquellas con cierta científicidad. Al final del día, tal vez no reporta tantas utilidades como en una primera aproximación.

Por último, permite escindir esos factores cuando no presentan utilidad para el análisis: en este sentido, en la medida en que el apoyo empírico *implica correspondencia con la realidad*, no tendrá incidencia —en principio, como regla general— el grado de convicción compartida o la permeabilidad de una generalización en una comunidad determinada³⁴.

Ahora bien, lo dicho hasta aquí permitirá efectuar un análisis de validez de las generalizaciones depurado y excluir generalizaciones espurias de nuestros argumentos sobre hechos. Sin embargo, una vez que nos encontramos con una generalización válida, se abre un nuevo campo de análisis para su utilización como garantía del razonamiento y para la evaluación del apoyo que, como tal, otorga a la conclusión.

4.2. Peligros en el uso de la generalización

Una vez que es posible afirmar que pretendemos utilizar una generalización válida, nos enfrentamos al segundo grupo de desafíos. Estos son muy variados e involucran cuestiones que van desde la fuerza con la que esa generalización apoya el paso desde el dato a la conclusión, la relevancia de la clase escogida para generalizar frente a otras también válidas posibles, la representatividad asignada a la generalización y muchos otros más.

Se trata, en definitiva, de problemas no necesariamente propios de la generalización —que ya fue calificada como válida— sino de posibles *peligros derivados de su utilización en una inferencia probatoria concreta*.

4.2.1. La solidez de las generalizaciones

Si el primer paso consiste en descartar las generalizaciones espurias, una vez admitida una como válida —dejemos de lado, por el momento, cualquier otra clase de problemas— debemos evaluar cuánto apoya la conclusión hacia la que se quiere llegar desde la información existente.

Usualmente se distingue, para efectuar este análisis, entre el fundamento y la fuerza de una generalización (GONZÁLEZ LAGIER, 2013: 76) o, en otras palabras, entre su fuerza de respaldo y su fuerza probatoria (ANDERSON, SCHUM y TWINING, 2015: 323). Aquí, en lugar de fuerza o fuerza probatoria, se hablará de solidez. Veamos con un ejemplo.

³⁴ Por ejemplo, los aficionados de River Plate podrían estar, en su totalidad, convencidos hace años de la validez de la generalización «los hinchas de Boca Juniors se retiran del estadio antes de que finalice el partido cuando su equipo se encuentra perdiendo». Sin embargo, de acuerdo con la noción de validez propuesta, este factor es irrelevante para su evaluación: si uno se detuviera en las salidas del estadio de Boca Juniors y constatará que los aficionados no se retiran, ni en su mayoría ni uno solo, antes de que termine el partido, la convicción compartida y asentada de los hinchas de River Plate no tendría incidencia alguna.

La generalización universal «*todos* los objetos que son lanzados al aire son *en todos los casos* atraídos a la tierra por la fuerza de gravedad» correlaciona a cada uno de los integrantes del antecedente «objeto lanzado al aire» con una característica «atraídos a la tierra» que se encuentra presente en todos ellos.

Esto implica que *el grado o la intensidad de la asociación detectada es sumamente alto*, el más alto posible. Este será, sin dudas, un factor que considerar para determinar la solidez de la generalización. Cuanto mayor sea el grado o intensidad de la asociación que expresa la generalización, mayor tenderá a ser su solidez. Sin embargo, sería un error derivarla únicamente de ese factor.

No solo resulta importante conocer ese dato, sino también es necesario contar con el fundamento cognoscitivo a partir del cual se lo extrajo. Si se formulara la generalización «los vecinos del barrio del Clot –Barcelona– tienen más de 60 años» a partir de la entrevista al 60% de la población de ese vecindario, pareceríamos estar en peores condiciones epistémicas que si lográramos preguntarle su edad al 80% de las personas que allí viven y arribar a esa conclusión. Este es el segundo factor que hay que considerar para determinar, en definitiva, la solidez de la generalización³⁵. Y como tal, constituye otro foco de posibles peligros en el uso de una generalización.

En este sentido, la mayoría de los autores y autoras parece coincidir en que la fuerza de la generalización, comprendida como la solidez que proporciona al enlace entre el dato y la conclusión, es proporcional a: 1) el fundamento cognoscitivo que *respalda la inducción* que da origen a la generalización y, 2) la intensidad o grado de probabilidad de la asociación subyacente (GONZÁLEZ LAGIER, 2013: 76; GASCÓN ABELLÁN, 2010 [1999]: 160; ROBERTS y ZUCKERMAN, 2004: 114).

Por estas razones, es adecuado detenerse a evaluar el juego que estos dos factores otorgan, de manera conglobada, para expresar el grado de apoyo inductivo que la generalización de que se trate brinda al salto del elemento de juicio a la conclusión. Pero ese campo es donde, precisamente, se hace presente el peligro de manipulación: asignar una mayor solidez a la generalización de la que verdaderamente posee, pretender atribuirle un mayor fundamento cognoscitivo a su respaldo o distorsionar la intensidad de la asociación.

4.2.2. *La selección de las generalizaciones*

Otro foco de potenciales riesgos en el uso de las generalizaciones consiste en su *selección*. Al respecto, vale decir que, en muchos casos, no resulta evidente cuál es la generalización, de entre muchas *válidas* posibles, que mejor se ajusta a la inferencia y, de hecho en la mayoría de las ocasiones existe más de una calificable como elegible.

³⁵ En el primer ejemplo, sin embargo, no se constató *cada uno de los integrantes* del antecedente, pero el tipo razonamiento demostrativo que subyace permite establecer con un *altísimo fundamento cognoscitivo* que todos los integrantes poseerán la característica.

De este modo, deberíamos escoger aquella/s generalización/es que nos conduzca/n a un mejor resultado epistémico —una más precisa correspondencia con la realidad—, pues de lo contrario incurriríamos en un uso inadecuado de la generalización: se insertaría en la inferencia una generalización que arrojaría resultados epistémicamente sub-óptimos. ¿Cómo inclinarnos, entonces, racionalmente por una/s o por otra/s? ¿Debemos preferir una generalización más «específica» pero con menor grado de solidez o, por el contrario, debemos confiar en aquellas generalizaciones con un nivel de generalidad mayor pero también con un alto grado de apoyo empírico?

Por ejemplo, para evaluar la credibilidad de un testigo un litigante podría alegar la generalización «las personas que declaran bajo juramento suelen decir la verdad». Sin embargo, la contraparte o el juez podría echar mano a aquella que establece que «la coherencia del relato suele ser indicador de veracidad», fundado en que la psicología del testimonio ofrece mejores garantías para valorar los indicadores de credibilidad. Asimismo, el juez de una etapa recursiva podría estimar que una generalización adecuada es aquella, más específica, que dispone que «las personas con antecedentes penales por falso testimonio no suelen ser veraces».

Veamos qué propuestas se han realizado para dirimir esta encrucijada.

4.2.2.1. Las generalizaciones específicas del caso y de contexto vs las generalizaciones “ad-hoc”

La distinción que trazan ANDERSON, SCHUM y TWINING entre generalizaciones de contexto y específicas del caso (2015: 325) tal vez nos guíe por un camino equivocado para el análisis de la selección de las generalizaciones.

Para los autores, «especialmente si son apoyadas por pruebas», las generalizaciones específicas del caso suelen triunfar o desplazar a las generalizaciones de contexto. Como ejemplos de las primeras mencionan aquellas vinculadas a los hábitos o al carácter *de una persona* y, concretamente, enuncian la siguiente: «el empleador-demandado en este caso regularmente discriminaba a mujeres en sus prácticas laborales» (2015: 326). Concluyen su punto al afirmar que «[u]na generalización específica-del-caso, respaldada por un testigo, sobre los hábitos de [X] a este respecto hubiese sido mucho más persuasiva que cualquiera de estas generalizaciones de contexto» (2015: 327). En una posición similar se ubica TARUFFO, para quien solo pueden utilizarse máximas de experiencia (a esta altura, llamémoslas generalizaciones) que se refieran *directamente* al caso específico, pues aquellas demasiado vagas y genéricas carecen de utilidad e incrementan el riesgo de malentendido (2009: 453). Es decir, la tesis de estos autores es que las generalizaciones específicas del caso son *epistémicamente preferentes* por sobre las de contexto. Podríamos identificar este argumento como el de la «preferencia epistémica por las generalizaciones específicas».

Para evaluar la plausibilidad del argumento, es importante detenerse en los ejemplos que proponen ANDERSON, SCHUM y TWINING para comprender qué es lo que entienden por generalizaciones específicas del caso. Así, enunciados del tipo «la per-

sona x tiene el hábito y) o «el empleador-demandado regularmente discriminaba a mujeres en sus prácticas laborales» serían, en su opinión, prototípicamente generalizaciones de esa clase.

Sin embargo, si se observa con atención, esos enunciados podrían integrar un argumento sobre hechos en calidad de datos o bien como conclusiones, pero no tendrían el rol de garantía del argumento. Veamos con mayor detalle la cuestión a partir del primer enunciado que se propone.

Para afirmar que la persona x tiene el hábito y , esto es, para arribar a esa conclusión fáctica, debemos contar con ciertos datos. Esos datos presumiblemente sean que x realiza y con una cierta frecuencia (diaria, z veces a la semana, semanal, z veces al mes, etc.). Para avanzar desde esos datos hacia la conclusión «la persona x tiene el hábito y », probablemente utilizemos, en el caso, una garantía de tipo conceptual (una definición de «hábito»).

Asimismo, también podríamos considerar que el enunciado «la persona x tiene el hábito y » se trata de un dato o afirmación que nos permitirá avanzar hacia una conclusión en un caso concreto. Por ejemplo, sabemos que existe una persona que tiene el hábito de robar utilizando una máscara de un personaje famoso y vestido con colores llamativos. En el caso concreto, nos encontramos con una persona acusada de un robo con idénticas características. El enunciado mencionado nos permitiría avanzar hacia la conclusión «la persona x es la autora de este delito» siempre y cuando utilizáramos una generalización que afirmara que «las personas suelen comportarse de acuerdo con sus hábitos».

Pretender que enunciados individuales de ese estilo («la persona x realiza la acción y diariamente» o «la persona x tiene el hábito y ») puedan ser considerados generalizaciones puede obedecer, por un lado, a una confusión entre las garantías o generalizaciones conceptuales y las garantías empíricas; por el otro, a la omisión de identificar los roles que cumplen cada uno de esos enunciados en un razonamiento sobre hechos, sin un análisis sobre su diversa estructura lógica y características definitorias.

Sin embargo, tampoco parecería aportar claridad a la tarea de seleccionar la generalización más adecuada la afirmación de que las generalizaciones «ad-hoc» o formuladas especialmente para el caso concreto no son confiables (TARUFFO, 2009: 453; GONZÁLEZ LAGIER, 2013: 59). Este argumento podría denominarse como el de «rechazo epistémico de las generalizaciones *ad-hoc*».

Al menos dos objeciones se pueden formular en su contra.

En primer lugar, no es del todo claro qué significa que una generalización sea formulada únicamente para el caso en concreto. En este sentido, el análisis de validez de una generalización, comprendido como constatación de apoyo empírico, es lógicamente previo a la utilización de la generalización en un razonamiento particular —como se señaló más arriba— de manera que parece difícil imaginar una generalización «para el caso concreto». Eventualmente, la discusión será si *este* o *aquel* caso es una instancia de esa generalización o si no existe otra generalización más adecuada.

A su vez, si la construcción de una generalización *ad-hoc* implica el intento de utilizar una con un antecedente singular (básicamente, uno presente en el caso del que se trate), se podrían formular idénticas consideraciones a las señaladas más arriba para criticar las generalizaciones «específicas del caso». En definitiva, que ese enunciado no puede ser considerado una generalización y, en todo caso, será una conclusión a la que se llega gracias a una generalización —por ejemplo, sobre los hábitos o comportamientos de las personas— y que podría servir, a su vez, como dato o afirmación para formular una nueva conclusión sobre un individuo en particular, necesariamente a través de otra generalización.

Sin embargo, podría ser imaginable un supuesto en que efectivamente se formulara una generalización *para el caso concreto*. Pero esto siempre y cuando sea comprendido en los siguientes términos: si se realizara una actividad epistémica de recolección de datos hasta ese momento inexplorada (por las razones que fuese) para fundamentar la validez de una generalización. Ejemplo: en un determinado caso, la defensa intenta demostrar, por algún motivo relevante, que una persona de x años, con las características a , b y c , tarda un tiempo $t1$ en recorrer caminando la distancia $d1$. Imaginemos que solo hay 100 personas de x años con las características a , b y c . Una de ellas es el acusado. La defensa, de manera privada, encuentra a otras 85 personas y les solicita que recorran esa distancia a pie. Todas tardan un tiempo $t1$. Asumamos, también, que las condiciones experimentales no son objeto de disputa. Luego de realizar esa actividad epistémica, la defensa está en condiciones de formular una generalización: «las personas de x años con las características a , b y c tardan un tiempo $t1$ en recorrer caminando la distancia $d1$ ».

La pretensión de catalogarla de «específica del caso» o «ad-hoc» dependería, únicamente, de que el primer lugar en donde fue utilizada se trató de *ese caso*. Pero su potencial epistémico excede los límites del proceso y se trata, lisa y llanamente, de una generalización válida.

De este modo, la etiqueta de generalización «específica del caso» o «ad-hoc» puede no resultar un parámetro satisfactorio para la selección de las generalizaciones.

En todo caso, frente a los escollos que presentan estas dos visiones del problema, tal vez resulte de alguna utilidad para la preferencia por una de entre dos o más generalizaciones en competencia, detenerse a observar el *grado de generalidad* del antecedente (de más general a menos, por ejemplo: «las personas...»; «los testigos...»; «los testigos con antecedentes penales...»).

Sin embargo, como se dijo anteriormente, el aumento o descenso de la generalidad puede venir de la mano, a su vez, con una variación de la intensidad de la asociación (por ejemplo: «las personas...casi siempre...», «los testigos... en muchas ocasiones...», «los testigos con antecedentes penales... a veces...»). A su vez, esa variación en el grado de generalidad también puede aparejar una modificación en la base empírica a partir de la cual se produce la generalización (i.e., la generalización

vinculada a los testigos con antecedentes penales tal vez tenga una mayor parte de los individuos de esa clase relevados, pues sería una categoría más reducida).

De este modo, parece difícil establecer un criterio claro *a priori* y, en todo caso, deberá ser un aspecto en el que la argumentación por la *preferencia epistémica* involucre tanto el grado de generalidad como la intensidad de la asociación, entre otros factores.

4.2.2.2. Las fuentes de las generalizaciones

Como se indicó más arriba, existe una gran preocupación en la doctrina respecto de la *fuerza* de las generalizaciones lo que, además de incidir en su validez, tendría injerencia *decisiva* en su solidez o confiabilidad como garantía del argumento (ANDERSON, 1999: 458-459; ANDERSON, SCHUM y TWINING, 2015: 325³⁶; GONZÁLEZ LAGIER, 2013: 59; GASCÓN ABELLÁN, 2010 [1999]: 160; BEX, 2011: 19). Ello pues, se afirma, existirían *en abstracto* fuentes de generalizaciones más confiables y otras menos confiables para formular inferencias. Esto traería consigo, en consecuencia, otro posible *foco de peligro* en el uso de las generalizaciones.

De manera sintética, se suele ubicar en el extremo de mayor confiabilidad a las leyes o reglas científicas y a los conocimientos expertos; en un punto medio, a las generalizaciones que son compartidas al interior de una sociedad determinada durante un tiempo prolongado y, por último, en el sitio de menor confiabilidad se ubica a las provenientes de la experiencia personal y el sentido común, a los que se suele calificar como prejuicios y estereotipos.

A su vez, existe una particularidad que ofrecen muchas de las generalizaciones que, en su mayoría, son utilizadas en el razonamiento probatorio, y es que parecen no poseer sustento empírico. Si uno excluye de la discusión aquellas que constituyen la cristalización de leyes científicas en el lenguaje vulgar («todo objeto que se eleva luego deberá bajar»), o las que provienen de un conocimiento necesariamente experto («la presencia de *x* puntos de coincidencia particulares indica que dos huellas dactilares son de la misma persona»), en las restantes resulta difícil evaluar *cuánto* es el apoyo empírico de la generalización («las personas que huyen de la escena del crimen suelen ser sus autores»). De hecho, uno podría genuinamente preguntarse si, en estos casos, no se pretende utilizar una generalización inválida ya que es difícil establecer *cuál* es su apoyo empírico.

Estas generalizaciones, se afirma, tendrían su origen en el sentido común, en las experiencias de las personas o en ciertas intuiciones compartidas: algo que suele identificarse como el *stock of knowledge*. A su vez, se considera que, por esa razón, este grupo de generalizaciones poseería un fundamento empírico que sería de muy difícil o imposible constatación y, también, que su grado de confiabilidad resultaría

³⁶ Las consideraciones que siguen son aplicables respecto del eje de *confiabilidad* propuesto por los autores, como así también respecto del eje de *fuerza*.

sensiblemente menor de ser válidas. Este grupo es identificado como el de las generalizaciones sintético-intuitivas (ANDERSON, 1999: 457; ANDERSON, SCHUM y TWINING, 2015: 331-332), en la medida en que sintetizan ciertas creencias o intuiciones de las personas.

Sobre este marco, creo conveniente hacer dos aclaraciones.

En primer lugar, la posición que postula la existencia de un cierto orden de prelación epistémica entre las distintas fuentes de generalizaciones, que llevaría a preferir una clase por sobre otra en punto a su confiabilidad, tal vez podría ser matizada o no necesariamente sostenerse de un modo tan extremo.

Si bien la actividad epistémica científica de recolección y análisis de datos puede ser muy rigurosa y reposar sobre teorías bien establecidas, y por esa razón suele ubicarse a las generalizaciones que de allí derivan como más confiables, muchas veces también carece de límites fácticos que resultan contraproducentes para la solidez de sus conclusiones (i.e., su campo de análisis es extremadamente extenso), mientras que áreas «no científicas» podrían producir prácticas epistémicas sumamente completas y confiables (por ejemplo, el supuesto analizado para las generalizaciones del caso concreto). De manera que, si bien la fuente podría funcionar como un indicio sobre la validez de la generalización, no se presenta como un factor sumamente relevante para el análisis de un caso en particular, ya que siempre deberá analizarse la generalización *en concreto* y, asimismo, frente a la existencia de dos generalizaciones en competencia de una misma clase de fuente directamente carecería de toda utilidad.

En segundo lugar, tampoco parece prudente descartar, *a priori*, todas aquellas generalizaciones cuya fuente y/o fundamento no podemos identificar o desconocemos, o que bien parecen referir a un sentido común compartido o una experiencia *intersubjetiva*, pues no es lo mismo confrontar una generalización determinada y concluir que no posee apoyo empírico, que evaluar la validez de una generalización y no poder formular una conclusión concreta. Esto último, en principio, requeriría una explicación adicional acerca de por qué deberíamos, frente a la disyuntiva de descartarla o aceptarla con cautela, inclinarnos por la primera, como sostienen algunas posiciones. Contrariamente, tal vez la falta de constatación de su falsedad podría ser un factor adicional para optar por su utilización, sumado a nuestras intuiciones sobre su corrección.

En alguna medida, algo de esto parece insinuar TARUFFO (2012: 252-254) al señalar tres condiciones de aplicabilidad para las nociones de sentido común: i) que se trate de nociones aceptadas en el ámbito social y cultural donde ha sido formulada la decisión; ii) que no sean falsas u opuestas al conocimiento científico, y iii) que no se encuentren en contradicción con otras nociones de sentido común. En este sentido, podría ser de utilidad un análisis a partir del eje del *grado de coincidencia* (ANDERSON, SCHUM y TWINING, 2015: 325), que busca discriminar las generalizaciones de acuerdo con cuán ampliamente podría ser aceptada o compartida al interior de la comunidad particular en la que una cuestión deba ser resuelta, que podría ir desde

un pequeño grupo de personas hasta la totalidad de la comunidad. En esta clase de casos, quizás, la exigencia de coincidencia pueda ser extremada para compensar la carencia de información acerca del apoyo empírico y la fuerza específica con la que cuenta la generalización.

Por último, se suele afirmar que el sentido común y las generalizaciones que de allí derivan suelen estar cargadas de «prejuicios» y «estereotipos» (TARUFFO, 2009: 445), y que por esa razón se debe ser extremadamente cuidadoso en su utilización.

En primer término, frente a afirmaciones de esa clase se impone señalar que, como vimos más arriba, «estereotipo» o «prejuicio» puede hacer referencia tanto a una generalización espuria como a una generalización no universal con apoyo empírico. De ese modo, es evidente que una generalización espuria, sea un estereotipo, un prejuicio o una falsa asociación, nunca debería integrar un argumento sobre hechos.

En segundo lugar, para los estereotipos considerados como generalizaciones válidas pero no universales, si bien es cierto que suelen ser tratados como generalizaciones empíricas por su estructura similar, también lo es que hay ciertos casos en los cuales poseen diferencias importantes.

De este modo, parece útil para analizar la validez y selección de un estereotipo tener presente la distinción entre estereotipos *normativos* y *descriptivos*. Los primeros son usados con la pretensión de definir y constituir roles, mientras que los segundos para proporcionar información sobre las características de un grupo (ARENA, 2016: 54-55), y se trata de aquellos que guardan similitud con las generalizaciones empíricas que aquí se mencionan.

Esta distinción permite analizar con criterios de relevancia y aceptación diversos cada una de las clases. Y también hace posible evaluar con mayor precisión algunos casos en los que se atacan ciertas generalizaciones bajo el rótulo de «estereotipos», y se los suele asociar a la falta de apoyo empírico, cuando en verdad se está hablando de estereotipos normativos y su inadecuación por otras razones, o bien se hace referencia a generalizaciones válidas pero se considera, también, inadecuada su selección³⁷.

³⁷ Por ejemplo, ANDERSON, SCHUM y TWINING afirman que en el límite inferior del eje de confiabilidad se encuentran «los prejuicios sin fundamentos que se basan en estereotipos falsos –tales como los prejuicios basados en género, raza, clase o edad–» y, en la nota al pie, mencionan un extracto del caso *Bradwell v. Illinois*, 83 U.S. 130 (1872) como un ejemplo de un prejuicio basado en un estereotipo falso: «El destino y misión fundamental de las mujeres es cumplir con los oficios nobles y benignos de esposa y madre. Esa es la ley del Creador. Y las reglas de la sociedad civil deben adaptarse a la constitución general de las cosas, y no pueden basarse en casos excepcionales» (2015: 326). Si se observa con atención, y se deja de lado el repudio que puede causar leer algo así, se podrá advertir que el tribunal emplea en su argumentación un estereotipo *normativo* («las mujeres deben ser esposas y madres») que no busca proporcionar información acerca del mundo, sino establecer ciertos deberes, de manera que el apoyo empírico que posea, o mejor dicho, su falta de apoyo empírico, no lleva a abandonar el estereotipo o considerarlo inválido, ya que su dirección de ajuste es grupo social-estereotipo (ARENA, 2016: 55). De hecho, podría suceder que la generalización, o el estereotipo descriptivo, que afirma «las mujeres son, en su mayoría, esposas/madres» sea válida, en el sentido de que describa adecuadamente la realidad, de

5. ALGUNOS CRITERIOS PARA LA UTILIZACIÓN DE LAS GENERALIZACIONES

En este último apartado se buscará enunciar, sin ánimos de exhaustividad, algunos criterios que pueden resultar útiles a la hora de trabajar con las generalizaciones en los procesos judiciales. Un estudio riguroso del asunto llevaría a la elaboración de un trabajo independiente, pero considero que pueden ofrecerse algunas opciones para comenzar el debate. Ojalá la propuesta sirva como punto de partida³⁸.

En primer lugar, antes de entrar en el asunto, es importante tener en cuenta que se habla aquí de *criterios* para la utilización de las generalizaciones pues, como destaca GONZÁLEZ LAGIER, el exceso de precisión en este tipo de tópicos corre el riesgo de establecer un nuevo criterio de valoración tasada de la prueba y, en cambio, los criterios más amplios permiten mantener las ventajas del sistema de libre valoración (2013: 64).

1. Ahora bien, uno de los primeros criterios que decanta de lo desarrollado en los apartados anterior es que una generalización para ser epistémicamente útil debiera poseer apoyo empírico sólido. Sin embargo, una vez superado ese primer filtro, vemos que la solidez es una cuestión gradual, una generalización puede ser *más o menos* sólida, de acuerdo con el apoyo con el que cuente. Así, probablemente no dudaremos en afirmar que una generalización que describa una asociación en la que se hayan analizado u observado todos los integrantes del grupo x que conforma el antecedente y se haya detectado la propiedad y en el 90% de los individuos que lo conforman, será más sólida que una generalización fundada en la observación de una característica en la mitad más uno de todos los integrantes del grupo. Sin embargo, ¿hasta dónde podemos exigir la actividad epistémica de recolección de datos respecto de todos los integrantes de la clase? ¿Y si la generalización respecto del grupo x únicamente se basó en la observación de una cantidad y de sus integrantes?

De esta manera, podría pensar que *un primer criterio* podría expresar que (i) una generalización será más sólida que otra cuando: (a) la actividad epistémica de recolección de información haya abarcado una mayor cantidad de integrantes de la clase y (b) la presencia de la característica y asociada a esos individuos se encuentre presente en un mayor número de casos.

manera que no podría afirmarse que se trate de un estereotipo «falso». De hecho, podría ser correcta esa descripción de la realidad por la existencia previa de un estereotipo normativo que imponía esos roles y que determinó históricamente la configuración de esa realidad. Pero la generalización empírica no sería inválida, en términos de apoyo empírico, aunque pudiera ser irrelevante, sospechosa o ilegítima por otros motivos (SCHAUER, 2003: 150-155). Para un análisis profundo sobre las diferencias entre estereotipos normativos y descriptivos, véase ARENA, 2016.

³⁸ Pueden verse otros criterios, algunos parcialmente coincidentes pero desde una perspectiva distinta a la de este trabajo, en TARUFFO (2002: 424-425).

2. A continuación, otro punto importante se trata de la *selección* de la generalización. Como se vio, la actividad de generalizar siempre se encuentra direccionada en un sentido, es decir, elegimos concentrarnos en una clase de las muchas posibles en las que puede ser ubicado un individuo. Y al hacer eso, descartamos otras posibles orientaciones de la generalización. De manera que la dirección de la generalización, en definitiva, se trata de una *elección*. Así, deberíamos ser capaces de *justificar racionalmente esa preferencia*. ¿Por qué incluir a un sujeto dentro del grupo «profesores universitarios» y excluirlo, por ejemplo, de la clase «testigos»?

A su vez, dentro de una misma dirección podríamos tener distintos *grados de generalidad*: «testigos», «testigos con antecedentes penales», «testigos con antecedentes penales por falso testimonio». ¿Es siempre preferible una clase más restringida o podemos/debemos optar por aquella generalización con mayor soporte empírico frente a dos en competencia?

Sobre este punto, un *segundo criterio de mayor relevancia epistémica de la generalización* parece presentarse como una herramienta útil para la toma de decisiones acerca de qué generalización escoger. Esto es, podría formularse un criterio que establezca que: (ii) una generalización válida desplazará a otra en competencia cuando haga más probable que su contrincante una mejor aproximación a la realidad, esto es, cuando su *solidez* sea mayor.

3. De este criterio puede desprenderse un tercero a modo de corrección. Si bien, en principio, aquella generalización de mayor solidez debería prevalecer en la ponderación, pueden existir casos (por ejemplo, generalizaciones cuyo antecedentes puedan ser categorías «sospechosas» como aquellas basadas en cuestiones de género) en los cuales debemos inclinarnos a descartar, *a priori*, esa clase de generalizaciones por —paradójicamente— la existencia de una generalización «compensatoria» (SCHAUER, 2003: 151-154) que nos indica que esa clase de generalizaciones suelen camuflar consideraciones de otras clases —i.e., prácticas discriminatorias— que nos llevarían a resultados *moralmente inadecuados*³⁹. De este modo, un *tercer criterio* podría ser: (iii) las generalizaciones relevantes pueden ser desplazadas cuando existan razones fuertes para sospechar que su utilización nos llevará a conclusiones moralmente inadecuadas⁴⁰.

4. También hemos visto que, muchas veces, las generalizaciones utilizadas en los razonamientos sobre hechos permanecen *in pectore* de la persona que formula el

³⁹ También respecto de la exclusión por razones no necesariamente epistémicas, DAHLMAN, 2016: 92-93.

⁴⁰ Un ejemplo tal vez permita exponer mejor la idea. Asumamos, solo para explicar el punto, que la generalización «las mujeres son amas de casa» cuenta con apoyo empírico, es decir, describe adecuadamente la realidad. El criterio bajo análisis nos indicaría que esta generalización, si bien es válida, podría ser el producto de ciertas prácticas históricas discriminatorias hacia las mujeres, que le impusieron a lo largo de mucho tiempo ese rol. En consecuencia, si nos valiésemos de esa generalización estaríamos validando la práctica que dio lugar a esa configuración del mundo y, por lo tanto, tendríamos buenas razones morales para rechazar su utilización.

argumento, esto es, no se explicitan total o parcialmente. Esto es lo que se denomina un razonamiento entimemático y que, en el campo de la prueba, podríamos denominar *entimema probatorio* (BONORINO, 2015). Sin embargo, también es cierto que es prácticamente imposible explicitar con precisión todos los pasos y componentes de nuestros argumentos. De hecho, en general en nuestro día a día, solemos enunciar las garantías que nos permiten dar el paso de los datos a la conclusión cuando somos desafiados a hacerlo (TOULMIN, 2007: 136; ANDERSON; 1999: 456; ANDERSON, SCHUM y TWINING, 2015: 140; FERRER BELTRÁN, 2007: 133). Siendo conscientes de ello, pero también de que en los procesos judiciales las consecuencias son mucho más graves que en nuestra vida diaria, podría pensarse que se reduciría el riesgo inherente al uso de generalizaciones si se las explicita para ser objeto de análisis y crítica. De este modo, se formularía un cuarto criterio que establezca que: (iv) las generalizaciones que emplean las partes y el juez deben ser explicitadas en la mayor medida posible; en particular, cuando (a) ha existido contradicción entre las partes, directa o indirectamente, sobre alguna/s de ella/s, y (b) cuando un decisor sobre los hechos epistémicamente responsable⁴¹ estime que la/s generalización/es involucradas posee/n una importancia particular⁴².

5. Asimismo, existen casos en los que las generalizaciones necesariamente ingresan al proceso a través de testigos expertos, por su fundamento eminentemente técnico-científico. En estos supuestos, es claro que se acepta la posibilidad de que las generalizaciones sean objeto de la actividad probatoria, es decir, que se produce prueba para acreditar la existencia de tal o cual generalización.

Sin embargo, también podría darse el caso en el cual, ya sea por la contradicción entre las partes o por la responsabilidad epistémica del juzgador de los hechos, se produzca prueba auxiliar para sostener la plausibilidad de una generalización determinada («prueba sobre la prueba», en los términos de FERRER BELTRÁN, 2007: 133). Muchas veces esto puede ser muy sencillo si se trata de un antecedente con un grupo reducido de individuos, esto es, cuando el grado de *generalidad* sea menor. Pero también es posible que esa clase de prueba sea de muy difícil, o imposible, concreción.

Esta circunstancia se suele dar en aquellas generalizaciones sintético-intuitivas, en la terminología de ANDERSON, SCHUM y TWINING. A pesar de eso, muchas de esas generalizaciones parecen ser consideradas, desde las intuiciones de quienes las utilizan, como correctas. A su vez, esas intuiciones parecerían llevarnos a pensar que la generalización, en verdad, no es inválida —en el sentido de falta de apoyo empírico— sino, únicamente, que no se realizó la actividad epistémica necesaria para constatarla.

⁴¹ Aquí se trata de comenzar a pensar el vínculo entre ciertas virtudes que deberían tener los decisores epistémicamente responsables y el análisis de las generalizaciones usadas en los casos concretos. Véase, entre muchos otros, AMAYA, 2015.

⁴² De un modo similar, TARUFFO (2006: 265) señala que resultaría impracticable la explicitación de cada uno de los pasos de nuestros razonamientos y, en esa medida, no habría un deber especial para explicitar la elección de una máxima de experiencia «si es evidente el fundamento que la misma encuentra en el patrimonio de conocimientos del sentido común».

De este modo, un quinto criterio podría ser: (v) las generalizaciones pueden ser objeto de actividad probatoria para demostrar su validez, pero también (a) pueden existir casos en los que pueda argumentarse y sostenerse que una generalización cuyo fundamento se desconoce sea aceptada como válida, (a1) siempre y cuando no sea confrontada por otra generalización válida, (a2) y no se acredite su falta de apoyo empírico.

6. Un último punto se vincula con el modo en que se deberían controlar, en el marco de la etapa recursiva, la utilización de las generalizaciones durante el proceso de conocimiento. Aquí pueden detectarse dos interrogantes: ¿todas las generalizaciones deben ser controladas? ¿cómo *rellenar* los huecos en los entimemas probatorios? La primera de las preguntas podría remitir a consideraciones similares a las expuestas para esbozar el cuarto criterio, pero vinculadas con los motivos de agravio de las partes recurrentes. Sin embargo, la segunda –que también podría trasladarse a la situación del juez del juicio– requiere alguna precisión adicional.

En esta dirección, una herramienta que se presenta, en principio, fructífera es el *principio de caridad*, que «insta a interpretar los textos argumentativos optando por la versión en la que sus argumentos resulten más sólidos» (BONORINO, 2015: 55). A pesar de que, a primera vista, puede resultar una herramienta adecuada, debe ser compatibilizada con el instituto de la carga de la prueba, o si se quiere ser más específico, con los de carga de producción y carga de persuasión, dependiendo del proceso en que se trata. Esto se debe a que, tal vez, una aplicación irrestricta o a modo de regla, podría llevar a resultados indeseados a la luz de ciertos principios rectores, como ser la presunción de inocencia. De tal modo, un sexto criterio puede formularse así: (vi) las generalizaciones deben ser controladas (a) en la medida en que sea materia de agravio por las partes recurrentes o (b) que un juez de revisión epistémicamente responsable estime que la/s generalización/es involucrada/s posee/n una importancia particular, y (c) los entimemas probatorios deben ser completados apelando al principio de caridad, a menos que (c.1) esto pueda resultar contrario a los principios de la carga de la prueba o a otros principios fundamentales.

6. CONCLUSIONES

La primera conclusión que se desprende de lo expuesto es que el término «máximas de experiencia» podría traer más problemas que soluciones. El entramado conceptual que lo ha rodeado decantó en una confusión acerca de su estructura y su función en el razonamiento probatorio. En esa medida, quizás en este caso valga la pena abandonar por completo el término y avanzar, de aquí en adelante, de la mano de las generalizaciones.

La segunda conclusión que podemos extraer de lo dicho hasta aquí es que, tal como señala TWINING (2006: 334), las generalizaciones son *necesarias pero peligrosas*.

Son necesarias porque son el «pegamento» que mantiene nuestros argumentos unidos. No podemos prescindir de ellas para razonar sobre hechos a partir de pruebas. Constituyen la garantía que nos permite dar el paso entre la información con la que contamos y la conclusión que pretendemos hacer valer como verdadera.

Sin embargo, como también vimos, son peligrosas porque pueden traer aparejados problemas de invalidez de la generalización y peligros en el uso de la generalización. En esta dirección, pueden provocar razones inválidas (falsas) o ilegítimas para aceptar una conclusión.

Uno de los principales problemas por los cuales estos riesgos pueden pasar desapercibidos es porque, en ocasiones, las generalizaciones que constituyen las garantías de los argumentos no son expresadas y se suele razonar de modo entimemático.

Las personas, en la vida cotidiana, suelen exteriorizar las generalizaciones que subyacen a sus inferencias cuando son desafiadas a justificar sus conclusiones. Sin embargo, en el marco de un proceso judicial preocupado por la búsqueda de la verdad, no deberían aceptarse, de modo inconsciente y en todos los casos, los entimemas probatorios, esto es, los que implican un razonamiento sobre los hechos y su prueba.

Con esta precaución, es posible detectar si nos encontramos frente a alguna clase de entimema y buscar, de ser necesario, la explicitación de la premisa conformada por la generalización, para así poder controlarla e identificar posibles peligros.

En términos ideales, deberíamos indagar acerca de todos los componentes necesarios para descartar, en primer lugar, las generalizaciones espurias y, a continuación, poder hacer una evaluación sobre su potencialidad epistémica, para lo que deberíamos conocer la fuerza de la generalización, la posibilidad de generalizaciones en competencia, etc.

El camino de las generalizaciones es largo, complejo y, en gran medida, inexplorado. Pero, aun así, si nos preocupa la búsqueda de la verdad de los hechos, debemos recorrerlo a paso firme.

7 BIBLIOGRAFÍA

- ACCATINO, D., 2019: «Teoría de la prueba: ¿somos todos “racionalistas” ahora?», en *Revus* [Online], 39.
- AGUILERA, E., 2016: «Jordi Ferrer y la tradición racionalista de la prueba jurídica: una mirada crítica», en *Isonomía*, 44: 163-189.
- , 2020: «Una propuesta de aplicación de la epistemología jurídica en la investigación del delito», en FERRER BELTRÁN, J. y VÁZQUEZ, C., *Del derecho al razonamiento probatorio*. Madrid: Marcial Pons: 17-44.
- ALCHOURRÓN, C., 2000: «Sobre derecho y lógica», en *Isonomía*, 13: 11-33.
- AMAYA, A., 2015: «Coherencia, virtud y prueba en el derecho», en PÁEZ, A. (coord.), *Hechos, evidencia y estándares de prueba. Ensayos de epistemología jurídica*. Bogotá: Ediciones Uniandes.
- ANDERSON, T., SCHUM, D., & TWINING, W., 2015: *Análisis de la prueba*. Traducción de Flavia Carbonell y Claudio Agüero. Madrid: Marcial Pons.

- ARENA, F., 2016: «Los estereotipos normativos en la decisión judicial. Una exploración conceptual», en *Revista de Derecho de la Universidad Austral de Chile*, Vol. XXIX, n° 1: 51-75.
- BACIGALUPO, E., 1994: *La impugnación de los hechos probados en la casación y otros estudios*. Buenos Aires: Ad-Hoc.
- BEX, F. J., 2011: *Arguments, stories and criminal evidence: A formal hybrid theory*. Dordrecht: Springer.
- BONORINO, P., 2015: «Entimemas probatorios», en *Doxa. Cuadernos de Filosofía del Derecho*, 38: 41-71.
- CAFFERATA NORES, J. I., 1998: *La prueba en el proceso penal*. 3ª edición, Buenos Aires: Depalma.
- CHIOVENDA, G., 1977: *Principios de derecho procesal civil. II*. Traducción de J. Casals Santaló. Madrid: Instituto Editorial Reus.
- COLOMA CORREA, R. & AGÜERO SAN JUAN, C., 2014: «Lógica, ciencia y experiencia en la valoración de la prueba», en *Revista Chilena de Derecho*, vol. 41, núm. 2: 673-703.
- COPI, I. M. & COHEN, C., 2013: *Introducción a la lógica*. Traducción de Jorge Alejandro Rangel Sandoval. México: Limusa.
- COUTURE, E. J., 1958: *Fundamentos de derecho procesal civil*. 3ª edición, Buenos Aires: Depalma.
- DE LA OLIVA SANTOS, A.; ARAGONES MARTÍNEZ, S.; HINOJOSA SEGOVIA, R.; MUERZA ESPARZA, J.; TOMÉ GARCÍA, J. A., 2007: *Derecho procesal penal*. 8º edición. Madrid: Editorial Universitaria Ramón Areces.
- DEI VECCHI, D., 2017: *La decisión de encarcelar preventivamente y otros peligros procesales: una aproximación a la prisión preventiva desde la óptica de la justificación de decisiones judiciales*. Buenos Aires: Ad-Hoc.
- DEVIS ECHANDÍA, H., 1981: *Teoría general de la prueba judicial*, Tomo I, Buenos Aires: Víctor P. de Zavalía Editor.
- FERRAJOLI, L., 1995: *Derecho y Razón. Teoría del garantismo penal*. Madrid: Trotta.
- FERRER BELTRÁN, J., 2007: *La valoración racional de la prueba*. Madrid: Marcial Pons.
- , 2017: «El control de la valoración de la prueba en segunda instancia. Inmediación e inferencias probatorias», en *Revus*, 33: 107-126.
- GASCÓN ABELLÁN, M., 2010: *Los hechos en el derecho. Bases argumentales de la prueba*. Madrid: Marcial Pons.
- GONZÁLEZ LAGIER, D., 2013: *Quaestio facti. Ensayos sobre prueba, causalidad y acción*. México: Fontamara.
- , 2018: «¿Es posible formular un estándar de prueba preciso y objetivo? Algunas dudas desde un enfoque argumentativo de la prueba». *Ponencia presentada en el Congreso Mundial de Razonamiento Probatorio*, Universitat de Girona, Girona (Recuperado de: https://www.academia.edu/37549075/Es_posible_formular_un_est%C3%A1ndar_de_prueba_preciso_y_objetivo_Algunas_dudas_desde_un_enfoque_argumentativo_de_la_prueba).
- , 2020: «¿Qué es el fundherentismo y qué puede aportar a la teoría de la prueba en el Derecho?» (Borrador). Recuperado de: https://www.academia.edu/42333749/Qué_es_el_fundherentismo_y_qué_puede_aportar_a_la_teor%C3%ADa_de_la_prueba_en_el_Derecho_borrador
- HEMPEL, C. G., 1983: *Filosofía de la ciencia natural*. Traducción de Alfredo Deaño. Madrid: Alianza Editorial.
- HERNÁNDEZ MARÍN, R., 2013: *Razonamientos en la sentencia judicial*. Madrid: Marcial Pons.
- HORVITZ LENNON, M. y LÓPEZ MASLE, J., 2004: *Derecho procesal penal chileno*. Tomo II. Santiago de Chile: Editorial Jurídica de Chile.
- IGARTUA SALAVERRÍA, J., 1996: *Valoración de la prueba, motivación y control en el proceso penal*. Valencia: Tirant Lo Blanch.
- , 2004: *El Comité de Derechos Humanos, la Casación Penal española y el control del razonamiento probatorio*. Pamplona: Civitas.123
- LLUCH, A., 2012: *Derecho probatorio*. Barcelona: José María Bosch Editor.
- MONTERO AROCA, J., 2008: *Proceso penal y libertad. Ensayo polémico sobre el nuevo proceso penal*. Navarra: Thomson Civitas.
- MUÑOZ SABATÉ, L., 1967: *Técnica probatoria. Estudio sobre las dificultades de prueba en el proceso*. Barcelona: Editorial Praxis S.A.

- NIEVA FENOLL, J., 2010: *La valoración de la prueba*. Marcial Pons, Madrid.
- NOBILI, M., 1969: «Nuove polemiche sulle cosiddette “massime d’esperienza”», en *Rivista italiana di diritto e procedura penale*, 12: 123-193.
- PALACIO, L. E., 2002: *La prueba en el proceso penal*. Buenos Aires: Abeledo Perrot.
- PARRA QUIJANO, J., 2007: *Manual de derecho probatorio*. 16ª edición. Bogotá: Librería Ediciones del Profesional LTDA.
- ROBERTS, P. & ZUCKERMAN, A., 2004: *Criminal evidence*. Oxford: Oxford University Press.
- ROXIN, C., 2000: *Derecho procesal penal*. Traducción de Gabriela E. Córdoba y Daniel R. Pastor, revisada por Julio B. J. Maier. Buenos Aires: Editores del Puerto.
- SCHAUER, F., 2003: *Profiles, probabilities and stereotypes*. Cambridge, Mass.-London: Harvard University Press.
- , 2004: *Las reglas en juego. Un examen filosófico de la toma de decisiones basadas en reglas en el derecho y en la vida cotidiana*. Madrid: Marcial Pons.
- SENTÍS MELENDO, S., 1979: *La prueba: los grandes temas del Derecho Probatorio*. EJE, Buenos Aires.
- SERRA DOMÍNGUEZ, M., 1969: «Contribución al estudio de la prueba», en *Estudios de Derecho Procesal*. Barcelona: Ariel.
- SCHUM, D., 1984: *Evidential Foundations of Probabilistic Reasoning*. Illinois: Northwestern University Press.
- STEIN, A., 2005: *Foundations of evidence law*. Oxford: Oxford University Press.
- STEIN, F., 1893: *El conocimiento privado del juez*. Traducción al español de Andrés De la Oliva Santos. Bogotá: Temis, 1988.
- TARUFFO, M., 2002: *La prueba de los hechos*. Traducción de Jordi Ferrer Beltrán. Madrid: Trotta.
- , 2006: La motivación de la sentencia civil. Traducción de Lorenzo Córdoba Vianello. México: Tribunal Electoral del Poder Judicial de la Federación.
- , 2009: «Consideraciones sobre las máximas de la experiencia», en *Páginas sobre justicia civil*. Traducción de Maximiliano Aramburo Calle. Madrid: Marcial Pons: 439-453.
- , 2012: *Proceso y decisión. Lecciones mexicanas de derecho procesal*. Madrid: Marcial Pons.
- TOULMIN, S., 2007: *Los usos de la argumentación*. Traducción de María Morrás y Victoria Pineda. Barcelona: Península.
- TWINING, W., 2006: «Narrative and Generalizations in Argumentation about Questions of Fact», en *Rethinking Evidence. Exploratory Essays*. 2º ed. Cambridge: Cambridge University Press: 332-343.
- UBERTIS, G., 2014: «Comparative perspectives: truth, evidence and proof in criminal proceedings», en *International Journal of Procedural Law*, Vol. 4, Núm. 1: 79-96.
- , 2017: *Elementos de epistemología del proceso judicial*. Traducción de Perfecto Andrés Ibáñez. Madrid: Trotta.
- VÉLEZ MARICONDE, A., 1982: *Derecho procesal penal*. Tomo 1. 3ª edición, 1ª reimpresión. Córdoba: Marcos Lerner Editora Córdoba.

LA DECISIÓN SOBRE LA QUAESTIO FACTI EN LOS ACUERDOS DE CULPABILIDAD*

Diana Veleda

Universidad de Buenos Aires
Universidad Torcuato Di Tella
diana_veleda@hotmail.com

RESUMEN: La autora discute una posición relativamente extendida en el ámbito del Derecho Procesal Penal, según la cual la implementación de procedimientos de omisión del juicio penal o acuerdos de culpabilidad implica, necesariamente, la adopción de una noción de verdad distinta a la correspondencia con la realidad. Para ello, primero defiende la vigencia y la importancia de ese concepto de verdad en estos mecanismos y, luego, argumenta que existen razones morales, independientes a la búsqueda de la verdad, que justifican exigir a la fiscalía, a la hora de proponer un acuerdo de culpabilidad, que satisfaga el mismo estándar probatorio necesario para requerir la realización de un juicio ordinario. Finalmente, reflexiona sobre las implicancias que aquella exigencia, junto con otras características propias de este tipo de procedimientos, tienen en el mejoramiento de las condiciones epistémicas en las que tomamos la decisión sobre la *quaestio facti* en el contexto de los acuerdos de culpabilidad.

PALABRAS CLAVE: acuerdos de culpabilidad, verdad como correspondencia, verdad consensual, coerción, estándares probatorios.

DECIDING ABOUT FACTS IN TRIAL AVOIDING MECHANISMS

* Agradezco a Máximo LANGER y Matías DÍAZ, por el acompañamiento y los consejos de siempre, y a Pablo ROVATTI y Pablo LARSEN, por su lectura y comentarios.

ABSTRACT: The author challenges a doctrine that, to some extent, has been well accepted in the science of criminal procedure. According to such doctrine, the plea bargain and other mechanisms implemented to avoid trial embrace a concept of truth that is different from the concept upheld by correspondence theories. As opposed to such doctrine, the author reaffirms the validity and importance of a correspondence concept of truth in the context of the mechanisms to avoid trial. The author then suggests that there are moral reasons —not derived from the general interest in seeking the truth— that would require a prosecutor willing to propose a plea deal to have the same amount of evidence that would be necessary to move forward with a regular trial. The author ultimately envisage the consequences of applying the proposed standard, along with other specific characteristics of these proceedings, towards the enhancement of the epistemic conditions, under which we decide about facts in trial avoiding contexts.

KEYWORDS: trial avoiding mechanisms, correspondence theories, consensual theories, coercion, standards of proof.

SUMARIO: 1. INTRODUCCIÓN.— 2. LA VERDAD EN LOS ACUERDOS, ¿UNA PREOCUPACIÓN FUERA DE LUGAR?. 2.1. ¿Una verdad consensuada?. 2.2. Sobre la adecuación descriptiva de la noción de verdad consensuada en los acuerdos. 2.3. Sobre la adecuación normativa de la noción de verdad consensuada en los acuerdos.— 3. SOBRE LAS CONDICIONES DE LEGITIMIDAD MORAL DE LOS ACUERDOS DE CULPABILIDAD: LA MEDIDA DE LA PENA Y LA SUFICIENCIA DE LA PRUEBA. 3.1. ¿Propuestas o amenazas? Sobre la legitimidad de la propuesta de un acuerdo de culpabilidad por parte de la fiscalía. 3.2. La medida de la pena: recompensas vs. penalidades. 3.3. La suficiencia de la prueba: el estándar necesario para llevar el caso a un juicio ordinario.— 4. REPERCUSIONES EN LA DECISIÓN SOBRE LA *QUAESTIO FACTI* EN LOS ACUERDOS DE CULPABILIDAD. 4.1. Un determinado nivel de comprobación de la culpabilidad del acusado. 4.2. Un particular contexto de justificación: la inexistencia de hipótesis rivales propuestas por la defensa. 4.3. Una manifestación de conformidad, en principio, epistémicamente confiable.— 5. CONCLUSIONES.— 6. BIBLIOGRAFÍA

RECOMMENDED CITATION: VELEDA, D., 2020: «La decisión sobre la *quaestio facti* en los acuerdos de culpabilidad», in *Quaestio facti*, 2: 155-184. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22461

1. INTRODUCCIÓN

De un tiempo a esta parte, existe un consenso más o menos generalizado en el ámbito del derecho probatorio acerca de que la decisión sobre los hechos de un caso penal debe estar guiada por criterios de racionalidad epistemológica¹. Es cierto que los juicios podrían terminar sin haber alcanzado la *verdad*², pero siempre es deseable

¹ Muchas obras son representativas de esta corriente epistemológica en el derecho probatorio. Solo por mencionar algunas, véase TARUFFO, 1992, FERRER BELTRÁN, 2007 y 2005; GASCÓN ABELLÁN, 2010; ALLEN, 2010; AMAYA, 2010 y LAUDAN, 2006.

² En razón de la vigencia del principio de inocencia y la intermediación de un estándar de prueba exigente para la comprobación de la culpabilidad, el proceso penal puede culminar normalmente aun-

que sus resoluciones definitivas reflejen lo sucedido en el mundo real —esto es, principalmente, que condenemos a los culpables y absolvamos a los inocentes—³. Más aun, en el caso de las condenas, su legitimidad y la de sus consecuencias dependen de ello⁴. De allí que, dado que el criterio que utilizamos para justificar la veracidad o falsedad de enunciados los fácticos en un proceso penal es la prueba⁵, el sentido de las decisiones que lo componen y concluyen debería estar guiado por el grado de corroboración que la hipótesis acusatoria haya obtenido a partir de las evidencias recolectadas⁶.

Sin embargo, la gran mayoría de las legislaciones procesales penales del mundo permite emitir un pronunciamiento sobre la responsabilidad penal de una persona con prescindencia de la producción de la prueba en el debate y a partir de alguna clase de acuerdo entre la fiscalía y el acusado y su defensa⁷. Estos mecanismos se encuentran generalmente orientados hacia fines utilitarios —tales como la simplificación o abreviación de los procesos—⁸ y, al margen de su mayor o menor éxito en la consecución de esos objetivos⁹, suelen ser criticados por razones relacionadas con

que la hipótesis triunfante no sea la que goza de mayor apoyo empírico. Si la hipótesis de culpabilidad no alcanza el estándar necesario para la emisión de una condena, «se presumirá la verdad de la hipótesis menos confirmada (*i.e.*, la de la inocencia)» (FERRER BELTRÁN, 2010: 98).

³ Para FERRER BELTRÁN, la investigación judicial pretende decir algo sobre el mundo y, por ello, intenta determinar el valor de verdad de determinadas premisas fácticas (2010: 81-84).

⁴ «En un sistema donde el Estado asume la función del control social por medio de la imposición de penas se encuentra, entre otras condiciones de legitimidad, la necesidad de que los elementos que permiten afirmar la existencia de un hecho punible estén verificados con certeza o probados más allá de la duda razonable. Como sucedáneo de esta regla, existe el deber del Estado y de sus representantes en el foro penal de investigar y determinar la *verdad* objetiva, y también, de pronunciarse a favor del acusado si no es alcanzada esta *verdad* más allá de la duda razonable. En otras palabras, existe la imposibilidad, desde un punto de vista normativo, de aplicar una pena sin tener acreditados con un determinado grado de precisión los hechos que habilitan tal potestad estatal» (DÍAZ, 2015: 103).

⁵ V., por todos, FERRER BELTRÁN, 2010: 82-83 y, en el mismo sentido, 2005: 29-38.

⁶ En este sentido, «la fijación judicial de los hechos debe recurrir necesariamente a los esquemas propios de la epistemología general» (GASCÓN ABELLÁN, 2010: 46) y «una hipótesis se considera verdadera (correspondiente con la realidad) cuando ha sido probada, y ha sido probada cuando, tras la valoración de la inferencia probatoria, alcanza un alto grado de credibilidad» y «los criterios de valoración de la prueba son paralelos a los criterios de confirmación de hipótesis propuestos por algunos filósofos de la ciencia» (GONZÁLEZ LAGIER, 2003: 49).

⁷ En un reciente trabajo empírico, LANGER muestra que desde 1970 el *plea bargaining* y otros mecanismos de omisión del juicio penal fueron implementados, al menos, en sesenta jurisdicciones a lo largo del mundo (2019: 3-10).

⁸ Se ha argumentado también que esta clase de procedimientos pueden hallar sustento en razones no-utilitarias o de *justicia*, tales como la necesidad de mitigar el castigo a quien admite su culpabilidad (DUFF, FARMER, MARSHALL y TADROS, 2007: 174).

⁹ No contamos con mucha información que permita juzgar concluyentemente si los procedimientos abreviados han colaborado con los fines que justificaron su implementación, pero los datos disponibles respecto del caso de la justicia nacional en la República Argentina no son alentadores. Un relevamiento estadístico determinó que «[e]n términos de la duración del proceso, el 20% de los procedimientos abreviados se resuelve en menos de dos meses y otro 20% entre 2 y 4 meses desde que el caso llega ante el tribunal oral. La mayoría de los caos (60%) toman más de 4 meses y algunos hasta 1

la justicia de las decisiones que de ellos resultan. En particular, se los critica porque facilitarían la punición de inocentes (FERRAJOLI, 1989: 748) e incitarían a la disminución de la pena justa de los autores culpables (SANCINETTI, 2010).

Con respecto a lo primero, los acuerdos facilitarían el castigo de inocentes, por un lado, porque serían procedimientos *coercitivos*¹⁰ que provocarían que los imputados decidan prestar conformidad con la propuesta de la fiscalía con el único fin de evitar posibles males mayores, tales como la aplicación de una sanción mucho más grave en el juicio ordinario¹¹. En segundo lugar, porque al prescindir de la actividad probatoria dificultarían la posibilidad de detectar esas eventuales falsedades. De hecho, sin evidencia a la vista y si ello sirve a sus intereses, la fiscalía y el imputado podrían acordar la existencia de un evento —y, con ello, prescindir de probarlo—, cuando en rigor aquel no existe o no ha ocurrido de ese modo¹².

Pese a todo esto, en muchas jurisdicciones, los acuerdos son la manera en la cual se procesa una gran parte de los casos que ingresan al sistema de justicia penal¹³. Una posible explicación de este fenómeno es que, en rigor, los acuerdos implicarían un abandono de la noción de verdad como correspondencia con la realidad y funcionarían a la luz de una noción de verdad *consensual*, según la cual la corrección de los enunciados fácticos no vendría determinada por su mayor o menor aproximación a lo ocurrido en el mundo, sino por la decisión misma de las partes de no controvertir los hechos del caso.

o 2 años desde que el caso llega ante el tribunal oral. El promedio de los casos que se resuelven por esta vía toman 301 días y la mediana es de 165 días desde que el caso llega ante el tribunal oral» (BERGMAN, FONDEVILA y LANGER, 2017: 62).

¹⁰ Solo por citar algunas opiniones en este sentido, para LANGBEIN, «[l]as semejanzas entre el sistema estadounidense moderno de *plea bargaining* y el antiguo sistema de tortura judicial son muchas y escalofrantes» (1978: 9) y para DÍAZ CANTÓN, «toda renuncia o conformidad [del acusado], para ser tal, precisa de la libertad total de decisión, inexistente por definición en [el] juicio abreviado», ya que «[e]s ilusorio pensar que la coerción puede ser eliminada del proceso, ya que la misma existencia del proceso la implica necesariamente» (2005: 252-253).

¹¹ Aquí nos interesa especialmente esta consecuencia indeseable del presunto carácter coercitivo de los acuerdos, vinculada con el incremento del riesgo de error en la decisión sobre los hechos, pero esa condición es problemática en sí misma, en razón de la garantía que prohíbe al Estado obligar al imputado a realizar declaraciones contra sí mismo. Esto, claro está, siempre que se acepte que esta garantía no tiene un valor meramente instrumental o epistémico —es decir, orientado únicamente a prevenir declaraciones falsas— sino también un valor en sí misma.

¹² De ahí que las implicancias del rol probatorio de las partes sea una de las peculiaridades que distingue al proceso judicial de otros contextos epistémicos (véase, en este sentido, FERRER BELTRÁN, 2010: 90 y ss. y TARUFFO, 2010: 194-195). Por ello se considera importante que, en aquellos sistemas que priorizan el control de las partes sobre la investigación de los hechos, sean desarrollados incentivos apropiados para contrarrestar esa tendencia a incursionar en comportamientos distorsionadores de la verdad (DAMAŠKA, 2003: 120).

¹³ LANGER analiza la relación existente entre el número de condenas obtenidas a través de mecanismos de evitación del juicio y el número total de condenas en veintiséis jurisdicciones del mundo y, en la mayoría de ellas, la tasa resultante excede el 50 por 100 de los casos (2019: 27 y ss.).

En este trabajo sostendremos que, por el contrario, la resolución de casos penales a través de acuerdos de culpabilidad no nos compromete, *necesariamente*, con la adopción de una noción distinta a la de la verdad como correspondencia con la realidad. En particular, afirmaremos que los acuerdos deben observar determinadas exigencias independientes de la búsqueda de la verdad, vinculadas con su legitimidad moral —o, si se quiere, con el objetivo de evitar la violación de los derechos del acusado en ese contexto—, y que la observancia de esas exigencias repercute en las condiciones en las que adoptamos la decisión sobre la *quaestio facti*. Más precisamente, afirmaremos que aquellas facilitan que la decisión sobre los hechos goce de una genuina justificación epistémica, esto es, sobre la base de prueba y de acuerdo con un concepto de verdad como correspondencia con la realidad.

Para ello, descartaremos primero una objeción relacionada con la posible falta de coherencia interna de nuestra propuesta, según la cual, dado el origen y objetivos de los acuerdos, la preocupación por la averiguación de la verdad —entendida como correspondencia con la realidad— se encontraría fuera de lugar (punto 2). En segundo lugar sostendremos que, para que la propuesta de un acuerdo no sea *coercitiva* ni viole los derechos del acusado, la fiscalía no puede imponer a aquel una penalidad por elegir ir a un juicio ordinario y, además, debe contar con una sospecha suficiente en su contra, en particular, aquella que le permitiría solicitar la realización del debate público (punto 3). En tercer lugar, analizaremos la manera en la cual la observancia de esas condiciones de legitimidad moral de los acuerdos puede repercutir en la decisión sobre la *quaestio facti* (punto 4). Finalmente, ensayaremos unas breves conclusiones (punto 5).

2. LA VERDAD EN LOS ACUERDOS, ¿UNA PREOCUPACIÓN FUERA DE LUGAR?

2.1. ¿Una verdad *consensuada*?

El enfoque de este trabajo y su propuesta podrían merecer una objeción preliminar, relacionada con la aparente inconsistencia entre una preocupación fuerte por la averiguación de la verdad y la naturaleza y los objetivos de los acuerdos de culpabilidad.

Como ya adelantamos, estos procedimientos se apoyan en la ausencia de contradicción o en el consenso al que las partes arriban sobre ciertos aspectos del caso, con el fin de evitar la producción de prueba que estiman innecesaria, al manifestarse seguras sobre cuál sería su resultado. De hecho, la mayor parte de los sistemas que contienen la posibilidad de arribar a acuerdos fácticos en casos penales tienen su origen en el *plea bargaining* anglosajón y, allí, el reconocimiento de la culpabilidad por parte del acusado a través del *guilty plea* tiene una significación peculiar: constituye una auténtica cancelación de la controversia sobre los hechos y con ella —en princi-

pio— deviene innecesaria la comprobación del contenido fáctico de la declaración¹⁴. Esto es connatural a un proceso concebido como *adversarial*, o bien como una *disputa* entre partes iguales que, además, son dueñas de la controversia que desarrollan (LANGER, 2005: 119)¹⁵, en el que el objetivo de la resolución de un conflicto prevalece, en términos generales, a la meta de la averiguación de la verdad (DAMAŠKA, 1997: 120)¹⁶. De allí que se considere que la noción de verdad, en dicho contexto, resulte «más [relativa] y consensual: si las partes llegan a un consenso sobre los hechos del caso, mediante acuerdos sobre declaraciones de culpabilidad o estipulaciones, no es tan importante determinar cómo ocurrieron los hechos realmente» (LANGER, 2018: 39-40).

Estas particularidades de la decisión sobre los hechos en los acuerdos de culpabilidad podría explicarse, no tanto por la mayor o menor importancia que se confiere, en ese contexto, a la verdad como correspondencia con la realidad¹⁷, sino por la adopción, en su reemplazo, de un concepto de verdad diferente, de naturaleza con-

¹⁴ Así, «analizado desde el modelo de la disputa [el *guilty plea*] cobra completo sentido», pues «si el proceso penal no es sino una contienda entre acusador y acusado, solo existe una controversia si este último se declara no culpable, mientras que si el imputado se declara culpable, ya no existe disputa alguna sobre la que el juez o el jurado puedan resolver, ya que el acusado se ha “allanado” a la pretensión del acusador, y entonces solo resta determinar la pena» (LANGER, 2005: 119).

¹⁵ El proceso penal estadounidense se acerca a lo que LANGER denomina como el modelo procesal *de la disputa*, en el que (i) el proceso penal es una lucha entre partes iguales —la fiscalía, por un lado, el imputado y su defensa, por el otro—, frente un tercero imparcial —el juez—; (ii) las partes son dueñas de la contienda, pues pueden disponer del contenido del litigio; (iii) en general, hay dos investigaciones paralelas que avanzan según el impulso que recibieren de cada parte y (iv) el objetivo del proceso es, al menos predominantemente, la resolución de un conflicto (LANGER, 2005: 114-123).

Es el dominio del caso en manos de las partes, como elemento característico de los modelos adversariales, lo que resulta central para este trabajo. Este elemento se encuentra presente en las diversas conceptualizaciones existentes sobre este tipo de modelo procesal. Puede leerse en DAMASKA (1997: 87 y siguientes) o en el *adversarial legalism* de KAGAN. En efecto, este consiste en un método de contradicción legal entre disputantes que invocan intereses contrapuestos y que se caracteriza por el *activismo* de los *litigantes*. En sus palabras, se trata de «un estilo de contradicción jurídica en el que la afirmación de demandas, la búsqueda de argumentos legales dominantes y la recolección y presentación de evidencia están dominadas no por los jueces o por oficiales estatales, sino por las partes que disputan intereses contrapuestos, de manera tal que, a través del *plea bargaining*, el *adversarial legalism* «permite a las partes disputantes amenazar a sus adversarios con costos muy altos y demoras, lo cual alienta la resolución de los casos a través de negociaciones informales (y muchas veces extorsivas)» (2001: 9 y 84, mi traducción desde el original).

¹⁶ En efecto, «el omnipresente fin de la resolución del conflicto constituye un argumento silencioso a favor del mayor control de las partes sobre el desarrollo del proceso. No cabe duda, por ejemplo, de que la equiparación entre decisión judicial y resolución de conflictos favorece la idea de que solo los hechos controvertidos deben constituir el objeto de la actividad probatoria. Las admisiones de culpabilidad, en los procesos penales, y la admisión de hechos, en los civiles, se convirtieron con bastante naturalidad en parientes cercanos» (DAMAŠKA, 1997: 120).

También se sostiene, sin embargo, que la opción entre la resolución de un conflicto y la averiguación de la verdad puede no ser siempre posible. Véase, por ejemplo, WEIGEND, 2003: 169.

¹⁷ Sobre el concepto de verdad como correspondencia con la realidad véase, por todos, GASCÓN APELLÁN, 2010: 51.

sensual¹⁸, según el cual la solución del caso propuesta por las partes podría «apartarse de lo verdaderamente acontecido» (GUZMÁN, 2005: 297)¹⁹. De acuerdo con ello, los enunciados fácticos comprendidos en el acuerdo serían verdaderos porque las partes no los controvierten y no porque se haya corroborado que esa es la manera en la que los eventos se desarrollaron en el mundo a través de la evidencia recolectada u otro criterio de justificación disponible²⁰. En este marco, una preocupación fuerte por la objetiva determinación de los hechos, en términos de correspondencia con la realidad, quedaría fuera de lugar. Discutiremos esta posición en lo que sigue.

2.2. Sobre la adecuación descriptiva de la noción de verdad *consensuada* en los acuerdos

Desde un punto de vista descriptivo, no es claro que los acuerdos de culpabilidad, sea cual fuere su origen, operen completamente divorciados de la noción de verdad como correspondencia con la realidad.

Esto es más evidente en los ordenamientos procesales de tradición europea-continental. Estos sistemas han incorporado los acuerdos de origen anglosajón pero los han traducido²¹ de acuerdo con las particularidades propias del modelo procesal

¹⁸ La idea de verdad *consensual* podría incluirse dentro de las teorías *pragmatistas* de la verdad, según las cuales «la verdad de un enunciado se concibe en términos de “aceptabilidad justificada”. Un enunciado es verdadero si está justificado creer que es verdadero porque sirve a algún fin (versión “instrumentalista”) o porque es aceptado (versión “consensualista”). En suma, un enunciado no está justificado porque sea verdadero, sino que es verdadero porque está justificado, o más exactamente, porque el criterio para aceptarlo como verdadero (el criterio de verificación) está justificado» (GASCÓN ABELLÁN, 2010: 51).

¹⁹ En su análisis del caso del *plea bargaining* estadounidense, BOVINO llega a una conclusión similar. Considera que, «al concederse a la declaración del imputado el mismo valor que al veredicto del jurado, la verdad se da por establecida. El *consenso*, en este contexto, *desplaza*, a la determinación judicial de la “verdad real”» (2005: 63).

²⁰ Es necesario aclarar, sin embargo, que la distinción entre criterios y conceptos de verdad es confusa en las teorías pragmatistas de la verdad (GASCÓN ABELLÁN, 2010: 51) y que, de hecho, la distinción de las nociones de conceptos y criterios de verdad, en general, no es un asunto sencillo. Al respecto, puede consultarse HAACK, 1978: 88-91.

A los fines de este trabajo, nos bastará con afirmar que los acuerdos de culpabilidad no nos obligan a abandonar un concepto de verdad como correspondencia y reemplazarlo por un concepto *consensual* y que, en todo caso, según veremos más adelante, el consenso de las partes o la falta de contradicción de la hipótesis acusatoria podría ser uno de los factores contextuales que modifique las condiciones de justificación de la hipótesis acusatoria, a la luz de un concepto de verdad como correspondencia. En este sentido, el trabajo rechaza la adopción del consenso como concepto de verdad, en sentido extremo, pero no en un sentido moderado, como uno de los *criterios* de justificación de un enunciado con pretensión de verdad. Agradezco a Daniel GONZÁLEZ LAGIER por esta precisión.

²¹ Las ventajas de la metáfora de la *traducción* frente a la noción de los *trasplantes* legales puede leerse en LANGER, 2018: 67-76.

subyacente²². En estos contextos, la preocupación por la averiguación de la verdad conserva un lugar que no es para nada despreciable.

La confesión del acusado en un acuerdo de culpabilidad no es otra cosa que una fuente de conocimiento acerca de los hechos o, mejor dicho, una prueba más de su culpabilidad que, si no es corroborada con otros elementos que respalden su veracidad, podría ser insuficiente para el dictado de una condena²³. De acuerdo con ello, «un acusado no puede poner fin a la etapa de la determinación de culpabilidad al admitir su responsabilidad ante el tribunal. Si bien la admisión de culpabilidad puede ser de mucha utilidad para el juez en la búsqueda de la verdad, lo cierto es que el juez todavía tiene la última palabra al respecto y en consecuencia puede decir: “le creo pero su confesión sola no es prueba suficiente más allá de toda duda razonable de que usted lo hizo”» (LANGER, 2018: 40-41)²⁴.

Pero tampoco el *plea bargaining* anglosajón resigna toda preocupación por la objetiva determinación de los hechos. Por lo pronto, las *Federal Rules of Criminal Procedure* estadounidenses exigen que la declaración de culpabilidad cuente con una *base fáctica* suficiente y ponen en cabeza del tribunal la realización de las indagaciones necesarias para determinar su verificación en cada caso —regla 11.b.3—²⁵.

Es cierto que la noción de base fáctica suficiente parece haber sido dotada de un contenido bastante flexible en la jurisprudencia estadounidense. Una de las principales referencias a la que suele acudir para argumentar que el *plea bargaining* descon-

²² Los sistemas procesales de origen europeo-continental se hallan tradicionalmente guiados por los principios de averiguación de la verdad, oficialidad y legalidad, generalmente asociados con una noción de disputa meramente *formal* entre las partes. A diferencia del caso estadounidense, estos procesos tienen más notas comunes con el modelo usualmente llamado *inquisitivo* o que LANGER denomina como de la *investigación oficial*, en el que (i) existe una única investigación; (ii) esta investigación es impulsada por un oficial del Estado; (iii) ese oficial no tiene un interés concreto en la resolución del litigio sino que actúa con objetividad, y (iv) las partes no pueden disponer materialmente del objeto del proceso (2005: 114-123).

²³ Como afirmamos en otra ocasión (VELEDA, 2019: 280-281), la obtención de un reconocimiento de culpabilidad por parte del acusado también jugó un rol relevante en los procedimientos europeo-continentales, pero lo hizo de un modo bastante diferente al del modelo anglosajón, asociado a la práctica de la tortura y al sistema de pruebas legales (MAIER, 2004: 448 y FOUCAULT, 2014: 223-226).

²⁴ Algunos ejemplos de esta clase de acuerdos, adaptados a la idiosincrasia de los procesos penales de tradición europea-continental, son el caso italiano (véanse, entre otros, LANGER, 2018: 93-103 y MAFFEI, 2004), el alemán (véanse, entre otros, LANGER, 2018: 82-93, DIAS, 2015 y WEIGEND, 2008) y el argentino (v., entre otros, LANGER, 2018: 103-108 y CÓRDOBA, 2005: 229-250).

²⁵ Regla 11.b.3.: “*Determining the Factual Basis for a Plea*. Before entering judgment on a guilty plea, the court must determine that there is a factual basis for the plea”.

La correspondiente nota del *Advisory Comitee* agrega: “*The court should satisfy itself, by inquiry of the defendant or the attorney for the government, or by examining the presentence report, or otherwise, that the conduct which the defendant admits constitutes the offense charged or an offense included therein to which the defendant has pleaded guilty. Such inquiry should, e.g., protect a defendant who is in the position of pleading voluntarily with an understanding of the nature of the charge but without realizing that his conduct does not actually fall within the charge*”.

sidera la verdad como correspondencia es el precedente *North Carolina v. Alford*²⁶. Allí, la Corte Suprema de los Estados Unidos resolvió que no presenta un problema desde una perspectiva constitucional la decisión de un tribunal de aceptar un *guilty plea* acompañado de una declaración de inocencia por parte del acusado. Aunque parece difícil conciliar esta posición con una pretensión dirigida a la objetiva determinación de los hechos, algunas reflexiones de la Corte Suprema de los Estados Unidos muestran que la preocupación por la precisión fáctica de la condena no se encontraba ausente. En particular, ese tribunal se dedicó especialmente a evaluar la trascendencia de la declaración de inocencia del acusado en función de la fortaleza del caso que la fiscalía tenía contra él. Al respecto, afirmó:

Aquí el Estado tenía un caso fuerte de homicidio en primer grado contra Alford. [...] En razón de la abrumadora evidencia en su contra, un juicio era precisamente lo que Alford ni su abogado deseaban. [...] Cuando su declaración de culpabilidad es vista a la luz de la evidencia que existía en su contra, la cual negaba, categóricamente, su proclamación de inocencia y le otorgaba al juez los medios para verificar si la declaración de culpabilidad estaba siendo efectuada de manera inteligente, su validez no puede ser cuestionada con seriedad (mi traducción desde el original).

De acuerdo con ello, aunque pensemos que la condena de Alford fue dictada en condiciones que están lejos de ser las ideales, no es adecuado afirmar que la decisión fue adoptada a la luz de una noción consensual de la verdad o con total despreocupación por la averiguación de lo realmente ocurrido. Las reiteradas referencias a la existencia de sólida evidencia sobre la culpabilidad del acusado muestran que la atención a lo realmente ocurrido está presente en el razonamiento de la Corte y que esto fue tenido en cuenta para restar credibilidad a la declaración de inocencia, todo ello a la luz de un concepto de verdad como correspondencia con la realidad.

También es cierto que existen otros factores, generalmente presentes en las formas robustas de *plea bargaining*, que pueden generar efectos distorsionadores de la verdad. Por ejemplo, en el caso estadounidense, la facultad que pueden tener los fiscales de imputar tipos penales adicionales a los que corresponderían al caso, la posibilidad de negociar la base fáctica de la acusación —*fact bargaining*—, las dificultades que enfrentan los jueces para corregir esas desviaciones o la circunstancia de que los fiscales valoren más la resolución eficiente de los casos que la precisión de sus decisiones, al menos en aquellos procesos que involucran delitos leves (LIPPKE, 2011: 221-223). Sin embargo, esto tampoco nos obliga a abandonar nuestro punto.

Que los acuerdos celebrados en aquellos términos sean un mal camino para la determinación objetiva de lo ocurrido no es equivalente a afirmar que el sistema entienda que los hechos contenidos en ellos ocurrieron de ese modo simplemente porque las partes así lo afirman. En todo caso, las diferentes condiciones de celebración de estos procedimientos —p. ej., la mayor o menor presencia de esos factores distorsivos, la mayor o menor disponibilidad del objeto del caso en manos de la fiscalía y el acusado, la mayor o menor rigidez del control judicial sobre la satisfacción de la

²⁶ “North Carolina v. Alford”, 400 U.S. 25 (1970).

base fáctica— mostrarán una correlativa preocupación, mayor o menor, por el conocimiento de lo realmente ocurrido. En otras palabras, los acuerdos no implican, necesariamente, la adopción de otra noción de verdad, pero sus características pueden reconocerle una importancia mayor o menor a su averiguación, en relación con los restantes valores en juego —p. ej., aquellos asociados a la eficiencia del sistema—. ²⁷

Una argumentación más o menos similar se utiliza para afirmar que la circunstancia de que un proceso sea considerado adversarial no basta para concluir que se encuentra guiado por un concepto distinto al de la verdad como correspondencia con la realidad o que no le interesa conocerla, en medida alguna. Para llegar a esas conclusiones, lo relevante es «si el ordenamiento exige efectivamente que la decisión se base en pruebas genuinas desde el punto de vista epistémico» y, en todo caso, «quedará por discutirse si la medida en que esa regulación propende a descubrir la verdad es o no satisfactoria, si se tutelan o no suficientemente otros valores, si se distribuye con justicia el riesgo de error, etc.; pero no cabrá decir que no le interesa la verdad sin más por el solo hecho de concebirse al proceso como *adversarial*» (DEI VECCHI, 2016: 281). De manera similar, la sola posibilidad de resolver un caso penal a través de un acuerdo de culpabilidad no conlleva, necesariamente, la adopción de una noción de verdad consensual, sino que las particulares circunstancias en las cuales aquellos permitan adjudicar la solución del caso exhibirán el mayor o menor peso conferido al objetivo de arribar a decisiones precisas en el aspecto fáctico. Será fundamental, entonces, determinar si los acuerdos exigen que la decisión definitiva se base en «pruebas genuinas desde el punto de vista epistémico» y en qué medida.

2.3. Sobre la adecuación normativa de la noción de verdad *consensuada* en los acuerdos

Por último, desde una perspectiva normativa, no sería posible afirmar que los acuerdos de culpabilidad se fundan en una noción consensual de la verdad y que, con ella, pueden habilitar soluciones absolutamente ficcionales, sin cuestionar seriamente su legitimidad. La averiguación de la verdad, entendida como correspondencia con la realidad, es una condición de justicia de las decisiones resultantes de un proceso penal en un Estado de derecho ²⁸.

²⁷ Existen muchos otros factores, además de los aquí enumerados, que pueden ser decisivos en el grado de autodeterminación del acusado a la hora de decidir declararse culpable y optar por una alternativa al juicio. Entre ellas, se encuentra la labor de los abogados defensores y su rol en el correcto asesoramiento del imputado acerca de la conveniencia de la herramienta en cada caso. Un análisis sobre la vigencia y el alcance del derecho a la asistencia técnica en el marco del *plea bargaining* estadounidense puede leerse en PETEGORSKY, 2013: 3631-3641. Agradezco a Carmen VÁZQUEZ por esta observación.

²⁸ Es cierto que la adecuación de la decisión sobre los hechos de un caso penal a lo realmente ocurrido puede no garantizar la *justicia* del resultado si, por ejemplo, la norma cuyo antecedente fáctico hemos corroborado es una norma injusta por otras razones —p. ej., fundada en prejuicios de clase, raza o religión—. La noción de justicia es utilizada aquí únicamente en ese primer sentido, relacionado con

De acuerdo con el principio de *verdad* o de la *verdad* como garantía (FERRAJOLI, 1989: 545), la corrección moral de una condena penal depende de la corrección fáctica del juicio de responsabilidad contenido en ella. En otras palabras, «la *verdad*, como correspondencia aproximada de una hipótesis con el objeto que refiere, constituye uno de los requisitos fundamentales que debe contener un sistema de persecución penal adecuado a un Estado de derecho, y respetuoso de la dignidad humana» y su búsqueda «debería aparecer como un objetivo relacionado con la pretensión de aplicar el poder punitivo en contra de una persona» o, más precisamente, «ante la intención de aplicar una pena estatal, debería surgir la obligación de determinar la *verdad* como garantía política del acusado», de manera que «el objetivo del proceso penal podría ser redefinido en términos de corroboración de la veracidad de la hipótesis acusatoria con pretensiones punitivas, en vez de búsqueda de la *verdad* en sí misma» (DÍAZ, 2017: 120-121).

Más allá del proceso penal, la *verdad* también funciona como un ideal regulativo general, pues el éxito del derecho como mecanismo ordenador de las conductas de las personas depende de la posibilidad de que las consecuencias establecidas en las normas jurídicas se apliquen, efectivamente, cuando ocurran los hechos a los cuales el derecho las condiciona. Así, «solo si el proceso judicial cumple la función de determinar la verdad de las proposiciones referidas a los hechos probados podrá el derecho tener éxito como mecanismo pensado para dirigir la conducta de sus destinatarios» (FERRER BELTRÁN, 2007: 30) y, en conexión con ello, el principio también funciona como una condición necesaria de la legitimidad pública del sistema de justicia, pues si este fuera «frecuentemente percibido como uno que condena al inocente y que absuelve al culpable no conseguiría ganarse el respeto, ni la obediencia, de los gobernados» (LAUDAN, 2006: 23).

En definitiva, la verdad es un valor con múltiples repercusiones y la mayor o menor aptitud de un proceso penal para alcanzarla es un parámetro válido para analizar la justicia de las soluciones que produce, sea cual fuera la tradición jurídica de la que provenga²⁹. Esto no significa desconocer la existencia de otros valores que también deben ser considerados y cuya mayor o menor incidencia se relacionará con las características de cada modelo procesal³⁰. En todo caso, lo central es reflexionar sobre cuáles son las condiciones de celebración de los acuerdos que podrían reflejar una ponderación aceptable de todos los valores en juego.

la coincidencia entre la culpabilidad penal declarada en el juicio y los eventos ocurridos en el mundo real. Esto es, como ya dijimos, que condenemos a los culpables y absolvamos a los inocentes. Agradezco al Profesor Jordi FERRER BELTRÁN por esta observación.

²⁹ Esta posición implica, desde una perspectiva moral, considerar a la verdad como un valor universal. Sobre la dicotomía universalismo-relativismo moral puede consultarse TIMMONS, 2002: 38-41. Desde una perspectiva más cercana al relativismo, se considera que existen distintos tipos o nociones de verdad y justicia, dependiendo de la cultura procesal subyacente. Por ejemplo, véase GRANDE, 2008: 155-156.

³⁰ Véase, por todos, DAMAŠKA, 1997: 115-127.

3. SOBRE LAS CONDICIONES DE LEGITIMIDAD MORAL DE LOS ACUERDOS DE CULPABILIDAD: LA MEDIDA DE LA PENA Y LA SUFICIENCIA DE LA PRUEBA

3.1. ¿Propuestas o amenazas? Sobre la legitimidad de la propuesta de un acuerdo de culpabilidad por parte de la fiscalía

En este apartado sostendremos que, para que la propuesta de un acuerdo de culpabilidad por parte de la fiscalía —o del órgano encargado de ello, según la legislación aplicable— no viole los derechos del acusado —o, según veremos, para que resulte moralmente aceptable—, deben verificarse, necesariamente³¹, dos condiciones: (i) que el acusado no sea amenazado con una pena exagerada si decidiera optar por la realización del debate ordinario y (ii) que la hipótesis acusatoria haya sido corroborada de conformidad con un estándar probatorio bien determinado, equivalente a aquel que resultaría exigible para llevar el caso a juicio. Estas exigencias se sustentan en razones independientes a una pretensión dirigida a la averiguación de la verdad, pero, como veremos más adelante, tienen implicancias concretas en el mejoramiento de las condiciones epistémicas en las cuales tomamos la decisión final. Veamos.

Una de las maneras más usuales de exhibir el carácter *inmoral* o *coercitivo* de los acuerdos es a través de su comparación con una escena compatible con lo que, normalmente, consideraríamos una amenaza. Tomemos, por ejemplo, el caso de una persona que es asaltada, a quien el ladrón exige que entregue sus pertenencias pues, si no lo hace, será asesinada o agredida de alguna manera. La posición del acusado frente a la propuesta de un acuerdo de culpabilidad sería asimilable a la del destinatario del asalto, porque también sería perjudicado si no procede como le piden. Es decir, si no renuncia a la realización del juicio, se enfrenta a la posibilidad de recibir una pena mayor. En ambos casos, el acusado y la víctima serían colocados frente a una opción injusta o meramente ficta, pues, en rigor, solo hay una respuesta racionalmente aceptable ante la alternativa propuesta, ya que, de otro modo, se verían colocados frente a una situación desventajosa —*i.e.*, una pena más alta o algún riesgo para su vida o su salud, según el caso—. Como explica KIPNIS:

Tanto el asaltante como el fiscal piden a las personas que realicen una elección difícil entre un perjuicio cierto y menor y un perjuicio incierto mayor. En la situación del asaltante, debo elegir entre la segura pérdida de dinero y la probabilidad difícilmente calculable de que el asaltante esté dispuesto y sea capaz de matarme si me resisto. Como imputado, me veo forzado a elegir entre un castigo cierto y menor y un castigo sustancialmente mayor con una probabilidad difícilmente calculable. Mientras la medida del perjuicio cierto disminuye y la magnitud de la

³¹ Es decir, se trata de condiciones *necesarias*, no *suficientes*. Existen otras razones por las cuales la propuesta de un acuerdo de culpabilidad por parte de la fiscalía podría violar los derechos del acusado, pero no todas ellas serán analizadas en este trabajo. La conclusión que se pretende alcanzar es que, ausentes las dos exigencias que sí se analizarán en este apartado, entonces la propuesta será ilegítima.

probabilidad del perjuicio mayor aumenta, deviene más y más razonable elegir el primero (1976: 93-106, mi traducción desde el original).

La analogía, así formulada, ha sido criticada con buenas razones.

En primer lugar, existen diferentes clases de situaciones en las que nos enfrentamos a propuestas «que no podemos rechazar» o casos en los que, frente a una alternativa presentada por un tercero, existe una única opción racionalmente aceptable o debemos elegir «entre dos males» y, pese a ello, no consideramos que esa decisión haya sido el producto de alguna clase de coerción o de una situación de injusticia. LANGER ilustra esta situación con un ejemplo:

[S]i B tiene una enfermedad mortal y el médico A le propone someterlo a un tratamiento que, si B colabora, le salvará la vida pero lo dejará sometido a intensos dolores por el resto de su vida, entonces B debe elegir entre dos males: morir o colaborar con un tratamiento que lo llevará a sufrir intensos dolores por el resto de su vida. Si B optara por realizar el tratamiento, B podría decir que la propuesta de A era una ‘propuesta que no podía rechazar’, o que no tenía otra opción que la de aceptar la propuesta de A. Pero si B aceptara esa propuesta, normalmente no consideraríamos su aceptación como involuntaria. De este modo, si A le realiza el tratamiento a B, B no podría luego demandarlo por sus intensos dolores. En otras palabras, hay muchas situaciones en las que optar entre dos alternativas malas no constituye una decisión involuntaria. La diferencia entre la pena del juicio y la pena por declaración de culpabilidad tampoco convierte, por sí misma, la decisión de los acusados en involuntaria (LANGER, 2006: 232)³².

¿Qué distingue el caso de la víctima del asalto respecto del paciente médico? En ambos casos nos encontramos ante propuestas en las que un sujeto A propone a un sujeto B hacer la acción X y, además, anuncia que hará Y —perjudicial para B— si B no hace X (WERTHEIMER, 1979: 277). En nuestros ejemplos, esas propuestas serían «dame tu dinero y, si no lo haces, morirás», o «sometete a este tratamiento y, si no lo haces, morirás». Sin embargo, aunque asumamos que tanto la víctima del asalto como el paciente se encontraban, en alguna medida, compelidos a aceptar la propuesta del asaltante y del médico, respectivamente, nuestra intuición nos indica que, mientras la situación en la que fue colocada la víctima del asalto es injusta, la situación del paciente no lo es³³.

³² La traducción al español, en cada cita de esta obra, corresponde a Agustín VARELA y E. Matías DÍAZ, sobre «Repensando el *Plea Bargaining*: La Práctica y la Reforma del Juzgamiento por los Fiscales en el Proceso Penal Estadounidense», en LANGER, Máximo, ¿Para qué sirve un modelo procesal? Ensayos sobre los sistemas acusatorio e inquisitivo, Marcial Pons, Madrid, en prensa.

³³ El ejemplo podría criticarse porque el médico no es responsable por el riesgo de muerte al que se encuentra sometido el paciente, mientras que el asaltante sí lo es respecto de su víctima. Sin embargo, existen otros supuestos en los que el proponente no ha colocado a la víctima en la situación de desventaja que fundamenta la coerción desde una perspectiva psicológica, pero igualmente consideramos injusto su proceder. Imagínese que A cae a un precipicio accidentalmente y logra sostenerse, pero no podrá hacerlo por mucho tiempo. Entonces, B ofrece su ayuda a cambio de que A le entregue una suma de dinero. En este caso, B no colocó a A en una situación de desventaja pero sí se aprovechó injustamente de ella. En cualquier caso, al margen de la relevancia que, para la teoría moral sobre la coerción, pueda tener la circunstancia de que haya sido el proponente quien haya colocado al sujeto receptor en la disyuntiva en la cual se encuentra, los ejemplos todavía funcionan para comunicar la diferencia que

A partir del concepto de coerción proporcionado por la ley contractual estadounidense³⁴, WERTHEIMER explica que, para afirmar que B fue objeto de coerción por parte de A para hacer X, debemos poder afirmar (i) que B fue *psicológicamente compelido* a aceptar hacer X y (ii) que debe resultar *incorrecto* para A compeler psicológicamente a B para que haga X. Así, con respecto a la teoría moral subyacente a esas condiciones jurídicamente establecidas, WERTHEIMER sugiere que la coerción puede ser detectada a través de la realización de dos *tests* independientes: uno psicológico y otro moral (1979: 271).

Apliquemos los *tests* al caso de la víctima y del paciente. En primer lugar, tanto la víctima del asalto como el paciente se encontraron *psicológicamente compelidos* a aceptar las respectivas propuestas. Como vimos, ambos enfrentaron una oferta en la que solo una de las opciones disponibles era razonablemente aceptable, con la expectativa de evitar un mal futuro. En este específico sentido, ambas decisiones fueron el producto de una situación de coerción. Desde esta perspectiva psicológica, afirmar que la víctima y el paciente «involuntariamente aceptaron» la propuesta constituye un enunciado meramente *descriptivo* respecto del contexto de su decisión (WERTHEIMER, 1979: 209-210). Sin embargo, nuestra intuición moral³⁵ nos sugiere que esas situaciones, aunque similares en el plano descriptivo, difieren sustancialmente en el plano normativo.

Según WERTHEIMER, la diferencia entre esas conductas puede advertirse si se indaga la legitimidad de lo que denomina el *plan unilateral declarado por el proponente*. Algunos planes de conducta conllevan la violación de un deber moral y, aunque la inmoralidad de un plan puede no ser una condición *necesaria* de la inmoralidad de una proposición bicondicional, sí es una condición *suficiente* (1979: 277)³⁶. En el caso

luego estableceremos entre la situación del fiscal y el asaltante, esto es, la circunstancia de que la propuesta no constituya una violación de derechos, por subyacer a ella un plan de conducta inmoral o por reducir las opciones originalmente disponibles por parte del sujeto receptor. Agradezco al profesor Jordi FERRER BELTRÁN por esta observación.

³⁴ En particular, *The Restatement of Contracts*, que define la coerción como «(a) cualquier acto injusto de una persona que compele una manifestación de aparente consentimiento de otra persona a una transacción determinada, sin su voluntad, o (b) cualquier amenaza injusta de una persona [...] que [induce] a otro a participar de una transacción bajo la influencia de dicho temor, desde que le impide ejercer su voluntad y su juicio» (WERTHEIMER, 1979: 271, mi traducción desde el original).

³⁵ Una de las formas de evaluar críticamente una teoría moral es su *atractivo intuitivo*, es decir, su conformidad con nuestras creencias acerca de la moralidad de nuestras acciones (TIMMONS, 2002: 14).

³⁶ FARREL y SPECTOR entienden que la verificación de una situación de coerción depende del análisis de determinados factores, tales como la eliminación coactiva de una opción de la cual originalmente disponía el sujeto. En la amenaza, «una opción disponible para el agente antes de la propuesta resulta eliminada una vez hecha ésta; en la propuesta, en cambio, se mantiene la opción original y se agrega por lo menos una nueva opción» (1992: 20). Sin embargo, la moralidad de la propuesta no es determinante, a su juicio, para afirmar la existencia de un caso de amenaza y podríamos encontrarnos ante una escena que no involucre una amenaza sino una oferta, pero que esta resulte inmoral (1992: 23-30). La distinción no afecta nuestro argumento. La fiscalía no debe amenazar al acusado pero tampoco puede realizar una oferta inmoral, en violación a los derechos que aquel tiene en el contexto de esa decisión.

del médico, su plan unilateral —realizar un tratamiento doloroso para tratar la enfermedad del paciente— no constituye la violación de ningún deber moral a su cargo, pero sí lo es el plan declarado por el ladrón —apropiarse del dinero de la víctima—. Esto quiere decir que, mientras que el médico tenía el derecho de someter al paciente a esa alternativa, el ladrón violó un deber moral al someter a la víctima de esa manera.

¿Qué sucede, entonces, con la propuesta de un acuerdo de culpabilidad formulada por la fiscalía? Según el enfoque que elegimos, es necesario analizar si la fiscalía, al proponer un acuerdo de culpabilidad, declara un plan unilateral legítimo o, por el contrario, viola los derechos del acusado³⁷.

Para ello, primero debemos dilucidar cuál es el plan unilateral declarado por la fiscalía. Al proponer un acuerdo, la fiscalía ofrece, en principio, la imposición de una pena reducida a cambio de la declaración de culpabilidad del acusado. Como contrapartida —es decir, para el caso en el que el imputado no acepte su ofrecimiento— la intención declarada por la fiscalía es la de promover la realización de un juicio ordinario y solicitar allí la imposición de una pena mayor. En consecuencia, el ofrecimiento de la fiscalía es una propuesta estructurada, a grandes rasgos, de la siguiente manera: «La fiscalía propone la pena X al imputado a cambio de que admita su culpabilidad y, si el imputado no acepta ese acuerdo, declara que solicitará la realización del juicio y la imposición de una pena mayor».

La moralidad de ese plan unilateral depende, entonces, de que la fiscalía no viole ningún derecho del acusado al solicitar la realización de un juicio y la imposición, en ese contexto, de una pena mayor. Para conocer si ello es así, es necesario determinar cuáles son las condiciones en las que, legítimamente, la fiscalía podría promover esas demandas.

3.2. La medida de la pena: recompensas vs. penalidades

Revisemos, en primer lugar, la cuestión vinculada con la pena en expectativa. Como vimos, la circunstancia de que el acusado espere una sanción mayor en el eventual juicio significa, para algunos, que su decisión de renunciar a ese derecho nunca será libre, ya que la intención de evitar un castigo mayor funcionará como un

³⁷ LANGER se refiere a este asunto a través de la noción de *base moral* —*moral baseline*— de los acuerdos, es decir, a qué tiene derecho moralmente el acusado frente a la propuesta de la fiscalía. Según explica, la base moral de los acuerdos de culpabilidad está compuesta por tres elementos. Primero, el fiscal «no debería efectuar propuestas de declaración de culpabilidad en los casos en los que un jurado razonable no podría encontrar al acusado culpable más allá de una duda razonable», segundo, «los acusados [también] tienen derecho a esperar que, si son condenados en un juicio, el sistema de justicia penal aplicará una pena justa adecuada a las características del delito y del condenado» y «[f]inalmente, como oficiales de justicia que son, los fiscales deben imputar a un individuo solo por delitos que describan adecuadamente la conducta realizada por él, [...] que no sea socialmente inocua» (2006: 236-240).

incentivo para prestar conformidad con el acuerdo, incluso si ello implica emitir una confesión total o parcialmente falsa.

Es necesario distinguir dos escenarios posibles. Imaginemos el caso en el que A es acusado por el delito de robo, que tiene una pena de un mes hasta seis años de prisión. Imaginemos también que, en razón de las circunstancias del hecho —y los restantes datos que el sistema considere relevantes para evaluar este punto—, la pena estimable para A sería de cuatro años de prisión. En este supuesto, la concreción de un acuerdo por dos años de prisión contendría una reducción de dos años respecto de aquella sanción que, por razones de justicia, sería esperable para A. En otras palabras, A resultaría beneficiado o recompensado por su decisión de evitar la realización del juicio. Sin embargo, si A fuese además amenazado con que, en caso de no aceptar el acuerdo, será solicitada en el juicio una pena de seis años de prisión, entonces la propuesta de la fiscalía también contendría una especie de penalización por la decisión de rechazar el acuerdo, esto es, un incremento de dos años en la sanción normalmente esperable en ese caso.

Recapitulando, al celebrar un acuerdo de culpabilidad, un acusado puede ser beneficiado con la reducción de la pena en expectativa, pero también puede ser amenazado con ser castigado por hacer uso de su derecho a ir a un juicio ordinario. En el primer caso, nos hallamos ante la concesión de una *recompensa por la renuncia al juicio* —*waiver reward*— y, en el segundo, ante la imposición de una *penalización por la elección del juicio* —*trial penalty*— (LIPPKE, 2011: 10-23)³⁸.

Ahora bien, a diferencia de la concesión de una *recompensa*, la imposición de una *penalización* es una conducta moralmente reprochable porque «los acusados tienen derecho a esperar que, si son condenados en un juicio, el sistema penal de justicia aplicará una pena justa adecuada a las características del delito y del condenado» (LANGER, 2006: 239). Entonces, si la propuesta de la fiscalía incluye esa clase de penalización, su plan unilateral declarado no es legítimo, mientras que si contiene una recompensa, sí lo es³⁹.

El respeto de esta condición también permite descartar que el ofrecimiento de la fiscalía constituya una *amenaza*, entendida como la eliminación coactiva de una

³⁸ La posibilidad de practicar esta distinción depende, claro, de la existencia de un procedimiento confiable para la determinación de la sanción esperable tras la celebración ordinaria del debate. Podría tratarse, por ejemplo, de la realización de una audiencia ante el tribunal en la que se establezca la sanción justa en expectativa, de acuerdo con las circunstancias del caso, junto con la obligación de respetarla como tope máximo en el eventual debate público (LIPPKE, 2011: 16-23).

³⁹ No pretenden aquí descartarse todas las objeciones que podrían dirigirse contra la práctica de disminuir la sanción esperable en razón de la celebración de un acuerdo. Como vimos, los acuerdos de culpabilidad también son criticados, desde un punto de vista moral, en razón de la reducción de la pena justa dirigida a los autores culpables (SANCINETTI, 2010). El único punto que pretendemos alcanzar es que, mientras la fiscalía ofrezca una recompensa por la omisión del juicio —*waiver reward*— y no una penalización por la elección del juicio —*trial penalty*—, no podría afirmarse que la «amenaza» de la imposición de esa pena es injusta, puesto que se trata de aquella sanción que sería normalmente esperable a la luz de los parámetros que la legislación establece para su determinación.

opción originalmente disponible para el receptor de la propuesta (FARREL y SPECTOR, 1992: 20). Volvamos al caso del asaltante. Es claro que el «ofrecimiento» del asaltante constituye una amenaza porque la víctima pierde injustamente una opción disponible antes de la realización de la propuesta: conservar sus pertenencias y no ser agredida en su vida o su salud. ¿Qué sucede en el caso del acusado? Si la fiscalía propone un acuerdo de culpabilidad, pero amenaza con solicitar la imposición de una pena excesiva en un eventual juicio, podríamos afirmar que, en alguna medida, elimina una opción originalmente disponible para el acusado, es decir, la de participar del juicio y obtener allí una pena justa. En cambio, si la fiscalía propone una recompensa por la evitación del juicio —*waiver reward*—, el acusado gana una nueva opción —la de realizar el acuerdo y obtener una sanción disminuida—, pero no pierde las alternativas originalmente disponibles —acudir al juicio y obtener una sanción justa—.

3.3. La suficiencia de la prueba: el estándar necesario para llevar el caso a un juicio ordinario

Nos queda analizar bajo qué condiciones la fiscalía tiene el derecho de solicitar la realización de un juicio respecto del acusado. Una de las condiciones *necesarias*, aunque no suficientes, de dicha potestad, es que la fiscalía cuente con una determinada sospecha dirigida al imputado, fundada en suficiente evidencia obtenida durante la investigación preliminar que determine la probable existencia del hecho y su participación en él.

Esto se explica, en buena medida, porque la intervención del acusado en el juicio público, además de ser su derecho, es una de las cargas que le resultan exigibles en razón de su sometimiento a un proceso penal. El Estado tiene la potestad de requerir su presencia allí y, de hecho, en la mayor parte de las legislaciones procesales penales del mundo, tiene la facultad de compelerlo a presentarse a través de la fuerza pública. Esta carga supone, claro, la restricción de otros derechos y libertades⁴⁰, de manera que debe ser exigida *razonablemente*, es decir, cuando el Estado cuente con elementos suficientes para justificar esa decisión⁴¹.

⁴⁰ Para PASTOR, «la persecución penal implica, desde el comienzo, el sometimiento del imputado a condiciones de semi-penalización que se manifiestan en ciertos padecimientos que encuadran en el llamado carácter idéntico a la pena que se atribuye ya al proceso penal: angustia, gastos, pérdida de tiempo y de trabajo, humillación, descrédito, etc.» (2009: 393).

⁴¹ En su análisis del caso estadounidense, LANGER agrega que, dado que el fiscal tiene el deber de buscar justicia, «no debería efectuar propuestas de declaración de culpabilidad en los casos en los que un jurado razonable no podría encontrar al acusado culpable más allá de una duda razonable», entonces, «si un fiscal no tiene, como mínimo, evidencia suficiente para sortear o pasar el umbral de una decisión absolutoria directa por parte del juez luego de que la fiscalía presentó su caso en juicio y antes de que la defensa presente su caso, el fiscal tiene el deber moral de no incitar al acusado a declararse culpable mediante una propuesta de declaración de culpabilidad» (2006: 236-237).

En definitiva, si el plan unilateral declarado por la fiscalía es la realización de un debate oral y público, la moralidad de esa pretensión dependerá de la observancia de la totalidad de las condiciones necesarias para llevarla adelante legítimamente. Entre ellas, se encuentra la existencia de una cantidad de prueba que sustente la hipótesis acusatoria con un grado de probabilidad determinado, según el cual resulte normalmente tolerable la carga exigible a todo acusado de someterse a un enjuiciamiento penal. De acuerdo con ello, la propuesta de un acuerdo de culpabilidad debe apoyarse en la satisfacción del mismo estándar de prueba aplicable a la decisión de remitir un caso penal a un juicio ordinario.

4. REPERCUSIONES EN LA DECISIÓN SOBRE LA *QUAESTIO FACTI* EN LOS ACUERDOS DE CULPABILIDAD

Como vimos antes, las condiciones de legitimidad de los acuerdos —relacionadas con la suficiencia de la prueba sobre la culpabilidad y la medida de la pena— tienen una justificación moral lógicamente independiente de una pretensión específicamente orientada a la averiguación de la verdad. Sin embargo, si las observamos debidamente, contribuirán al mejoramiento de la precisión fáctica de las decisiones resultantes de los acuerdos.

En otras palabras, las exigencias antes descriptas, que los acuerdos deben necesariamente observar para no violar los derechos del acusado, repercuten en las condiciones en las cuales tomamos la decisión sobre la *quaestio facti* en estos contextos y permiten afirmar que la condena resultante de los acuerdos podría gozar de una justificación genuina desde un punto de vista epistémico, esto es, sobre la base de prueba y de acuerdo con una noción de verdad como correspondencia con la realidad.

Para ello, contaremos con *(i)* un determinado nivel de comprobación de la culpabilidad del acusado, a través de las evidencias recolectadas durante la investigación preliminar, *(ii)* la inexistencia de hipótesis rivales a la acusatoria y la expresa aceptación de ésta por parte del acusado y *(iii)* una manifestación de conformidad que, en principio, es epistémicamente confiable.

4.1. Un determinado nivel de comprobación de la culpabilidad del acusado

Según vimos, para proponer un acuerdo de culpabilidad, la fiscalía debe haber satisfecho un estándar probatorio asimilable al que resultaría exigible para llevar el caso a juicio.

La intermediación de un estándar de prueba implica que el juzgador sobre los hechos deberá determinar si las evidencias proporcionan a la tesis de la acusación

un grado de corroboración suficiente para aceptarla. Efectivamente, un estándar de prueba «especifica el umbral mínimo que ha de ser satisfecho a los efectos de aseverar que una hipótesis ha sido probada» (LAUDAN, 2006: 104), de manera que, al adoptar una decisión sobre los hechos probados⁴², el juzgador deberá dilucidar si la hipótesis puede considerarse comprobada con el grado de confirmación de que se disponga (FERRER BELTRÁN, 2007: 47; HAACK, 2013: 75).

Es generalmente aceptado que en el proceso penal existen diferentes clases de estándares, según cuán exigentes sean con respecto al grado de acreditación que debe haber alcanzado la hipótesis para poder ser aceptada⁴³. La existencia de un estándar de prueba supone también una determinada distribución del riesgo de error en la decisión asociada con él (FERRER BELTRÁN, 2018: 406 y 425). En particular, una de las razones que puede justificar una determinada distribución del riesgo de error a través de un estándar más o menos exigente viene dada la diferente gravedad de las equivocaciones posibles (*idem*: 409-410). Por ejemplo, en el caso de la decisión definitiva en un caso penal, el costo de las condenas erróneas y de las absoluciones erróneas⁴⁴.

¿Cuán costosos son los falsos positivos y los falsos negativos en la resolución que discutimos? Aquí se impone una aclaración. La decisión sobre la legitimidad de la propuesta de la fiscalía no es, estrictamente, una decisión sobre una cuestión fáctica o que, al menos de manera directa, esté sometida a un estándar de prueba determinado. Son las exigencias morales a las que se encuentra condicionado el ofrecimiento de la fiscalía las que, según vimos, obligarían al juzgador sobre los hechos a determinar si la hipótesis acusatoria goza del grado de corroboración que hubiera sido necesario para la adopción de un pronunciamiento distinto, esto es, el de llevar el caso a juicio. Entonces, lo que el tribunal debe indagar es si fue satisfecho el estándar probatorio asociado con esa otra resolución, pues existen razones morales para conectarla con la determinación de proponer un acuerdo de culpabilidad. Así, la definición del estándar de prueba y la consecuente distribución del riesgo de error serán aquellas correspondientes a la decisión de llevar un caso a juicio.

Entre los costos que pueden ser evaluados para determinar la gravedad del error en la decisión de llevar un caso a juicio o, como contrapartida, en la decisión de no celebrarlo, se encuentran las restricciones a los derechos y libertades del acusado que representa la realización de ese acto, por un lado, y los costos asociados a la decisión

⁴² FERRER BELTRÁN, 2007: 41 y ss.

⁴³ Efectivamente, «[e]n cada punto del procedimiento [penal] en que a un juez o a un jurado le corresponde hacer una determinación de derecho o una relativa a los hechos del caso, hay un EdP [—estándar de prueba—] asociado a esa decisión» y, cuando se emplean estándares de prueba, «subyace implícitamente la expresión de los costos relativos a los éxitos y los errores que pueden derivarse de la decisión respectiva» (LAUDAN, 2006: 133-134).

⁴⁴ Por ejemplo, una determinada posición sobre los costos relativos a las decisiones definitivas de un proceso —es decir, a la dicotomía condena-absolución— puede verse en LAUDAN, 2006: 111 y ss.

de clausurar anticipadamente una investigación, por el otro⁴⁵. En ambos casos, los errores son de menor gravedad en comparación con aquellos asociados al pronunciamiento final. Someter a una persona a juicio injustamente no es igualmente grave que condenarla injustamente y decidir no promover un debate tampoco es igualmente grave que absolver injustamente a alguien, máxime cuando ese pronunciamiento es generalmente provisional y no impide reeditar la pretensión de realizar el juicio en un futuro.

En ese contexto, usualmente se considera que la elevación de un caso a juicio se encuentra condicionada a la recolección de evidencia que haga presumir como *probable* que podría condenarse al acusado de conformidad con el estándar aplicable a la decisión final⁴⁶. Para solicitar la realización de un juicio «no se le requiere al acusador que demuestre con certeza la culpabilidad del acusado», sino que «[b]astará aquí que demuestre la probabilidad de que la persona sometida a proceso haya cometido el hecho», noción que, «en punto al requisito cognoscitivo, [tiene] cierta similitud con el estándar probatorio de la *preponderancia de la prueba*» (LIMARDO, 2018: 99)⁴⁷, según el cual cabe elegir el «enunciado que parezca relativamente “más probable”, sobre la base de los medios de prueba disponibles» (TARUFFO, 2008: 137-138), o bien con la idea de la *probabilidad prevaleciente* como criterio de selección entre hipótesis contradictorias o incompatibles, según la cual resulta aceptable seleccionar aquella que «reciba el apoyo *relativamente* mayor sobre la base de los elementos de prueba conjuntamente disponibles» (TARUFFO, 2011: 299).

La elección de esas categorías u otras no es aquí decisiva⁴⁸, aunque sí es deseable transmitir al juzgador sobre los hechos los parámetros propios de una decisión

⁴⁵ La idea de que la decisión de llevar un caso a juicio no debe ser adoptada arbitrariamente, en razón de los padecimientos que ello implica para el sujeto acusado, puede leerse en MAIER, 2004: 362, BERTELOTTI, 2009: 103-122, entre muchos otros. También se ha sostenido que la necesidad de que la decisión de llevar un caso a juicio se encuentre debidamente fundada responde a la necesidad de evitar, en la mayor medida posible, absoluciones incorrectas con fuerza de cosa juzgada (DEU, 2007: 48).

⁴⁶ Por ejemplo, para CAFFERATA NORES la elevación a juicio «requerirá probabilidad» (2008: 11-12) y para MAIER, la apertura a juicio debe producirse cuando «aparezca como *probable* una condena, esto es, cuando exista la probabilidad acerca de la existencia de los elementos de la imputación deducida en la acusación que, eventualmente, permitan condenar al tribunal de juicio» (2011: 358-359).

⁴⁷ El autor propone, sin embargo, que para evitar la realización de un *pre-juicio* para decidir sobre la necesidad del juicio ordinario, el control del mérito de la acusación durante la etapa intermedia debiera consistir en un «control negativo de arbitrariedad, en donde la tarea del juez que lo efectúa se concentra en corroborar que no exista arbitrariedad, por lo que solo en los casos en que resulte evidente que la base probatoria de la acusación es insuficiente para alcanzar el estándar cognoscitivo requerido y, en esa medida, arbitraria, deberá proceder a cumplir con su función de filtro e impedir que esa acusación prosiga» (LIMARDO, 2018: 122-123).

⁴⁸ No es claro que lo ideal sea tener definiciones precisas de los estándares, aunque ello sea posible (v. HAACK, 2013: 70 y siguientes) y la posibilidad misma de formular un estándar preciso, que funcione como umbral en la decisión sobre los hechos probados, ha sido controvertida. A favor de esta posibilidad, véase FERRER BELTRÁN, 2018: 405-419 y 2007: 139 y siguientes y, en contra de esa posibilidad, véase GONZÁLEZ LAGIER, inédito.

intermedia, necesarios para determinar el nivel de evidencia exigible, en términos relativos⁴⁹. En este caso, el grado de corroboración de la hipótesis acusatoria debe ser menor a aquel exigible para la adopción de la condena, pero mayor a aquel aplicable a las decisiones iniciales del proceso —por ejemplo, la de comenzar una investigación o poner en conocimiento del acusado los hechos que se le imputan—.

Podría objetarse que las evidencias reunidas durante la investigación preliminar son de baja calidad, pues son recolectadas unilateralmente por las policías o las fiscalías y no son sometidas al control propio del debate⁵⁰, y que, en esas condiciones, la hipótesis acusatoria, aunque carente de rivales, puede no ser otra cosa que una «mala explicación». Efectivamente, si entendemos la decisión sobre los hechos probados como una inferencia a la mejor explicación, consistente en un proceso de maximización de la coherencia, la elaboración de una base de coherencia pobre puede provocar que la explicación finalmente escogida sea la mejor entre aquellas disponibles, pero una mala explicación al fin de cuentas (AMAYA, 2013: 19-20).

Ahora bien, ese riesgo podría ser compensado si la observancia de ese estándar es sometida a un control intersubjetivo fuerte y si, en ese control, se contempla la noción de completitud del cuadro probatorio.

Respecto de lo primero, es claro que el acusado y su defensa deben ser puestos en conocimiento de la totalidad de la evidencia reunida por la fiscalía en su contra. De esa manera, podrán controlar esa prueba y, en su caso, proponer las hipótesis alternativas que consideren necesarias. Efectivamente, «[l]a omisión de revelar la prueba favorable a la defensa interfiere con un mecanismo central para controlar que los fiscales no efectúen propuestas de acuerdo en casos débiles» (LANGER, 2006: 272)⁵¹. Volveremos a referirnos a este punto en el siguiente apartado.

⁴⁹ Como explica GONZÁLEZ LAGIER, «[t]odo el mundo entiende que el criterio de confirmación ‘más allá de toda duda razonable’ es más exigente que el de la “prueba preponderante” o el de “prueba clara y convincente” [...]. Al menos, los estándares deberían transmitir la información de si la autoridad jurídica quiere establecer un nivel de exigencia mayor o menor, aunque no logren precisar cuál y, por tanto, si se ha alcanzado suficiente prueba acabe dependiendo de la estimación del juez y su buen criterio» (inédito).

⁵⁰ Para LANGER, el elemento común entre las diferentes adaptaciones de los procedimientos de omisión del juicio implementados a lo largo del mundo es la *administratización* del proceso. Según explica, «las condenas basadas en el consentimiento del acusado y en la evidencia recolectada en procedimientos administrativos le otorga mayor poder a la policía, a otras agencias administrativas y a la fiscalía, quienes los desarrollan en la medida en que no son cuestionados ni responsabilizados a través de un juicio y dado que sus decisiones de detener e imputar a un individuo por un determinado delito pueden no ser accesibles o transparentes para terceros, los ciudadanos o los medios de comunicación» (2019: 25).

⁵¹ En el ámbito estadounidense, ha sido discutida la vigencia y, también, la suficiencia de la regla de *disclosure* establecida en el precedente *Brady v. Maryland*, 373 U.S. 83, 86 (1963), en el ámbito del *plea bargaining*. De acuerdo con esa regla, se considera una violación al debido proceso, fundada en las enmiendas quinta y catorceava de la Constitución de los Estados Unidos, que la fiscalía no revele prueba de descargo que es decisiva para la decisión sobre la culpabilidad o la pena. Una argumentación en favor de la vigencia de dicha regla en el ámbito del *plea bargaining*, sobre la base de fundamentos diversos

En segundo lugar, al margen del conocimiento e intervención del propio acusado y su defensa, es central el control por parte del tribunal sobre la observancia del estándar necesario para proponer el acuerdo y, en el ejercicio de ese control, podría resultar útil la idea del *peso* de la prueba o de la *completitud* del cuadro probatorio⁵². Esta noción «mide al grado de completitud (*i.e.*, la riqueza) del conjunto de elementos de juicio con el que se adopta la decisión», es decir que, «mientras que el valor probatorio de una hipótesis da cuenta de la probabilidad de que esa hipótesis sea verdadera en términos comparativos con sus hipótesis rivales, el peso probatorio da cuenta de la probabilidad de acierto de la decisión, en función de la riqueza de los elementos de juicio tomados en consideración para decidir», de modo tal que, «incorporando el requisito de la completitud del peso probatorio ya no se produciría la consecuencia de tener que declarar probada una hipótesis con muy escaso apoyo probatorio, únicamente por ser relativamente mejor que su contraria» (FERRER BELTRÁN: 418, nota 31).

En consecuencia, si el tribunal que controla la satisfacción del estándar necesario para llevar el caso a juicio —al analizar la legitimidad de la propuesta del acuerdo— incluyera la dimensión del peso o de la completitud del cuadro probatorio, podría evitarse, en buena medida, el problema asociado a la posibilidad de que la hipótesis seleccionada sea, en rigor, una «mala explicación»⁵³.

4.2. Un particular contexto de justificación: la inexistencia de hipótesis rivales propuestas por la defensa

Al margen del estándar probatorio necesario para proponer el acuerdo —que, como ya vimos, es equivalente a aquel exigible para llevar un caso a juicio—, estos procedimientos cuentan también con una declaración a través de la cual el acusado no controvierte que los hechos se desarrollaron de conformidad con el relato contenido en la hipótesis de la fiscalía⁵⁴. Entonces, en el contexto epistémico final, en el

—desde la teoría *unconstitutional conditions*, hasta las reglas de los contratos, el derecho de propiedad y el debido proceso legal— puede leerse en BLANK, 2000: 2060-2085. Una argumentación en torno a la posible *insuficiencia* de la revelación según “Brady” para el caso de los acuerdos de culpabilidad, puede verse en DOUGLASS, 2001.

⁵² HAACK se refiere a este punto como la *completitud* del cuadro probatorio —*comprehensiveness*— (2014: 60, 218, entre otras).

⁵³ Podría seleccionarse, para el caso, alguno de los estándares intermedios propuestos por FERRER BELTRÁN en su catálogo, que incorporan la dimensión del peso probatorio. Por ejemplo, aquel según el cual «[u]na hipótesis sobre los hechos se considerará probada cuando se den, conjuntamente, las siguientes dos condiciones: a) Que la hipótesis ofrezca una mejor explicación de los hechos cuya ocurrencia se trata de probar que la hipótesis de la parte contraria, a la luz de los elementos de juicio existentes en el expediente judicial; y b) Que el peso probatorio del conjunto de elementos de juicio relevantes incorporados al proceso sea completo (excluidas las pruebas redundantes)» (2018: 418).

⁵⁴ Esto es así, claro está, en los acuerdos en los cuales la conformidad del acusado implica la aceptación de los hechos y de su participación en estos. El análisis que efectuamos en este punto no es aplicable

que adoptamos la decisión conclusiva para el caso, nos encontramos no solo con que el acusado no ha propuesto una hipótesis rival a la acusatoria, sino con que la acepta expresamente.

Existen diferentes teorías que explican el mecanismo de adopción de decisiones sobre los hechos en casos penales como un proceso de selección entre hipótesis alternativas, según su capacidad explicativa respecto de la evidencia disponible.

Por ejemplo, para ALLEN, la decisión sobre la responsabilidad penal del acusado consiste en la selección de una versión *plausible* de la culpabilidad y, además, en la inexistencia de una versión *plausible* de su inocencia (2010: 123). En particular, en un juicio penal, «[l]as partes tratan de exhibir las debilidades de los argumentos de sus oponentes pero, principalmente, intentan mostrar que su propia versión de la realidad es verdadera» y «la funcionalidad de la prueba en los procesos judiciales opera comparativamente respecto de las versiones de los hechos presentadas por las partes» (*Ibid.*: 134-135). Si aceptamos que esta es la manera en la cual decidimos sobre los hechos en un caso penal, es razonable prever que, frente a la aceptación expresa por parte del acusado de la hipótesis formulada por la acusación, esta se presente como la explicación más plausible acerca de lo ocurrido.

Desde esta perspectiva, la justificación es sensible al contexto y el grado de corroboración que exigimos para aceptar una hipótesis como verdadera no es equivalente en todos los escenarios. Por ejemplo, las teorías que describen la decisión sobre los hechos como una abducción o una inferencia a la mejor explicación (ALLEN y PARDO, 2008: 229-231 y AMAYA, 2013: 14-18)⁵⁵ advierten que la elección de la mejor hipótesis es sensible al contexto y a las explicaciones alternativas disponibles (AMAYA, 2010: 38⁵⁶ y ALLEN y PARDO, 2008: 231-233).

a aquellos procedimientos abreviados en los cuales esa conformidad no incluye dicha aceptación. Por ejemplo, en el sistema estadounidense, no sería aplicable ante un *plea de nolo contendere*. A través de este, el imputado declara que no va a discutir el asunto de su culpabilidad o inocencia (KAMISAR, LAFAYE, ISRAEL, KING y KERR, 2008: 1328).

⁵⁵ Para AMAYA, se trata de un proceso de maximización de la *coherencia* y, con ello, de selección de la explicación más *coherente* (2013: 4-5).

⁵⁶ Como bien explica AMAYA, existen diferentes posiciones acerca de qué es aquello que varía con el contexto y, según SCHAFFER, aquello que cambia con el contexto son, justamente, las alternativas epistémicas —*contrast set*— que la hipótesis debe superar (2005: 115-129).

En un sentido similar, GRIMALTOS explica la posición de LEWIS sobre el asunto, según la cual la consideración de posibilidades que impliquen la negación de nuestras creencias «hace que ya no pueda atribuirse el conocimiento que en el otro contexto, en el que no se las ha considerado, no tenía inconveniente en atribuirse» y, si bien la primera no acuerda con el segundo en que la existencia de conocimiento sea sensible al contexto, también asume que esas consideraciones prácticas pueden tener un efecto indirecto en aquel. En particular, que «la consideración de otras posibilidades, menos remotas, más probables y que en un primer momento he ignorado pueden producir el efecto del que habla Lewis», pero solo «si el hecho de considerar tal posibilidad hace que mi grado de certeza disminuya» (GRIMALTOS, 2009: 47-49).

Entre los rasgos del contexto que son relevantes para fijar los estándares de justificación, AMAYA ubica los *rasgos dialécticos* de la explicación y advierte que «[l]as prácticas de justificación tienen lugar en un contexto dialéctico que constriñe qué es lo que puede darse por sentado y qué es lo que, por el contrario, requiere ser justificado. El hecho de mencionar una hipótesis que tiene capacidad de derrotar a aquella que ha sido, provisionalmente, aceptada como justificada, obliga a incrementar el nivel de exigencia o, lo que es lo mismo, el umbral de justificación» (2010: 40-41). Como contrapartida, la expresa aceptación por parte del protagonista de los hechos de la descripción que de ellos hizo la fiscalía debilita los rasgos dialécticos del contexto en el que se discute la hipótesis acusatoria, lo cual permite disminuir, razonablemente, el grado de justificación exigible a la hora de tomar la decisión final.

Esto no significa que la falta de contradicción de la hipótesis acusatoria obligue a aceptarla, sino simplemente que, si la decisión sobre los hechos de un caso penal puede ser explicada como el proceso de selección de la hipótesis más plausible o de la explicación más coherente acerca de la evidencia disponible, la inexistencia de una hipótesis rival propuesta por el acusado modifica sensiblemente el contexto en el cual realizamos esa determinación final.

En otras palabras, si nos preocupara la circunstancia de que la condena resultante de esta clase de procesos pudiere gozar de un soporte en pruebas menor a aquel en el que se apoyan las resoluciones que emanan de un juicio ordinario —porque estas han satisfecho, por ejemplo, el estándar *más allá de toda duda razonable*—, lo cierto es que ello podría encontrar una explicación razonable en las especificidades del contexto epistémico que las rodea. Es decir, dados sus rasgos *no dialécticos*, es razonable y aceptable una disminución en el grado de exigencia de justificación de la decisión final.

Lo afirmado en este punto no significa que el caso nunca haya atravesado el análisis de hipótesis rivales. Ya vimos que el acuerdo, al estar sometido a un determinado estándar de probabilidad, puede haber involucrado el rechazo de hipótesis alternativas a la de la acusación en etapas procesales anteriores —esto es, en el análisis correspondiente a la examinación de la legitimidad de la propuesta de la fiscalía—⁵⁷. En cambio, la intención aquí es colocarnos frente al momento de decisión final del caso, que ya cuenta con el acuerdo completo y, así, con la expresa decisión del acusado de no controvertir la hipótesis acusatoria.

4.3. Una manifestación de conformidad, en principio, epistémicamente confiable

Se ha afirmado que una de las razones por las cuales una persona inocente podría decidir declararse culpable es “la conclusión de que es preferible una pena leve y

⁵⁷ Véase *supra* nota 53.

personalmente asumible, que el riesgo —aunque no sea elevado— de una pena más grave y difícil de asumir”, razón por la cual resulta recomendable “combatir las acusaciones exageradas o infundadas, generadoras de miedos exagerados o infundados en el imputado” (LASCURAÍN SÁNCHEZ y GASCÓN, 2018: 15 y 21).

Ahora bien, sostuvimos antes que la legitimidad del *plan unilateral declarado* por la fiscalía, al proponer un acuerdo de culpabilidad, depende de la determinación de las condiciones bajo las cuales aquella puede demandar la realización de un juicio y la imposición, allí, de una pena mayor. Según vimos, la fiscalía puede ofrecer una recompensa por la renuncia al juicio —esto es, una disminución del castigo justo que es esperable allí—, pero no una penalidad por la elección de esa alternativa (véanse puntos 3.1 y 3.2 de este trabajo). Esto contribuye a disminuir el riesgo de que los acusados se vean *psicológicamente* compelidos a prestar conformidad en razón de la expectativa de imposición de una pena exagerada si optan por el juicio. La conciencia de que, en el eventual debate, solamente podrá imponerse la pena justa —según las características del hecho y las condiciones personales del acusado— o una inferior⁵⁸, podría disminuir los incentivos para que los acusados inocentes se declaren culpables, pues removería buena parte del sustento al temor por un eventual castigo desproporcionado.

Algo similar sucedería a raíz de la adecuada revelación de la totalidad de la prueba relevante recolectada por la fiscalía⁵⁹. Cuando el acusador es obligado a revelar la totalidad de la prueba de cargo y de descargo, reunida durante la investigación, los acusados inocentes son menos susceptibles a ser compelidos a confesar en razón del temor infundado a una condena errónea, ya que poseen información precisa acerca de la fortaleza del caso de la fiscalía y sobre el riesgo cierto de participar de un juicio ordinario (PETEGORSKY, 2018: 3613-3614)⁶⁰.

⁵⁸ La mayoría de las legislaciones procesales contiene reglas que prohíben a los jueces imponer sanciones superiores a aquellas solicitadas por los acusadores, de manera que el imputado, si no acepta el *wavier reward*, tendría garantizado que la pena en el eventual juicio no excederá, por decisión del tribunal, aquella determinada como justa en el trámite del acuerdo de culpabilidad.

⁵⁹ La revelación de prueba de cargo suele ser considerada parte del derecho de defensa en los procedimientos de origen europeo-continental. Más precisamente, este derecho «no se difundió desde las jurisdicciones del *common law* hacia las del *civil law*. De hecho, fue adoptado más temprano por las jurisdicciones del *civil law* que por las del *common law*, no necesariamente porque las primeras estuvieran más orientadas al debido proceso que las últimas, sino probablemente porque el escenario institucional del sistema del *civil law* con su formal tratamiento de la etapa previa al juicio hizo más patente la necesidad de establecer y expandir este derecho una vez que le fue dado al acusado un rol más prominente en la resolución de los casos penales». Además, «las jurisdicciones del *civil law* han dado un mayor alcance al derecho que las del *common law*, en tanto las primeras dan a la defensa un temprano acceso a toda la investigación oficial que dirigen los funcionarios públicos. Pero una vez más, este alcance más amplio no es resultado de una mayor orientación hacia el debido proceso en el sistema del *civil law* sino de la manera en que el asunto ha sido pensado dentro de un conjunto de prácticas institucionales preexistentes y una determinada concepción del proceso penal» (LANGER y ROACH, 2013: 8).

⁶⁰ En sus palabras, «*by forcing disclosure of exculpatory evidence, a prosecutor cannot bluff her way to conviction by misrepresenting the strength of the government's case*» (PETEGORSKY, 2013: 3613). Las

La conformidad manifestada en estas condiciones resultaría, *en principio*, epistémicamente confiable⁶¹.

5. CONCLUSIONES

5.1. La observancia de determinadas condiciones mínimas de legitimidad moral en el marco de los acuerdos de culpabilidad, aunque no se funde en una concreta pretensión orientada a la averiguación de la verdad, repercute en las condiciones epistémicas en las cuales tomamos la decisión sobre la *quaestio facti* en ese contexto.

5.2. Los acuerdos de culpabilidad son procedimientos que permiten el dictado de una condena con prescindencia de la realización de un juicio y de la producción y discusión de la prueba en ese marco. Esto es posible, centralmente, en razón de la *decisión* del acusado de renunciar a su derecho a la celebración de ese juicio y a la corroboración de su culpabilidad, en dicho contexto, de acuerdo con el estándar probatorio normalmente aplicable —p. ej., más allá de una duda razonable—. Pese a ello, la circunstancia de que la fiscalía deba alcanzar un estándar probatorio bien determinado para proponer válidamente la celebración de un acuerdo —aquel equivalente al que resulte necesario para enviar un caso a juicio—, la inexistencia de hipótesis rivales a la acusatoria y una confesión que, en principio, es epistémicamente confiable, muestran que la decisión final sobre la *quaestio facti* en estos casos no viene asociada a un concepto de verdad consensual, sino que puede gozar de una genuina justificación epistémica, esto es, sobre la base de prueba y a la luz de una noción de verdad como correspondencia con la realidad.

5.3. Las conclusiones de este trabajo podrían objetarse observando que, al fin y al cabo, frente a la decisión definitiva de condena, que usualmente se halla sometida a un estándar probatorio muy exigente, terminamos conformándonos con un estándar menor, reservado para otro momento procesal, tal como es la apertura a juicio. Si nos preocupa la precisión fáctica de las decisiones resultantes del juicio penal, ¿por

implicancias de esa clase de *bluff* por parte de la fiscalía ha sido objeto de estudios y experimentaciones. Un interesante examen empírico sobre los efectos de esa clase de engaños —*of being bluffed by the dealer*— en la decisión de un acusado de aceptar un acuerdo de culpabilidad o confesar pueden verse en WILFORD, 2014.

⁶¹ Hablamos de una confesión confiable *en principio* porque la eliminación de las *trial penalties* es una condición necesaria, pero no suficiente, para garantizar la veracidad de esa declaración. Existen otros factores de riesgo —p. ej., malas prácticas policiales, tales como interrogatorios extensos, o bien encarcelamientos preventivos prolongados, entre otros—, que podrían incidir en esa decisión y cuyo análisis excede los limitados objetivos de este trabajo. Sin embargo, el objetivo de esta sección no es desarrollar exhaustivamente la totalidad de las condiciones que podrían incentivar confesiones falsas y, a nuestros fines, basta con asumir como cierta la premisa propuesta por LASCURAÍN SÁNCHEZ y GASCÓN, según la cual que el temor por una condena injusta y exagerada podría constituir uno de ellos. La determinación de la totalidad de las condiciones bajo las cuales una persona podría estar dispuesta a confesar falsamente demanda la realización de un análisis empírico que excede los alcances de este trabajo.

qué no exigir que la decisión definitiva tras un acuerdo de culpabilidad se encuentre sometida al mismo estándar aplicable al juicio ordinario —certeza, más allá de una duda razonable, etcétera—?

En primer lugar, porque la condición de legitimidad de una condena dictada tras un acuerdo de culpabilidad no es, al menos principalmente, un determinado grado de comprobación de la culpabilidad del acusado, sino la decisión de éste de renunciar al juicio. Esto no quiere decir que el estándar *más allá de toda duda razonable* no pueda ser alcanzado en estos trámites, pero nuestra argumentación se encontró dirigida a demostrar un punto anterior, esto es, que la evaluación de la validez de esa renuncia no es independiente a la valoración de cierto grado de comprobación de la culpabilidad del acusado, de manera tal que, a pesar de la omisión del juicio, la decisión definitiva, a raíz de ese y otros factores, podría tener suficientes garantías epistémicas y no ser el producto de la confianza ciega en las afirmaciones de las partes. En otras palabras, ello no nos obliga, necesariamente, a concluir que la decisión definitiva tras el acuerdo —*i.e.*, la condena— esté sometida a ese estándar de corroboración menor. Como ya vimos, la decisión que está sometida a ese estándar de corroboración menor es una anterior y diferente —la de llevar el caso a juicio—, con la cual se encuentra conectada, por razones morales, la evaluación de la legitimidad moral de la propuesta de la fiscalía.

6. BIBLIOGRAFÍA

- ALLEN, Ronald J., 2010: «Versión plausible de culpabilidad sin otra alternativa plausible: Regla de decisión en el proceso penal» en PARCERO, Juan A. CRUZ y LAUDAN, Larry (comps.), *Prueba y estándares de prueba en el derecho*, UNAM, Instituto de Investigaciones Filosóficas.
- AMAYA, Amalia, 2010: «Diez tesis acerca de la coherencia en el Derecho», *Revista Discusiones*, número X —*La Coherencia en el Derecho*—: 21-64.
- 2013: “Coherence, evidence and legal proof”, *Legal Theory*, 19, Cambridge University Press: 1-43.
- BERGMAN, Marcelo, FONDEVILA, Gustavo y LANGER, Máximo, 2017: ¿A quién y cómo se juzga en la Ciudad Autónoma de Buenos Aires? Una radiografía de la Justicia Nacional en lo Criminal y Correccional, Centro de Estudios Latinoamericanos sobre Inseguridad y Violencia, Universidad Nacional de Tres de Febrero.
- BERTELOTTI, Mariano, 2009: «El control del mérito sustantivo de la investigación», en ANITUA, Gabriel I. y TEDESCO, Ignacio F. (comps.), *La cultura penal – Homenaje al Profesor Edmundo S. Hendler*, Buenos Aires: Editores del Puerto: 103-122.
- BLANK, Daniel P., 2000: «Plea Bargain Waivers Reconsidered: A Legal Pragmatist’s Guide to Loss, Abandonment and Alienation», 68 *Fordham L. Review*: 2011-2095.
- BOVINO, Alberto, 2005: «Procedimiento abreviado y juicio por jurados», en MAIER, Julio B. J. y BOVINO, Alberto (comps.), *El procedimiento abreviado*, Buenos Aires: Editores del Puerto: 53-95.
- CAFFERATA NORES, José I. y HAIRABEDIÁN, Maximiliano, 2008: *La prueba en el proceso penal*, Buenos Aires: Lexis Nexis.
- CÓRDOBA, Gabriela, 2005: «El juicio abreviado en el Código Procesal Penal de la Nación Buenos Aires», en Maier, Julio B. J. y Bovino, Alberto (comps.), *El procedimiento abreviado*, Buenos Aires: Editores del Puerto: 229-250.

- DAMAŠKA, Mirjan R., 1997: *Evidence Law Adrift*, Yale University Press, New Haven & London. Citado por su traducción al castellano de Joan Picó I Junoy, *El derecho probatorio a la deriva*, Madrid: Marcial Pons, 2015.
- 2003: «Epistemology and legal regulation of proof» en *Law, Probability and Risk*, Oxford University Press.
- DEI VECCHI, Diego, 2016: «La prueba judicial como conocimiento: una caracterización poco persuasiva», en FERRER BELTRÁN, Jordi y VÁZQUEZ, Carmen (coeds.), *Debatendo con Taruffo*, Madrid: Marcial Pons.
- DÍAS, Leandro, 2015: «Los acuerdos en Derecho penal en Karlsruhe y Estrasburgo: análisis de las recientes sentencias del Tribunal Constitucional Federal Alemán y del Tribunal Europeo de Derechos Humanos Buenos Aires», en *Pensar en Derecho*, nro. 6.
- DÍAZ CANTÓN, Fernando, 2005: «Juicio abreviado vs. estado de derecho Buenos Aires», en MAIER, Julio B. J. y BOVINO, Alberto (comps.), *El procedimiento abreviado*, Buenos Aires, Editores del Puerto, pp. 251-276.
- DÍAZ, E. Matías, 2017: *Sobre la verdad y la protección de la inocencia en el proceso penal*, Buenos Aires: Ad Hoc.
- DOUGLASS, John G., 2001: «Fatal Attraction? The Uneasy Courtship of Brady and Plea Bargaining», *50 Emory L.J.*: 437-5118.
- DUFF, A., FARMER, L., MARSHALL, S. y TADROS, V., 2007: *The Trial on Trial: Volume 3. Towards a Normative Theory of Criminal Law*, Oxford - Portland, Hart Publishing.
- FARRELL, Martín y Spector, Horacio, 1992: «Dos tesis sobre amenazas y ofertas», en *Cornblit, Oscar (comp.), Libertad y Mercado*, Buenos Aires: Macchi.
- FERRAJOLI, Luigi, 1989: *Diritto e ragione. Teoria del garantismo penale*, Roma, Ed. Laterza & Figli. Citado por la traducción al castellano de ANDRÉS IBÁÑEZ y otros, *Derecho y Razón. Teoría del garantismo penal*, novena edición, Madrid: Ed. Trotta, 2009.
- FERRER BELTRÁN, Jordi, 2005: *Prueba y verdad en el derecho*, Madrid: Marcial Pons.
- 2007: *La valoración racional de la prueba*, Marcial Pons, Madrid
- 2010: «El contexto de la decisión sobre los hechos probados en el derecho», en PARCERO, Juan A. Cruz y LAUDAN, Larry (comps.), *Prueba y estándares de prueba en el derecho*, UNAM, Instituto de Investigaciones Filosóficas.
- 2018: «Prolegómenos para una teoría sobre los estándares de prueba. El test case de la responsabilidad del Estado por prisión preventiva errónea», en PAPAYANNIS, Diego M. y PEREIRA FREDES, Esteban (eds.), *Filosofía del derecho privado*, Madrid, Marcial Pons.
- FOUCAULT, Michel, 2014: *Obrar mal, decir la verdad*, Buenos Aires: Siglo XXI.
- GASCÓN ABELLÁN, Marina, 2010: *Los hechos en el derecho*, Madrid: Marcial Pons.
- GONZÁLEZ LAGIER, Daniel, «¿Es posible formular un estándar de prueba preciso y objetivo? Algunas dudas desde un enfoque argumentativo de la prueba», inédito.
- 2003: «Hechos y argumentos (Racionalidad epistemológica y prueba de los hechos en el proceso penal) (II) », en *Jueces para la democracia*, número 47.
- GRANDE, Elisabetta, 2008: «Dances of Criminal Justice: Thoughts on Systemic Differences and the Search for the Truth», en JACKSON, John, LANGER, Máximo y TILLERS, Peter, *Crime, Procedure and Evidence in a Comparative and International Context – Essays in Honour of Professor Mirjan Damaška*, Oxford: Hart Publishing.
- GRIMALTOS, Tobías, 2009, «Creencia, aceptación y conocimiento» *Episteme NS*, vol. 29, nro. 1,.: 35-50.
- GUZMÁN, Nicolás, 2005, «La verdad y el procedimiento abreviado», en MAIER, Julio B. J. y BOVINO, Alberto (comps.), *El procedimiento abreviado*, Buenos Aires: Editores del Puerto.
- HAACK, Susan, 1978, *Philosophy of logics*, Cambridge: Cambridge University Press.
- 2013: «El probabilismo jurídico: una disensión epistemológica», en VÁZQUEZ, Carmen (ed.), *Estándares de prueba y prueba científica*, Madrid: Marcial Pons, 2013.: 65-98.
- 2014: *Evidence matters: science, proof, and truth in the law*, New York: Cambridge University Press.
- KAGAN, Robert A., 2001, *Adversarial legalism: the American way of law*, Harvard University Press.

- KAMISAR, Yale, LAFAVE, Wayne R., ISRAEL, Jerold H., KING, Nancy J. y KERR, Orin S., 2008: *Advanced Criminal Procedure*, EEUU: Thompson/West.
- KIPNIS, Kenneth, 1976: «Criminal Justice and the Negotiated Plea», *Ethics*, 86: 93-106.
- LANGBEIN, John H., 1978: «Torture and Plea Bargaining», *The University of Chicago Law Review*, vol. 46, pp. 3 y siguientes. Citado por su traducción al castellano: «Tortura y plea bargaining», en MAIER, Julio B. J. y BOVINO, Alberto (comps.), *El procedimiento abreviado*, Buenos Aires: Editores del Puerto, 2005: 3-29.
- LANGER, Máximo, 2004: «From Legal Transplants to Legal Translations: The Globalization of Plea Bargaining and the Americanization Thesis in Criminal Procedure», *Harvard International Law Journal*, 2004, vol. 45. Citado por su traducción al castellano de Lucas TASSARA: «De los trasplantes a las traducciones legales: la globalización del plea bargaining y la tesis de la 'americanización' en el proceso penal», *Revista Discusiones*, nro. 21, 2018: 25-134.
- 2005: «La dicotomía acusatorio-inquisitivo y la importación de mecanismos procesales de la tradición jurídica anglosajona. Algunas reflexiones a partir del procedimiento abreviado» en MAIER, Julio B. J. y BOVINO, Alberto (comps.), *El procedimiento abreviado*, Buenos Aires: Editores del Puerto.
- 2006: «Rethinking Plea Bargaining: The Practice and Reform of Prosecutorial Adjudication in American Criminal Procedure», 33 AM. J. CRIM. L. 223, 248. Traducido al castellano por Agustín VARELA y E. Matías DÍAZ, en «Repensando el Plea Bargaining: La Práctica y la Reforma del Juzgamiento por los Fiscales en el Proceso Penal Estadounidense», en LANGER, Máximo, ¿Para qué sirve un modelo procesal? Ensayos sobre los sistemas acusatorio e inquisitivo, Madrid: Marcial Pons, en prensa.
- 2019: «Plea bargaining, trial-avoiding conviction mechanisms, and the global adminitratization of criminal convictions», *Annu. Rev. Criminol.*, DOI: 10.1146/annurev-criminol-032317-092255.
- LANGER, Máximo y ROACH, Kent, 2013: «Derechos en el procedimiento penal: un estudio sobre la convergencia de los sistemas y el derecho del imputado a acceder a los elementos de prueba», en Mark Tushnet, Thomas Fleiner & Cheryl Saunders, HANDBOOK OF CONSTITUTIONAL LAW 273, Routledge.
- LASCURAÍN SÁNCHEZ, Juan Antonio e INCHAUSTI, Facundo Gastón, 2018: «¿Por qué se conforman los inocentes?», *InDret – Revista para el análisis del Derecho*, Barcelona.
- LAUDAN, Larry, 2006: *Truth, Error and Criminal Law*, Cambridge University Press. Citado por su traducción al castellano por Carmen Vázquez y Édgar Aguilera, *Verdad, error y proceso penal*, Madrid: Marcial Pons, 2013.
- LIMARDO, Alan, 2018: «Control judicial de la acusación: Alcance y operatividad respecto del mérito de la prueba», en RÚA, Gonzalo (dir.), LIMARDO, Alan y ROVATTI, Pablo (coords.), *Estudios sobre el sistema penal adversarial*, Buenos Aires: Ed. Didot: 95-126.
- LIPPKE, Richard L., 2011: *The Ethics of Plea Bargaining*, Oxford: Oxford University Press.
- MAFFEI, Stefano, 2004: «Negotiations 'on Evidence' and Negotiations 'on sentence' – Adversarial Experiments in Italian Criminal Procedure», *Journal of International Criminal Justice* 2.
- MAGARIÑOS, H. Mario, 1999: «El 'juicio previo' de la Constitución Nacional y el 'juicio abreviado' – Ley 24825», en *Cuadernos de Doctrina y Jurisprudencia Penal*, Ad Hoc, año V. nro. 9-B.
- MAIER, Julio B. J., 2004: *Derecho Procesal Penal*, t. I. Fundamentos, Buenos Aires: Editores del Puerto.
- 2011: *Derecho procesal penal*, T. III, Buenos Aires: Editores del Puerto.
- PASTOR, Daniel, 2009: *El plazo razonable en el proceso del Estado de Derecho*, Buenos Aires: Ad Hoc.
- PETEGORSKY, Michael Nasser, 2013: «Plea Bargaining in the Dark: The Duty to Disclose Exculpatory Brady Evidence during Plea Bargaining», 81 *Fordham L. Review*: 3599-3650.
- REDLICH, Allison D. y SHTEYNBERG, Reveka V., 2016: «To plead or Not to Plead: A Comparison of Juvenile and Adult True and False Plea Decisions», *Law and Human Behavior*, vol. 40, nro. 6: 611-625.
- SANCINETTI, Marcelo A., 2010: «Avenimiento y mediación: ¿la pena como 'objeto de negocios jurídicos'?», en *elDial.com*, Doctrina, DC138 F.
- SCHAFFER, Jonathan, 2005: «What shifts? Thresholds, Standards, or Alternatives?», en Preyer, Gerhard & Peter, Georg (eds.) *Contextualism in Philosophy: Knowledge, Meaning and Truth*.

- TARUFFO, M., 1992: *La prova dei fatti giuridici*, Milano: Giuffrè Editore. Citado por la versión castellana de FERRER BELTRÁN, Jordi, *La prueba de los hechos*, Trotta, Madrid, 2011.
- 2008: *La prueba*, Madrid: Marcial Pons.
- 2010: «Conocimiento científico y estándares de la prueba judicial», en PARCERO, Juan A. Cruz y LAUDAN, Larry (comps.), *Prueba y estándares de prueba en el derecho*, UNAM, Instituto de Investigaciones Filosóficas.
- TIMMONS, Mark, 2002: *Moral Theory: An introduction*, Oxford: Rowman and Littlefield.
- VELEDA, Diana, 2019: «Delimitación del control judicial en el avenimiento. Su relación con el carácter acusatorio del proceso penal de la CABA y con el mejoramiento de las condiciones de legitimidad de los acuerdos», en DÍAZ, E. Matías, PEREL, Martín (dirs.) y VELEDA, Diana (coord.), *Aspectos relevantes del juicio en el procedimiento penal de la CABA*, Buenos Aires: Ad-Hoc.
- WEIGEND, Thomas, 2008: «The Decay of the Inquisitorial Ideal: Plea Bargaining Invades German Criminal Procedure», en JACKSON, John, LANGER, Máximo y TILLERS, Peter (eds.), *Crime, Procedure and Evidence in a Comparative and International Context – Essays in Honour of Professor Mirjan Damaška*, Hart Publishing.
- WERTHEIMER, Alan, 1979: «The prosecutor and the gunman», *Ethics*, volume 89, nro. 3, 269-279.
- WILFORD, Miko M., 2014: «Bluffed by the dealer: Distinguishing false pleas from false confessions», *Graduate Theses and Dissertations*, Iowa State University, 13877.

LA RELACIÓN ENTRE LA PRÁCTICA DE LOS SISTEMAS PENALES Y LA DETERMINACIÓN DE LOS ESTÁNDARES DE PRUEBA*

Santiago Eyherabide

Universidad de Buenos Aires

seyherabide@gmail.com.

RESUMEN: Este artículo ofrece una propuesta sobre cómo el estándar de prueba se configura en la práctica cotidiana de los sistemas penales. Sostengo que, junto a la formulación de orden general en el nivel normativo, existe una dimensión práctica del estándar de prueba para las distintas modalidades de comisión de delitos que es establecida por los tribunales. Luego de computar los efectos que producen la vaguedad de la formulación del estándar y la imprecisión de los resultados probatorios, se argumenta que, en función de los métodos regulares de investigación de los distintos delitos, existen «conjuntos probatorios de referencia» y «contextos de desdicha» que estructuralmente se repiten en los procesos y determinan cómo los culpables e inocentes acusados de un delito quedan involucrados en un proceso. Luego, son ciertas configuraciones de tales conjuntos las que, en la práctica de los sistemas judiciales, se exige para dictar una condena. Por otra parte, se desarrolla cómo los modelos de juicio por jurados y de jueces profesionales intentan garantizarse la satisfacción del estándar. Finalmente, se analizan los efectos que producen los errores de los operadores judiciales en la reunión de la prueba y en las inferencias que realizan con ella.

PALABRAS CLAVE: prueba; estándar de prueba.

* Agradezco a Diego Papayannis, mi tutor en la Maestría de Razonamiento Probatorio de la Universidad de Girona y amigo, con quien discutí y enriquecí mucho las ideas centrales de este trabajo.

THE RELATIONSHIP BETWEEN THE CRIMINAL JUSTICE SYSTEM PRACTICES AND THE DETERMINATION OF THE STANDARD OF PROOF

ABSTRACT: This article offers a proposal on how the standard of proof shapes itself in the daily practice of the criminal justice system. I claim that, together with the general formulation at the normative level, there is a practical dimension of the standard of proof for the different ways the commission of a crime is established by courts. After computing the effects produced by the vagueness of the formulation of the standard and the imprecision of the evidentiary results, it is argued that, based on the regular methods of investigation of the different crimes, there are «evidentiary reference sets» and «contexts of misfortune» structurally repeated in the cases that determine how the defendant, whether guilty or innocent, could be involved in a process. Therefore, it is certain configurations of such sets, in the practice of judicial systems, that are required to issue a sentence. On the other hand, it is developed how the models of trial by juries and professional judges try to guarantee the satisfaction of the standard. Finally, the effects produced by the errors of the judicial operators in the gathering of evidence and in the inferences they make with it are considered.

KEYWORDS: proof; standard of proof.

SUMARIO: I. PRIMERA PARTE.— 1.INTRODUCCIÓN— 2.CONCEPTO Y FUNCIONES DEL ESTÁNDAR: 2.1. Probar. 2.2. El estándar.— 3. EL ENFOQUE DE LARRY LAUDAN SOBRE LA FUNCIÓN DEL ESTÁNDAR DE PRUEBA.— 4. EFECTOS DE LA VAGUEDAD DEL ESTÁNDAR Y DE LA IMPRECISIÓN DEL RESULTADO PROBATORIO.— II.SEGUNDA PARTE.— 1. EL CONJUNTO PROBATORIO DE REFERENCIA, LAS PRUEBAS SINGULARES Y LOS CONTEXTOS DE DESDICHA.— 2. LA MALA CALIDAD EPISTÉMICA DEL SISTEMA: 2.1. Incapacidad para reunir pruebas robustas. 2.2. Mala calidad de las inferencias— III. TERCERA PARTE.— 1. LA DIMENSIÓN PRÁCTICA DEL ESTÁNDAR Y LOS DOS PROCEDIMIENTOS TRADICIONALES PARA DECIDIR: JUECES PROFESIONALES Y EL JUICIO POR JURADOS LEGOS.— IV. CUARTA PARTE.— 1. UNA BREVE REFLEXIÓN FINAL SOBRE LA VAGUEDAD DEL ESTÁNDAR Y LA POLÍTICA CRIMINAL.— 2. CONCLUSIÓN.— IV: BIBLIOGRAFÍA

RECOMMENDED CITATION: EYHERABIDE, SANTIAGO, 2020: «La relación entre la práctica de los sistemas penales y la determinación de los estándares de prueba», in *Quaestio facti*, 2: 185-223. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22457

I. PRIMERA PARTE

1. INTRODUCCIÓN

Me propongo analizar distintas variables que afectan de modo relevante la configuración del estándar de prueba para la condena penal.

Se acepta en general que el estándar de prueba para una decisión puede ser establecido con distintos niveles de exigencia y que una consecuencia significativa del nivel elegido será su incidencia sobre las dos clases de errores que en ocasiones se cometerán en un sistema penal (concretamente, la cantidad de personas inocentes que resultarán condenadas -falsos positivos- y la cantidad de personas culpables que resultarán absueltas —falsos negativos—).

El punto de partida es que cuanto más exigente sea un estándar de prueba, menos errores se producirán por condena de inocentes pero más errores sobre absolución de culpables e, inversamente, a menor exigencia del estándar, habrá más culpables condenados pero también más inocentes condenados.

De acuerdo con esta tesis, la selección del nivel de exigencia donde habrá de establecerse el estándar de prueba debe tener en cuenta fundamentalmente la razón proporcional socialmente aceptada de falsos negativos y falsos positivos.

Ahora bien, en la presentación de este modelo, en general, se presupone que existe libertad para elegir el punto concreto donde se desea fijar el estándar de prueba.

En este trabajo me concentraré justamente en un grupo de variables que afectan, o bien la configuración del estándar o bien el set de opciones disponibles para establecer el estándar de prueba en un sistema judicial.

La primera variable que intentaré computar dentro del modelo será la vaguedad insalvable que presenta la fórmula lingüística con la que se plasma el estándar de prueba. En segundo lugar, me centraré en el margen de imprecisión, también insalvable, que presenta el nivel de acreditación alcanzado respecto de una proposición fáctica a partir de un conjunto probatorio dado.

Luego introduciré una cuestión vinculada con el modo en que los culpables y los inocentes resultan sospechosos de haber cometido un delito en la práctica de los sistemas judiciales. A partir de ejemplos, veremos que para ciertos delitos, regularmente será muy difícil discernir a los inocentes acusados de los culpables acusados, mientras que para otros delitos ello será bastante sencillo. Desarrollaré aquí cómo estas configuraciones de la prueba para los inocentes y los culpables según cada delito condiciona fuertemente las razones proporcionales de inocentes condenados y culpables absueltos que el estándar puede establecer.

A continuación, introduciré dos variables relacionadas con la calidad epistémica del sistema judicial que generan nuevos condicionamientos a las posibilidades de establecer el estándar y de elegir la razón proporcional de errores. Estas variables son la capacidad para reunir pruebas robustas y la habilidad para realizar inferencias adecuadas a partir de ellas. Veremos los efectos de un mal desempeño del sistema en estos dos aspectos desde el punto de vista de su operación práctica.

Finalmente, analizaré la implementación del estándar de prueba, es decir, la cuestión de *cómo* cada sistema intenta garantizar en la práctica que el estándar de prueba en un caso concreto se encuentre satisfecho. Me aproximaré a esta cuestión destacan-

do las diferencias que existen entre los sistemas de jueces profesionales y los de juicio por jurados legos. A partir de este análisis sostendré la conclusión de que ese aspecto metodológico, al que llamaré indistintamente «dimensión práctica» o «dimensión procedimental» del estándar de prueba es en realidad constitutiva del estándar mismo, a la par de su establecimiento mediante una expresión en palabras, a la que llamaré, también indistintamente, «dimensión de pauta» o «formulación lingüística» del estándar de prueba.

2. CONCEPTO Y FUNCIONES DEL ESTÁNDAR

2.1. Probar

Para poner adecuadamente en contexto los objetivos que he trazado en la introducción, es necesario comenzar observando cómo llegamos a necesitar el estándar de prueba. Para ello hay que partir, muy sintéticamente, de la noción de prueba.

Entendemos el término «probar» en distintos sentidos (GASCÓN ABELLÁN, 1999: 83, FERRER BELTRÁN, 2003: 27-28). Me interesa aquí destacar que «probar» involucra una actividad por la que intentamos dar cuenta, demostrar a un observador (que puede ser el propio sujeto que realiza la actividad), que cierto dato o hecho es o se ha producido de cierto modo o que cierta ley de la naturaleza es de cierto modo.

En derecho, donde en general intentamos probar datos o hechos; tales datos suelen ser de distinto tipo. GONZÁLEZ LAGIER (2005: 22) ha realizado una excelente clasificación de aquello que intentamos probar en los procesos judiciales y que abarca desde acciones y omisiones hasta estados de cosas y estados psicológicos. También, en ocasiones, leyes de la naturaleza.

Por otra parte, esa actividad tendiente a demostrar ante un observador que cierto dato es de cierto modo tiene una estructura determinada. Probar, es una actividad que se realiza mediante «argumentos probatorios».

De acuerdo con ANDERSON, SCHUM, TWINNING (1991: 95) un argumento probatorio contiene cuatro elementos: (1) *un dato* a partir del cual se realiza (2) *una inferencia* por la cual se pretende dar apoyo a (3) *una conclusión*. A su vez, la inferencia siempre está basada en (4) *una generalización* que le da sustento. Este esquema, con terminología ligeramente distinta, es también explicado por GONZÁLEZ LAGIER (2003: 35), partiendo todos de las ideas de TOULMIN (1964) ¹.

Por más arbitrariedad o sin sentido que ciertos argumentos probatorios puedan presentar, como sucede con la prueba mágica (GASCÓN ABELLÁN, 1999: 8), p ej. el

¹ Obviaré aquí las etapas de razonamiento creativo y razonamiento crítico con las que se opera para investigar y descubrir. Al respecto véase también ANDERSON, SCHUM, TWINNING (1991, Cap 2).

chaman despertó con frío y esa es una señal que prueba que el acusado es culpable, su estructura se mantiene.

Asimismo, prácticamente siempre, el hecho que se desea probar involucra más de un argumento probatorio basado en distintos datos, así como la construcción de argumentos que se nutren de varios datos a la vez. Los argumentos suelen ser también escalonados o de varios pasos; y todo ello provoca que en la prueba de un hecho, y muy especialmente en la prueba de los hechos que interesan a los procesos penales, se construya un árbol argumental.

Este andamiaje puede organizarse esquemáticamente con forma piramidal con los *probanda* últimos en la cúspide, luego los penúltimos, etc. y dar lugar a un esquema ascendente de razonamientos y argumentos compuestos por numerosos datos probatorios, inferencias realizadas con ellos, conclusiones (intermedias o finales), las generalizaciones que le dan sustento y la prueba auxiliar vinculada a las credenciales probatorias de los datos asociados con los *probanda*².

A esta compleja estructura hay que añadir que el tipo de razonamiento que subyace a los argumentos probatorios individualmente considerados, razonamiento inductivo³, impide también las conclusiones seguras, por lo que hay que aceptar, forzosamente, que las conclusiones sobre hechos probados son siempre, en su esencia, probabilísticas.

Todas estas características dan por resultado que un conjunto de pruebas vinculadas a la acreditación de un hecho será casi siempre *incompleto* (nunca tenemos toda la prueba), *no concluyente* (un elemento de prueba podría favorecer simultáneamente más de una conclusión), *ambiguo* (no podemos afirmar con seguridad qué nos está diciendo la prueba), *discrepante* (algunos elementos favorecen una proposición mientras que otros favorecen otra proposición) y finalmente, con *diversos grados de credibilidad imperfectos* (ANDERSON, SCHUM, TWINNING, 1991: 303).

En definitiva, por más exitosamente trabajada que esté la prueba de un hecho, su grado de acreditación siempre está sujeto a algún margen de error, y los razonamientos probatorios efectuados con esas pruebas de ningún modo brindan resultados seguros y definitivos. Al contrario, el nivel/grado/umbral/punto/límite de acreditación que se logra sobre un hecho es esencialmente variable.

² En ANDERSON, SCHUM, TWINNING, 1991: 182 puede verse gráficamente representado un esquema como el indicado.

³ Una caracterización de las clases de razonamiento excede los límites de este trabajo. Para la afirmación del texto, basta tener en cuenta que el razonamiento probatorio encierra, desde una perspectiva de lógica deductiva, una falacia de afirmación del consecuente, por lo que no puede garantizar analíticamente la verdad de las conclusiones.

2.2. El estándar

Este es el terreno donde surge, frente a la necesidad de tomar decisiones, la necesidad de los estándares de prueba.

En el proceso penal se toman múltiples decisiones. De ellas, la condena o absolución del acusado es la decisión más destacada, pero en realidad son muy numerosas las decisiones que deben adoptarse y en las que hay hechos que forman parte fundamental de la justificación de tales decisiones (formular la imputación inicial, detener preventivamente al acusado, allanar una propiedad, intervenir un teléfono, realizar compulsivamente medidas de prueba, elevar un caso a juicio oral, etc.). En efecto, usamos estándares para innumerables decisiones de nuestra vida cotidiana (CALLEN, 2008: 166).

Ahora bien, como los momentos para tomar las decisiones no son prorrogables al infinito y como el hecho que debe servir de fundamento a la decisión no está demostrado de un modo cierto, objetivo o absoluto, surge la necesidad del estándar de prueba en derecho, es decir, el establecimiento de un punto, o de un área cuanto menos, en la que ciertos datos o hechos son considerados como suficientemente demostrados a los efectos de la fundamentación (jurídicamente válida) de una determinada decisión. Es algo así como la decisión relativa a «cuánto» se debe probar un hecho para tomar una decisión de modo legalmente justificado.

En este sentido, FERRER BELTRÁN (2007: 139) ha dicho que se trata de «fijar el umbral a partir del cual aceptaremos la hipótesis como probada. Es decir, debemos determinar el grado de probabilidad suficiente para dar por probada la hipótesis».

Para Larry LAUDAN el término se refiere a

Las teorías o modelos cuyo propósito es proponer una serie de condiciones de membresía para PP [Proposiciones Probadas] que efectivamente posean las cualidades de restringir al máximo la vaguedad de sus términos y de referirse al aspecto que debe revestir la evidencia (en el sentido antes mencionado) que le proporciona soporte a p (LAUDAN, 2011: 319).

Más sencillamente, el propio LAUDAN lo ha definido diciendo que «un estándar de prueba especifica el umbral mínimo que ha de ser satisfecho a los efectos de aseverar que una hipótesis ha sido aprobada» (LAUDAN, 2006: 104). TARUFFO (2010: 247) lo asocia con la idea de «grado de confirmación» que se exige para tener por probado un hecho.

AGUILERA (2008: 108), que parte de una definición como la de Laudan, explica luego que

al establecer un determinado E P [estándar de prueba] nos encontramos tomando una decisión respecto de las condiciones de configuración que la evidencia en apoyo de la proposición en cuestión debe satisfacer para ciertos fines. Sabemos también que la posibilidad de modular la intensidad de esas condiciones a la baja o a la alta, nos permite hablar de estándares más severos que otros.

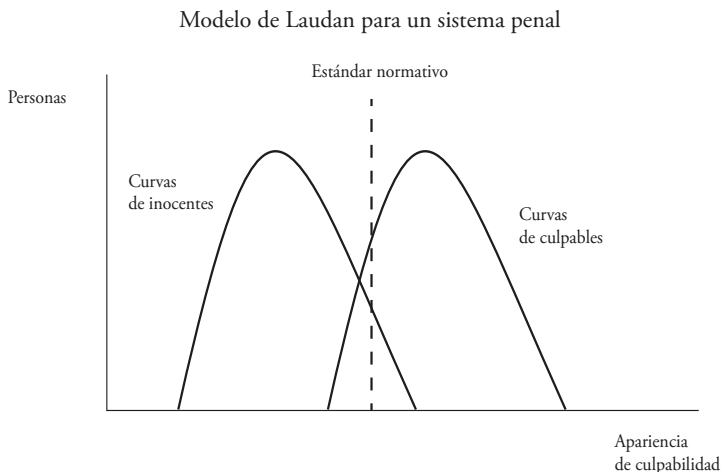
3. EL ENFOQUE DE LARRY LAUDAN SOBRE LA FUNCIÓN DEL ESTÁNDAR DE PRUEBA

Una consecuencia de las características de la prueba enunciadas en el punto anterior es que, aunque el objetivo en la decisión final de un juicio es que los culpables sean sancionados y los inocentes absueltos, en algunos casos personas inocentes podrían pese a ello tener prueba en su contra suficiente para que se tenga por probada su culpabilidad por encima del estándar de prueba establecido; y que por otra parte personas culpables podrían, pese a ser culpables, enfrentar acusaciones que no logran reunir prueba suficiente para superar el estándar. En definitiva, al menos en algunas ocasiones, errar es un mal del que es imposible escapar por completo.

Larry LAUDAN (2006) es uno de los autores que se ha detenido en esta cuestión y la traducción de su libro al español motorizó el debate en nuestra lengua. Él parte como vimos de una definición del estándar de prueba como un *umbral mínimo* de justificación que se debe satisfacer para afirmar que una hipótesis se encuentra probada. Es decir, una caracterización semejante al resto de las enumeradas.

Pero advierte enseguida en el estándar de prueba una función central: la de la distribución de los errores que se cometen en el sistema.

LAUDAN introduce como punto de partida una relación susceptible de representación gráfica, sobre la que trabajaré repetidamente en este texto. Se trata de la situación de las personas inocentes y culpables a las que se les imputa un delito y su representación en un eje de coordenadas según su apariencia de culpabilidad. La representación gráfica, aproximadamente, es la siguiente (LAUDAN, 2006: 109):⁴



⁴ LAUDAN, 2006: 109.

El modelo representa sobre el eje x la apariencia de culpabilidad que se logra obtener con la prueba respecto de las personas acusadas de un delito. El eje vertical contabiliza la cantidad de personas. Las curvas de inocentes y culpables muestran la distribución de las cantidades de personas de cada tipo (inocentes y culpables) en función de la apariencia de culpabilidad resultante de la prueba disponible. La línea vertical punteada ilustra el punto en el que se ha decidido fijar el estándar de prueba para la condena penal. A su derecha las personas serán condenadas. A su izquierda serán absueltas.

Como se ve allí, si moviéramos la línea del estándar de prueba, variaría la cantidad de personas condenadas y absueltas, incluyendo por supuesto cantidades de inocentes condenados y de culpables absueltos⁵. Un estándar situado más a la izquierda provocará la condena de más culpables pero también de más inocentes. Un estándar situado más a la derecha en el eje de la apariencia de culpabilidad importará la absolución de más inocentes pero también la absolución de más culpables.

Para LAUDAN (2006: 60), entonces, el estándar de prueba es un elemento central de lo que denomina la *doctrina de la distribución del error*, junto a otros conceptos que para él poseen también esa misma función de distribuir los errores que se cometen en un sistema penal (los otros son la presunción de inocencia, el beneficio de la duda y la carga de la prueba, no me detendré en ellos aquí)⁶.

Así, la fijación del nivel de exigencia de un estándar es una decisión sobre la distribución del error judicial que establece, en definitiva, cuántos inocentes condenados y cuántos culpables absueltos habrá en un sistema penal (Laudan, 2006: 109). Luego, se propone encontrar una forma no arbitraria de determinar un estándar de prueba para la decisión de condena o absolución de un acusado que, dando todo el beneficio de la duda que se fije como necesario, no arroje más falsos positivos (condenas erróneas) que los que la sociedad esté dispuesta a aceptar.

Para él esta decisión no afecta, sin embargo, la cantidad total de errores del sistema. Antes bien, la minimización del total de errores del sistema debe lograrse reuniendo previamente los mejores y más robustos conjuntos de prueba que sea posible a través de reglas procesales que favorezcan adecuadamente ese propósito. Es esto último lo que define la minimización del error (LAUDAN, 2006: 173)⁷.

⁵ Asumiré para los efectos de este trabajo que es posible desplazar el estándar de prueba de un nivel de exigencia a otro manteniendo todo lo demás constante. Un desarrollo profundo del espectro de consecuencias que en realidad podría generar un cambio del estándar de prueba puede verse en KAPLOW, 2012.

⁶ LAUDAN desarrolla estos conceptos en el cap. IV, pp. 137 yss.

⁷ En el mismo sentido FERRER BELTRÁN, 2013 (en particular nota 13)

4. EFECTOS DE LA VAGUEDAD DEL ESTÁNDAR Y DE LA IMPRECISIÓN DEL RESULTADO PROBATORIO

Deseo introducir aquí las primeras dos variables que impactan de modo relevante en el modelo explicado. Específicamente, veremos cómo la vaguedad afecta el «punto» donde se fija el estándar y cómo la imprecisión del resultado probatorio afecta la «línea» de las curvas de inocentes y culpables.

GONZÁLEZ LAGIER (2020: 8) asume una caracterización similar a las que he mencionado antes sobre el concepto de estándar de prueba. Él lo define como un criterio que debe funcionar como *umbral* que permita discriminar entre lo que se da por suficientemente probado (y lo que no) a los fines de tomar una decisión. Pero es importante para el propósito de este trabajo destacar que para González Lagier cierto grado de vaguedad siempre acompaña la formulación del estándar. En efecto, demostrar que los intentos de fijar un estándar objetivo y preciso no es posible constituye el núcleo central de su texto y terminará por afirmar que la vaguedad es llanamente insalvable en la formulación del estándar de prueba.

En línea con ello, de aquí en adelante, tomo como punto de partida para este trabajo que a nivel normativo se puede formular un estándar genérico de culpabilidad para un sistema penal en su conjunto, más o menos estricto, pero siempre con grados de vaguedad que no se pueden eliminar. Acepto entonces la observación de GONZÁLEZ LAGIER respecto de este punto.

Esto ya tiene una primera implicancia. Dada la vaguedad del estándar, existirá un conjunto de casos que estarán comprendidos por su alcance tanto si los imputados son condenados como si son absueltos, pues se hallan dentro de esa *zona gris* que crea la vaguedad que necesariamente acompaña al estándar de prueba formulado para un sistema dado.

Esta cuestión puede presentarse a primera vista en tensión con el principio de inocencia. Si un caso está correctamente resuelto como absolución, teniendo en cuenta la regla de decisión *in dubio pro reo*, ese mismo caso jamás podría estar correctamente resuelto como culpabilidad sin contrariar esta pauta. Sin embargo, el problema radica en que por regla en estos casos en realidad no hay duda o incertidumbre sino seguridad, tanto para quienes creen situarse por debajo del estándar de condena, como para quienes creen estar por encima de él. En efecto, cotidianamente en los sistemas penales las instancias superiores revocan sentencias (de absolución o de condena) de la instancia anterior y dictan una solución de signo contrario sin que ello provoque luego conflictos a la luz de principios constitucionales ni se inicien sumarios o procedimientos de destitución de los jueces que habían resuelto «erróneamente».

Se trata, en definitiva, de un problema de interpretación en el que los distintos intérpretes no albergan ninguna duda acerca de que lo que una determinada proposición significa es lo que cada uno de ellos sostiene.

Otro inconveniente diferente que se presenta es la imposibilidad de establecer con exactitud el grado de corroboración que una hipótesis ha alcanzado con la prueba disponible. En otras palabras, la imprecisión que reviste la «apariencia de culpabilidad» dado un conjunto probatorio en un caso concreto.

Aquellas características que presentan los conjuntos probatorios y que expresara anteriormente siguiendo a ANDERSON, SCHUM Y TWINNING, 1999 (*incompleto, no concluyente, discrepante y con diversos grados de credibilidad imperfectos*) resultan en un margen de indeterminación en el grado de corroboración obtenido luego de producida la prueba. En algunos aspectos este margen puede ser atenuado o agravado según la calidad que tenga un sistema (tanto en su diseño como en su ejecución), pero es imposible eliminarlo por completo.

De este modo, para cada sujeto representado en el gráfico (sea de la curva de inocentes o de culpables), su apariencia de culpabilidad en función del conjunto de pruebas que enfrenta, será al menos levemente imprecisa, levemente indeterminada.

Vale aclarar que también frente a esta cuestión alguien podría sostener que la prueba avala la culpabilidad del acusado con mayor grado de respaldo mientras que otra persona sostiene que la prueba avala la culpabilidad en una medida menor. Pero si ambos están dentro del margen de valoración racional del conjunto probatorio, tampoco existirán objeciones viables desde la perspectiva de la regla *in dubio pro reo*. Es que, en definitiva, la norma no puede desconocer la relatividad del conocimiento (relatividad que es general, más allá del ámbito judicial), ni la variabilidad de las conclusiones que distintos observadores somos capaces de extraer sobre un acervo probatorio dado.

En definitiva, la vaguedad del estándar y la imprecisión del resultado probatorio no se traducen necesariamente en actitudes proposicionales de duda por parte de los sujetos cognoscentes. Quizás desde el punto de vista de su actitud interna, tanto los que creen ver en un conjunto probatorio la certeza de culpabilidad como aquellos que extraen una certeza de inocencia (o bien una duda que los llevará a absolver), tienen muy clara y firme la decisión que cabe tomar en virtud de la prueba disponible y el alcance semántico que debe darse al estándar de prueba.

Si todos ellos han operado en la valoración probatoria conforme a las reglas dadas, aunque seguramente uno tendrá razón y el resto estará equivocado sobre lo que efectivamente sucedió en el mundo, dado el nivel de conocimiento disponible en un momento dado sobre cierta materia o tema, todos pueden haber operado correctamente. Ello se debe a que tanto la vaguedad del estándar como la imprecisión del resultado probatorio alcanzado crean áreas o segmentos donde debería haber puntos o líneas.

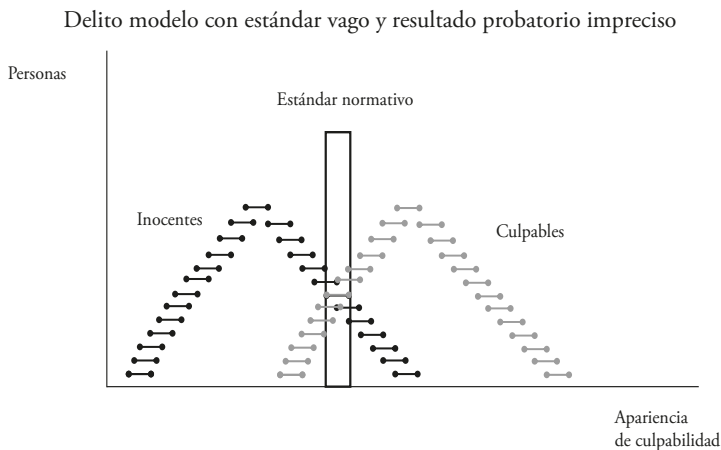
Así, el límite que separa la revocación de una decisión y su reemplazo por una de signo contrario sin otras consecuencias, de la que implica una revocación por aplicación de la doctrina de la arbitrariedad y que en algunos casos podría llegar incluso al inicio de un procedimiento disciplinario o penal contra quien tomó la decisión reputada como equivocada, se encuentra en si el juez que decidió «erróneamente» se

encontraba dentro del campo semántico del estándar y cumpliendo los parámetros, criterios, y pautas aceptados para valorar la prueba dentro del contexto de conocimientos y técnicas dadas, o si directamente tomó una resolución completamente arbitraria e irracional.

Ahora bien, ¿cómo se traduce la vaguedad de la formulación del estándar gráficamente en el esquema anterior? Transforma el estándar en un segmento, en lugar de un punto específico.

Por otra parte, ¿cómo se traduce la indeterminación del punto exacto de respaldo probatorio logrado con la prueba? Gráficamente, ensancha el grosor de las curvas correspondientes sobre el eje horizontal (ya no serán esas delgadas líneas sino unas gruesas franjas). Ello sucede porque cada sujeto (cada valor del eje y), sea inocente o culpable, enfrenta una apariencia de culpabilidad constituida por un segmento -el margen de imprecisión- y no por un punto. La franja se forma al trazar líneas que unen los puntos de apariencia de culpabilidad mínimos y máximos de cada uno de los sujetos.

Veamos la representación gráfica de ambos efectos:



El gráfico incorpora, por un lado, la vaguedad del estándar de prueba, que por esa razón queda expresado como un segmento en el x (apariencia de culpabilidad) para cada valor de y. Lo propio sucede con las curvas de inocentes y culpables, pues el resultado probatorio en cuanto a apariencia de culpabilidad es también un segmento para cada valor de y.

Aceptar que un estándar siempre estará formulado de este modo y que los resultados precisos de los conjuntos probatorios son también más o menos inciertos entraña incertidumbre jurídica. De hecho, bajo ciertas condiciones, un caso exactamente igual, podría resultar condena en un momento t_1 y absolución en t_2 . Y ello, como

indiqué anteriormente, solo será arbitrariedad si hubo errores graves o una decisión deliberadamente tendenciosa, pero habrá simplemente *poder* judicial cuando la decisión sea el resultado de un trabajo realizado de buena fe dentro de los procedimientos correctos desde una perspectiva legal y metodológica. En este sentido, entiendo que la representación anterior no es otra cosa que el reflejo gráfico de la idea del *poder de interpretación y del poder de comprobación* de FERRAJOLI (1990: 38 y 129), ese ámbito donde el resultado probatorio de un elemento de prueba o de un acervo probatorio a la luz del antecedente de hecho de una norma, efectuadas todas las intervenciones correspondientes (defensa, fiscal, juez) arroja un ámbito de penumbra en el que el decisor puede elegir sin impugnabilidad racionalmente dirimible la solución que cabe adoptar.

Ahora bien, al incorporar la vaguedad del estándar y la imprecisión del resultado probatorio hemos hecho endógenas al modelo dos cuestiones que son relevantes y que estaban fuera de él. Asimismo, estas cuestiones afectan el potencial del estándar de prueba para fijar un valor deseado preciso en la razón proporcional entre inocentes condenados y culpables absueltos.

Estas dos variables, en definitiva, implican que con el estándar se podrán fijar márgenes de la razón proporcional de errores, valores aproximados de ella. Pero no una ratio puntual.

No obstante, es importante señalar que un estándar que ocupe un segmento muy ancho es socialmente muy perjudicial porque convierte a la justicia penal en un sistema de azar. En este punto resultan esenciales los aportes de Jordi FERRER (2007: 139 y 2013: 4) para reducir la vaguedad en la formulación lingüística del estándar (y para demostrar que con él se puede establecer múltiples niveles de exigencia).

II. SEGUNDA PARTE

Vale detenerse ahora en otras dos cuestiones que LAUDAN explícitamente menciona pero que no forman parte de lo que pretende desarrollar en su modelo.

La primera cuestión es la forma y ubicación que toman las curvas de acusados inocentes y culpables. Expresamente señala que «la forma precisa de la curva no es importante para nuestro propósito: su pico puede ser más alto, sus colas más pequeñas, su dispersión puede ser menor, su posición específica a lo largo del eje de las X puede ser diferente... » (LAUDAN, 2006: 107)⁸.

⁸ Él retomará esta cuestión más adelante (p. 173) pero desde una perspectiva de diseño del sistema, no desde una perspectiva operativa como la que desarrollaré ahora. Volveré sobre esto luego.

La segunda cuestión que LAUDAN (2006: 115) menciona es que la distribución de los errores no depende solo del estándar de prueba sino también de «cuán robustas sean las pruebas, así como cuán válidas sean las inferencias que con base en ellas realizan los miembros del jurado». Más allá de aclararlo, sin embargo, tampoco sobre este aspecto continúa el análisis.

A continuación, me propongo estudiar justamente esas dos variables que no forman parte del análisis de Laudan y observar gráficamente el importante impacto que pueden tener (y tienen). Así, representando gráficamente algunos supuestos sobre *la forma precisa de las curvas* de inocentes y culpables y sobre la calidad epistémica de un sistema, trataré de mostrar qué importantes consecuencias tendrían ellas en la distribución del error judicial entre condenas y absoluciones erróneas e incluso, y más importante, en su cantidad total.

1. EL CONJUNTO PROBATORIO DE REFERENCIA, LAS PRUEBAS SINGULARES Y LOS CONTEXTOS DE DESDICHA.

Si hacemos entonces de la «forma precisa» de las curvas y, sobre todo, de «su ubicación precisa a lo largo del eje de las X» el centro de nuestra atención, vale preguntarse cuáles son las variables que determinan esa forma y esa ubicación.

ALLEN (2014: 211), ha reparado puntualmente sobre esta cuestión. Él señala que la forma de las curvas que utilizamos en los modelos dependen en realidad de dos variables empíricas, no analíticas: en primer lugar los tamaños relativos de los subconjuntos de casos que dan lugar a las curvas y en segundo lugar de la probabilidad asignada a los casos por quienes juzgan los hechos. En efecto, desde allí, asumiendo que los subconjuntos de casos de las curvas que efectivamente son sometidos a proceso tienen tamaños diferentes comienza a modelar los gráficos con curvas bien diferentes entre sí en los que la curva de inocentes es considerablemente más pequeña que la de culpables.⁹

Allen realiza luego consideraciones muy relevantes sobre la interacción de los conjuntos de casos y la forma de las curvas con la fijación del estándar de prueba sobre los que volveré en la última parte de este trabajo, pero en este apartado deseo centrarme en esos aspectos empíricos que son relevantes tanto para la conformación de los conjuntos de casos, como para la probabilidad que los juzgadores de los hechos les asignan.

⁹ Allen trabaja a lo largo de todo su artículo, indistintamente para casos civiles y penales, con curvas de «actores» y «demandados» a los que les asiste razón (deserving plaintiffs and deserving defendants), no de inocentes y culpables. Considero que ello no afecta de modo relevante la posibilidad de trasladar sus argumentos sin más a curvas de «inocentes» (demandados a los que les asiste razón) y culpables (casos en los que asiste razón a los actores).

Para ello debe partirse de algunas características del funcionamiento de los sistemas penales en la práctica que considero que definen, en gran parte, cómo son esas curvas.

Comencemos con un ejemplo imaginario —que quizás se parezca a uno real— que transcurre en la sala de un Tribunal, durante la lectura de un fallo, tras la realización del juicio oral por un hecho de secuestro extorsivo:

— Presidente del Tribunal: ... *Condenando, a Cristian Marcos Hoyos, alias Opi, de las demás circunstancias personales obrantes en autos, a la pena de 8 años y seis meses de prisión, por encontrarlo autor penalmente responsable...*

— ¡De ser un imbécil! Completó a viva voz «El Guri» mientras la palma de su mano golpeaba el atril desde el que acompañaba a Opi, esperando la lectura de su propia condena (sería la quinta en cuestión de meses, incluyendo una a cadena perpetua).

— ¡Ooooootra vez! balbuceó Hoyos «Opi» mientras el Presidente del Tribunal, que había interrumpido su lectura del fallo, alzaba la vista por encima de sus lentes con gesto severo.

— *Te dije que no usaras el teléfono del tipo para nada que no sea pedir la gaita.* ¡Gato! ¡Paloma muerta!, le espetó el Guri.

— *Le cambié el chip antes de hacer otros llamados gil.* Le respondió Opi por centésima vez ante el mismo comentario.

— *Bueno, ahora tenés 8 años y seis meses para darte cuenta de que eso da igual pavo.* Cerró Guri con una carcajada estrepitosa.

— *Y usted va a tener 12,* pronosticó el Presidente del Tribunal, visiblemente fastidiado, al Guri.

Si observamos el episodio en lo que aquí interesa, lo que sucedió es que Guri le reclamaba a Hoyos un error de *desempeño* vinculado a la ejecución del hecho de secuestro extorsivo, (bajo modalidad *express*) que habían cometido juntos. En rigor, le reclamaba por su actuación posterior, pues en lugar de vender el teléfono sustraído a la víctima a un tercero extraño, lo había conservado y lo había utilizado (aunque luego de cambiarle el chip de abonado).

Guri sabía que, tal como luego sucedió, las autoridades requerirían a la empresa prestataria de telefonía que informe sobre los chips de líneas telefónicas que hubieran impactado en el aparato sustraído a la víctima. Si los secuestradores comenzaban a usar el teléfono ellos mismos -error usual en la década del 2000-, por más que utilizara chips a nombre de terceros, a través de los interlocutores frecuentes quizás lograrían identificar al usuario real del teléfono. Con esa individualización podrían practicar una rueda de reconocimiento entre el usuario del teléfono y la víctima del secuestro que, en caso de dar positivo, los colocaba ante una cuasi-certidumbre de condena.

Lo que Gurito en definitiva conocía, y que quiero caracterizar aquí, era una parte de lo que denominaré el «contexto probatorio de referencia» (al menos para Argentina y para ese momento dado) que suele recabarse en los secuestros extorsivos llamados *express* (de duración breve -2 a 12 horas-, en general sin lugar de cautiverio

fijo -p ej. circulando en auto, o en un pasillo de una villa, y en los que la víctima fue liberada con vida tras el pago del rescate)¹⁰.

Y conociendo ese «contexto probatorio de referencia» para los secuestros, Gurito había tratado de contrarrestar la posibilidad de ser detectado para asegurar su impunidad, solo que Opi, en lugar de deshacerse del teléfono como le habían indicado, se limitó a cambiar el chip y comenzó a utilizarlo él.

Un caso con este acervo probatorio que no tenga otros aspectos que lo hagan más complejo (reconocimiento ocular defectuoso, inconsistencias del relato de las víctimas, etc.), es un caso que, en la práctica de un Tribunal de jueces profesionales, en el estado de la jurisprudencia en Argentina en ese momento, seguramente satisfaría las exigencias para dictar una condena.

Aun a riesgo de asumir demasiado, sostendré que al menos en la gran mayoría de los casos por infracción de un tipo penal (p ej. tráfico de estupefacientes, abuso sexual, corrupción, contaminación ambiental, etc.) los mecanismos de detección de los hechos y el conjunto de evidencia que suele reunirse guarda bastante parecido entre los casos y es el resultado de algunas de las pruebas críticas —que se repiten en cada investigación por el mismo delito— lo que define si se alcanza una condena o no.

Desde ya, existen pruebas especiales o investigaciones o inferencias extraordinarias, etc., que se presentan en algunos casos excepcionales. No pretendo aquí negar la singularidad que puede tener un caso, tampoco el alto profesionalismo o las especiales destrezas de un investigador, fiscal, defensor o juez. Pero así como existen formas estereotipadas de cometer ciertos delitos, existen formas regulares de investigarlos y formas recurrentes de valorar los resultados de esas investigaciones. Y esto absorbe la inmensa mayoría de los casos judiciales¹¹.

Por otra parte, algo similar sucede con los inocentes que resultan implicados en casos penales, pues también a su respecto suelen repetirse las circunstancias que los terminan por involucrar en un caso.

¹⁰ Véase «Informe de análisis: Base de datos sobre secuestros extorsivos en Capital Federal (período 2001-2008)» disponible en <https://www.mpf.gob.ar/protex/files/2016/03/informe-estadistico-secuestros-extorsivos.pdf>.

¹¹ Los protocolos de investigación y guías de buenas prácticas que proliferan cada vez más en las fuerzas de seguridad y organismos de acusación para una multiplicidad de delitos y medios de prueba avalan esta conclusión. V. gr., en Argentina, la Unidad Fiscal Especializada en Violencia contra las Mujeres de la Procuración General de la Nación tiene publicados en su web, exclusivamente respecto de muertes violentas de mujeres, una docena de protocolos de investigación (regional, nacional y provinciales). Ver <https://www.mpf.gob.ar/ufem/protocolos-y-guias/>.

Veamos ahora un ejemplo con estupefacientes¹², con dos de las modalidades típicas de inicio de estos casos¹³:

a) El primero es la realización de una denuncia anónima, seguida de tareas de investigación de las fuerzas de seguridad que observan y documentan el circuito de comercio y aprovisionamiento de estupefacientes en simultáneo con intervenciones telefónicas en las que se va conociendo cómo se realizan cotidianamente las operaciones de comercio. Con algo de suerte, también permiten averiguar cómo, de quién y de dónde se proveen los «dealers» de la droga para vender. Cuando ello está suficientemente documentado se allanan los domicilios y requisan los vehículos de los miembros de la organización. El cuadro debe completarse con el hallazgo de droga en inmuebles o vehículos de alguno de los acusados porque tanto las observaciones de las tareas de investigación, como el lenguaje críptico de las conversaciones telefónicas resultan, sin ese hallazgo, evidencia equívoca. Con el hallazgo de droga en poder de los acusados, en cambio, se argumenta que los eventos de intercambio observados desde la distancia y el sentido de las referencias extrañas de las conversaciones telefónicas eran, efectivamente, vinculadas con estupefacientes.

b) El otro mecanismo típico es el hallazgo de droga en tránsito a través de los medios de control del estado, sea en frontera o en controles del interior del territorio por operativos rutinarios de control en rutas. En estos casos, la propia droga secuestrada, el acta del secuestro y las declaraciones de los funcionarios que intervinieron, constituye el núcleo habitual de evidencia.

Por otra parte, el modo en el que inocentes pueden resultar vinculados a estas investigaciones suele provenir, para los casos del tipo –A–, de su presencia inesperada en un inmueble donde se encuentra droga (sobre todo si es mucha) y en los casos de tipo –B– de su presencia en el medio de transporte en el que se encuentra la droga (ej. acompañante del auto que la transporta droga), pues a veces están involucrados en el hecho y a veces lo desconocen.

Asimismo, como evidencia atípica o singular en esta clase de delitos puede surgir algún acusado que desee acogerse al régimen del arrepentido, o encontrarse documentos digitales o físicos con información, o un vendaval de evidencia de descargo por medio de declaraciones testimoniales de amigos y familiares de alguno de los acusados (generalmente el que fue hallado de modo inesperado) que sostienen su inocencia y explican razones de su presencia desafortunada.

¹² No puedo fundar con estudios empíricos los ejemplos que propongo a continuación. No obstante, pudiendo estar equivocados, basta con que sean aceptados como ilustración de situaciones posibles y más o menos plausibles, pues mi interés se limita a que me permitan explicar algunos conceptos relevantes luego.

¹³ Dejaré de lado casos sofisticados, que pueden involucrar estructuras organizadas, circuitos de contrabando desde el extranjero, distribución mayorista o exportación, lavado de dinero y connivencia con autoridades.

Los ejemplos pueden seguir, pero por el momento basta hasta aquí para tomar de este ejemplo y el de «Opi y Gurito» las variables que creo centrales para definir la forma y ubicación de las curvas de culpables e inocentes para un delito: llamaré *conjunto probatorio de referencia* de un delito a ese caudal de prueba generalmente presente en los «casos tipo» de un delito, *prueba singular* a la evidencia distinta de la usualmente presente en tales casos y *contexto de desdicha* a las situaciones desafortunadas en las que personas inocentes pueden verse enfrentadas a prueba en su contra en una causa y las convierte en imputados por el hecho¹⁴.

Dos modalidades genéricas importantes de *contextos de desdicha* son los casos amparados por causas de justificación o de inculpabilidad. Las personas que se encuentran en esas situaciones regularmente tendrán que enfrentar un proceso penal y será allí, luego de producida bastante prueba (incluyendo la que demuestra el hecho ilícito y la autoría del acusado) donde los extremos de la justificación o inculpabilidad se incorporen al caso y terminará por demostrarse la falta de responsabilidad.

Pues bien, entiendo que, analizando un delito en particular, la ubicación y forma de las curvas en el gráfico de culpables e inocentes para ese delito dependen centralmente:

a) La de *culpables*, del conjunto probatorio de referencia del delito según su modalidad tipo de comisión.

b) La de inocentes, de los contextos de desdicha para ese delito según su modalidad tipo de comisión.

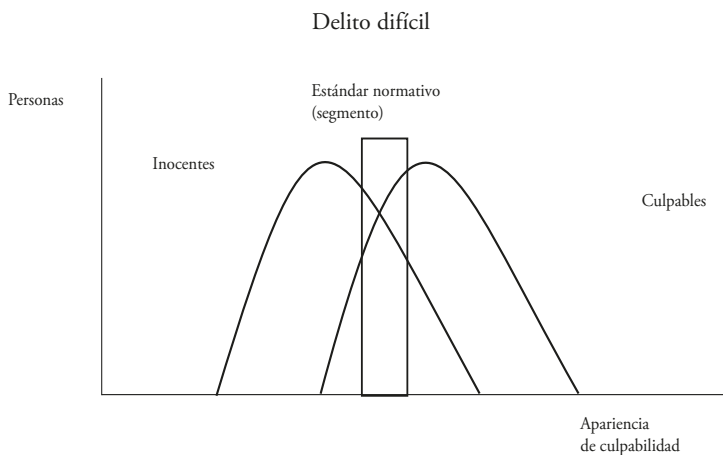
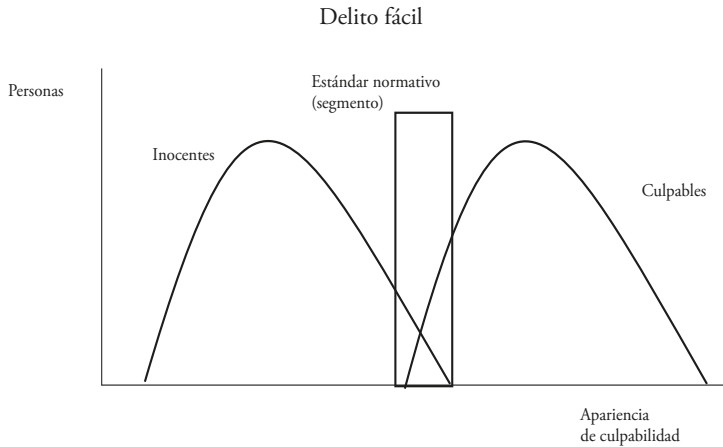
Como es fácil de advertir, para cada delito los conjuntos probatorios de referencia y los contextos de desdicha pueden ser muy diferentes.

Habrán *delitos fáciles* donde los conjuntos probatorios de referencia sean sólidos y los contextos de desdicha escasos (algo así podría suceder con los delitos tributarios, ambientales y con muchos de los originados en lo que Kaplow denomina auditing como mecanismo de detección (KAPLOW, 2012: 830) por ser medios sofisticados y profesionalizados de detección del ilícito) y *delitos difíciles* donde los conjuntos probatorios de referencia sean o bien escasos o bien muy semejantes con las situaciones que enfrentan quienes están en contextos de desdicha (algo así podría suceder con ciertos tipos de robos, abusos sexuales, secuestros —y en general los que dependen de reconocimientos oculares—, los casos de corrupción, etc.).

Creo que, gráficamente, y dado que preliminarmente parece claro que no hay inconvenientes en definir las curvas de inocentes y culpables tomando como referencia un único delito, lo dicho hasta aquí podría reflejarse del siguiente modo¹⁵:

¹⁴ He simplificado bastante las cosas por los fines de este trabajo. El contexto de desdicha puede provenir muchas veces de indeterminación en los tipos penales o de otras razones, como ciertos sesgos de selectividad del sistema, pero no son esenciales para lo quiero mostrar aquí.

¹⁵ En estos gráficos y en los siguientes no ensancharé las curvas solo por una cuestión de claridad de los gráficos, dado que hago referencia a cuestiones que no se vinculan directamente con ese tema. Pero



Como puede observarse, en los delitos fáciles el área de superposición entre la curva de inocentes y la de culpables disminuye, las curvas se alejan y es entonces menos probable que dado un estándar de prueba haya condena de inocentes o absolución de culpables (la ubicación del estándar en estos gráficos es casual, podría estar más a la derecha o más a la izquierda). De hecho, para un delito dado, las curvas podrían incluso no tener área de intersección y si el estándar ha sido debidamente fijado, no existir ninguna condena de inocentes. Es más, si no existen en la práctica contextos de desdicha, la curva de inocentes misma podría no existir para ciertos delitos.

la imprecisión del resultado probatorio y sus efectos estarán siempre presentes.

En los delitos difíciles, en cambio, el área de superposición de las curvas de inocentes y culpables es mucho mayor y, a paridad de estándar y de calidad epistémica del sistema con los delitos fáciles, la tasa de errores será mayor.

Esto obedece, como indiqué, a *conjuntos probatorios de referencia* complejos y la presencia, al mismo tiempo, de abundantes *contextos de desdicha*. Veámoslo con más supuestos.

Los abusos sexuales, por ejemplo, son delitos cuyo conjunto probatorio de referencia tiene como elemento central la declaración de la víctima (RAMÍREZ ORTIZ, 2020).

Existen al menos dos «casos tipo» de abuso: contra allegados de la víctima y contra desconocidos de la víctima. Estimo que hay diferencias relevantes entre ambos «casos tipo». Por un lado, los abusos cuando el autor es un extraño: aquí generalmente no habrá mayores cuestionamientos sobre el hecho y la debilidad de estos casos radica en que muchas veces dependen con exclusividad de prueba de reconocimiento ocular, que presenta tasas de error importantes (DUCE: 2013 y 2017, DIGES y PÉREZ MATA: 2014), lo que colocará la curva de inocentes posiblemente implicados en un punto del eje de «apariencia de culpabilidad» muy cercana y con mucha intersección con la de los verdaderos culpables. En los casos de abuso por allegados, muchas veces el hecho no puede ser probado por otros medios y es difícil desmentirlo si se es inocente, por lo que nuevamente, la consecuencia es que las curvas de inocentes y culpables podrían tener una importante área de superposición¹⁶.

En corrupción también hay problemas tanto de *conjunto probatorio de referencia* como de *contextos de desdicha*. Si el caso tipo es el estallido de un escándalo periodístico o la denuncia de un rival político, muchas veces es muy difícil obtener prueba dirimente adicional al análisis de actuaciones administrativas con serias irregularidades formales (de contrataciones, concesiones de beneficios, etc.). Al mismo tiempo, las denuncias por corrupción pueden formar parte de las estrategias de disputa política y los expedientes administrativos con irregularidades son habituales también cuando no se comete un delito de corrupción, lo que genera contextos de desdicha significativos. Además, por las características de los hechos, los testigos dirimientes suelen ser muy escasos (casi todos los testigos son funcionarios y en cuanto declaran haber advertido irregularidades se ven expuestos a pasar de testigos a imputados).

De acuerdo con lo dicho hasta aquí, la forma y ubicación de las curvas sí influye muy seriamente en la cantidad y distribución de errores del sistema: las curvas de *delitos fáciles* tendrán escasa (pudiendo llegar hasta nula) intersección y consecuentemente habrá pocos errores en el sistema si se fija el estándar en los grados de exigencia adecuados y los *delitos difíciles*, en cambio, tendrán una importante cantidad de erro-

¹⁶ En este último supuesto, sin embargo, seguramente la curva de inocentes sea cuantitativamente muy pequeña, lo que también incide fuertemente en la razón proporcional de errores cuando se modifica el estándar.

res más allá de dónde se coloque el estándar, dada la importante área de intersección de las curvas.

Por cierto, en línea con esto, vale aclarar que definir las curvas para todo un sistema penal, tiene un propósito didáctico a los fines de ilustrar cómo funciona el estándar, pero no es empíricamente posible (pues si existe algún modo de obtener «curvas promedio» entre los distintos delitos, ellas no reflejarían la realidad de ningún conjunto real de inocentes y culpables acusados de ningún delito en particular).

Una aclaración también importante, es que esas variables son también en parte definidas por el Estado, pues su actividad de control (sea por detección o por auditing e incluso por las metodologías forenses de investigación que utiliza y aprueba utilizar) es determinante de contextos probatorios de referencia. En otras palabras la ley no solo por vía de fijación del estándar sino también por su regulación sobre los modos de establecer los contextos probatorios de referencia, y con ello las curvas posibles y sus formas, define la distribución y cantidad del error. Y también los textos penales, con el modo en que establecen sus elementos constitutivos, delimitan, según el estado de la técnica investigativa disponible, los conjuntos probatorios de referencia y los contextos de desdicha.

En efecto, probablemente varios de los delitos de peligro abstracto posean la estructura que tienen pensando en el logro de contextos probatorios de referencia más fáciles (o económicos) de lograr antes que en el propósito de las conductas que estrictamente se desea sancionar.

2. LA MALA CALIDAD EPISTÉMICA DEL SISTEMA

Pretendo ahora poner a prueba algunos supuestos sobre cómo afecta la calidad epistémica del sistema a la cantidad y distribución de los errores teniendo en cuenta lo desarrollado hasta aquí. Esta segunda cuestión, en palabras de Laudan, versa sobre la capacidad para reunir *pruebas robustas* (que según los términos del apartado anterior serían buenos conjuntos probatorios de referencia) y para realizar con ellas *inferencias válidas*.

Vale pensar que un sistema de baja calidad epistémica producirá más errores que uno de alta calidad epistémica. Intuitivamente, para que esto suceda, basta con una mayor área de intersección de las curvas de inocentes y culpables.

Entonces, podría pensarse en primer término que una baja calidad epistémica de un sistema incrementa el área de superposición entre inocentes y culpables desplazando la curva de culpables hacia la izquierda, pues se reúne a su respecto una menor calidad de evidencia y por tanto de apariencia de culpabilidad. También podría desplazar la curva de inocentes hacia la derecha, cuando el esclarecimiento de una situación inicialmente confusa requiere refinar la evidencia o las inferencias y se fracasa en ello (v. gr. causas de justificación e inculpabilidad).

Esto es efectivamente lo que sucedería si esa mala calidad epistémica fuera debido al diseño del sistema, a su regulación normativa. Una mala calidad epistémica en este sentido impedirá el ingreso de evidencia relevante o prohibirá inferencias potencialmente válidas y de esa manera provocará condicionamientos en la apariencia de culpabilidad que puede lograrse (sea para los inocentes o para los culpables), determinando así efectos en la forma y ubicación de las curvas.

Esa clase de debilidad epistémica es la que analiza LAUDAN (2006: 173), poniendo en tela de juicio distintas normas sobre admisibilidad y valoración de la prueba. Esa clase de limitaciones, sin embargo, no me interesan aquí. Dejaré esa cuestión como una constante, presumiendo que el sistema está adecuadamente diseñado.

Por otra parte, lo que para LAUDAN es una constante y presupone correctamente realizado es la práctica de la reunión de prueba y de las inferencias que se efectúan con ella una vez dado el diseño del sistema: «Tales inferencias no reflejan errores imputables al jurado que, supongamos, es muy cuidadoso y racional en sus valoraciones» (énfasis añadido). De este modo, el proceso ejecutivo de la reunión y valoración de la prueba, la calidad de epistémica de los operadores reales en los casos concretos, resulta así ajeno a su interés. Justamente al contrario, ello es lo que será aquí objeto de análisis.

Vale aclarar que estos puntos de partida implican que mientras para Laudan la forma y ubicación de las curvas son variables en juego, para mí, en este apartado, esa forma y ubicación estará dada, pues presupongo el diseño del sistema como una constante.

Por otra parte, Laudan presupone a los operadores (el jurado) como una constante. Es decir, presupone que las decisiones de las personas concretas son exactamente las que resultan de una adecuada reunión y análisis de la prueba en los niveles que permite el diseño del sistema. Así, según ese presupuesto, los operadores reales estarán reuniendo efectivamente la prueba que el sistema permite reunir y valorándola del mejor modo que el sistema permite valorarla, por lo que la apariencia de culpabilidad en concreto y en la práctica, con este presupuesto, será la misma (la óptima que permite el sistema en abstracto).

Sin embargo, como esta será aquí la variable (no la constante), las decisiones concretas de los operadores en la práctica sí pueden apartarse de la apariencia de culpabilidad óptima que resulta de una adecuada reunión y análisis de la prueba (reflejada en las curvas). En otras palabras, no vamos a dar por presupuesto que el operador es *muy cuidadoso y racional en sus valoraciones* como hace Laudan y veremos qué ocurre cuando ello sucede.

Nuestro problema bajo estudio será que al menos ocasionalmente el sistema operativamente falla en la práctica para reunir la evidencia disponible y/o en el momento de su valoración. Desde esta perspectiva, las curvas de culpables e inocentes óptima que el sistema permite obtener (junto a las otras variables que la determinan y que hemos visto en el apartado anterior) no se desplazan en el eje de la evidencia por el hecho de que investigador o el juez sea incapaz de realizar correctamente su tarea.

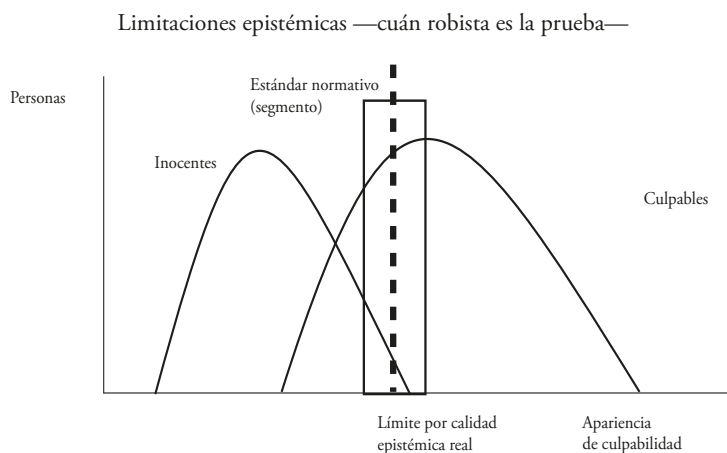
Corresponde observar entonces desde esta perspectiva práctica, cómo impacta la mala calidad en función de sus dos dimensiones («fuerza de la prueba» e «inferencias válidas»).

Estimo que ese impacto se produce, por un lado, afectando el nivel de apariencia de culpabilidad que se puede lograr en los hechos en relación con el óptimo que se podría alcanzar dada la estructura del sistema y, por otro, creando representaciones falsas de las curvas de inocentes y culpables y/o del propio estándar. Veremos cada uno de estos efectos

2.1. Incapacidad para reunir pruebas robustas

Para el amplio (seguramente mayoritario) conjunto de casos en los que la incorporación de más prueba guarda correlación con una mayor apariencia de culpabilidad, los déficits de reunión de prueba limitarán la apariencia de culpabilidad que se puede alcanzar; y con ello el punto (el segmento, mejor dicho) donde el estándar de prueba puede estar ubicado en la práctica. Es que, si el sistema es de mala calidad, habrá evidencia que será incapaz de reunir y posteriormente valorar. Por tanto, los valores máximos conseguibles en el eje x para esta clase de casos no estarán dados por el máximo disponible sino por las limitaciones propias que acarrea no poder obtener cierta prueba.

Gráficamente ello puede reflejarse del siguiente modo:



La línea punteada expresa un límite a la apariencia de culpabilidad que puede lograrse en virtud de la evidencia que puede reunirse. Por limitaciones en la calidad epistémica del sistema, no es posible que la corroboración de una hipótesis vaya más allá del límite indicado.

Como puede observarse en el gráfico, si el estándar de prueba es colocado más allá de su limitación epistémica, se obtendrán un 100% de absoluciones. El dilema que esto plantea es que para condenar culpables debe bajarse el estándar hasta el máximo que permite la limitación epistémica y esta no necesariamente sería la opción socialmente deseable si pudiera elegirse otro estándar más exigente

He colocado el segmento estándar normativo de ese sistema de modo que ocupe valores a ambos lados de la limitación epistémica por cantidad de evidencia pero, en realidad, podría darse un supuesto en el que la prueba que se reúne nunca satisfaga el estándar y podrían haber supuestos inocuos, si las limitaciones para reunir prueba se encuentran a la derecha de donde culmina el segmento-estándar. Salvo en este último supuesto, un estándar limitado por la línea verde que representa la incapacidad para obtener prueba que incremente la apariencia de culpabilidad, será en los hechos un sistema que conducirá por regla general (dependiendo de la forma y ubicación de las curvas) a condenar más inocentes y más culpables, pues habrá que conformarse con un estándar ubicado más a la izquierda.

Es interesante observar también otro conjunto de casos. Aquellos dentro de la curva de inocencia para los que la demostración de tal inocencia requiere de un plus tras la acreditación de que el acusado ha cometido el hecho típico. A veces, sin detenernos demasiado, solemos asumir que más evidencia equivale a más apariencia de culpabilidad en todos los casos. Sin embargo, en múltiples ocasiones ello no es así. Por ejemplo, usualmente en los supuestos de causas de justificación o de inculpabilidad (que ya desde el punto de vista analítico, en la teoría del delito, se analizan con posterioridad a la verificación de que ha existido una acción típica por parte del autor), la evidencia que respalda la falta de responsabilidad penal también suele presentarse como una instancia ulterior de indagación y de prueba, cuando ya se ha reunido la evidencia que respalda la comisión de un hecho *prima facie* punible.

En esos casos, entonces, la prueba de la causa de justificación o de inculpabilidad, suele llegar al proceso después de haber reunido la evidencia sobre la comisión del hecho y la autoría del acusado. Es decir, secuencialmente, se probaría primero que la persona *es* autora y luego, al robustecer más la prueba, se terminaría por acreditar que la persona es en realidad inocente (v. gr. prueba de los requisitos de la legítima defensa, pericias psiquiátricas, etc.).

Estos son ejemplos sobre cómo la debilidad epistémica en reunir *más y mejor* evidencia puede implicar también mayor condena de inocentes. Lo propio puede ocurrir con la prueba de descargo que traen los acusados luego de su ingreso al proceso. Aquí también sumamos más evidencia cuyo efecto, en general, es disminuir la apariencia de culpabilidad.

2.2. Mala calidad de las inferencias

La segunda dimensión de la mala calidad epistémica desde una perspectiva práctica que deseo ejemplificar es la de realizar malas inferencias a partir de la prueba reunida.

En este aspecto clave es necesario recordar, como se señaló al inicio, que un argumento probatorio está compuesto por los siguientes elementos: un dato -hecho-, la conclusión que se pretende extraer de ese dato, una inferencia que los conecta y una generalización que le da fundamento a la inferencia (ANDERSON, SCHUM Y TWINING, 1991 y GONZÁLEZ LAGIER, 2003).

Ciertamente, este problema es más importante que el anterior. Lo que se le «hace decir» a la prueba (lo que pretendemos inferir), se basa en generalizaciones que no están allí para ser tomadas por los operadores (o sí están pero ni los mejores sistemas logran aprehenderlas con certidumbre de modo masivo). Esas generalizaciones, en base a la mayoría de la prueba que se suele reunir en las investigaciones penales, las formulamos, discutimos, contradecimos y reformulamos los propios operadores judiciales.

Tomando un ejemplo de ANDERSON, SCHUM y TWINING (1991) ¹⁷ supongamos que en un caso se discute una inferencia sobre la influencia —instigación— que una persona X ejerció sobre su pareja Y para que cometiera un homicidio partiendo del dato (prueba) de que X era considerablemente mayor que Y.

La máxima de experiencia o de sentido común de que las personas mayores en la pareja pueden influir sobre los menores (que de hecho podría hasta ser llanamente falsa y las cosas ser al revés o depender totalmente de otras variables) puede ser formulada con distintos grados de seguridad y de generalización, por ej.: a) las personas mayores en la pareja suelen influir un poco o b) las personas mayores regularmente influyen mucho, o c) las parejas mayores siempre influyen totalmente en las decisiones del que es menor.

Claramente, si la tercera generalización es la correcta, el fiscal gana su caso solo demostrando la edad de cada involucrado. Pero esta es muy probablemente falsa o, mejor dicho, directamente irracional. La primera inferencia en cambio, aunque prueba muy poco, es una generalización más probablemente correcta.

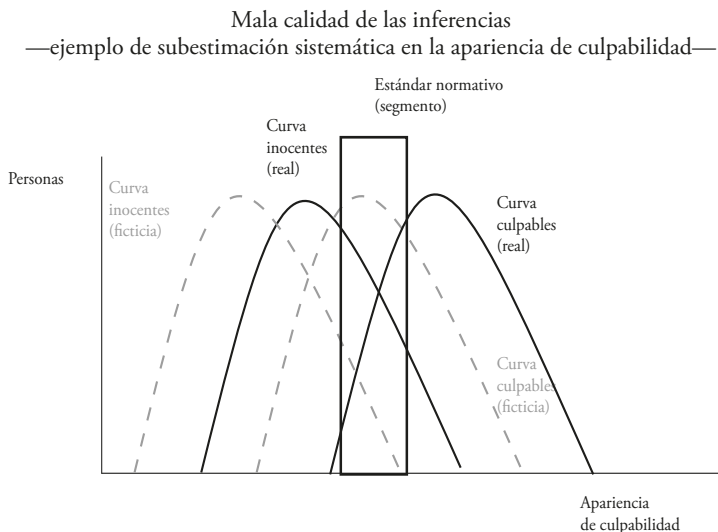
Esa clase de generalizaciones (no científicas, ni tan arraigadas en la experiencia general), son muy abundantes dada la limitada evidencia que suele haber en las causas y juegan un papel crucial en la suerte de los casos. Es en la formulación y discusión de estos fundamentos para nuestras inferencias donde existe más espacio para la duda y donde muchas de las habilidades habitualmente asociadas a los abogados tienen lugar (retórica, empatía, etc.).

¹⁷ El ejemplo está basado en el caso Thomson vs Bywaters que se desarrolla a lo largo de todo el libro, pero ver especialmente aquí, cap. 10, pp. 321 y ss.

Bien, pues, ¿cómo una mala calidad en la formulación y aceptación de estas inferencias y generalizaciones impacta en el gráfico? Creando estimaciones de apariencia de culpabilidad de inocentes y culpables ficticias; y generando por consiguiente puntos de valoración en un caso concreto que se encuentran totalmente fuera de la curva correspondiente (apariencias de culpabilidad directamente arbitrarias). Así ese caso será situado en un valor de apariencia de culpabilidad que estará totalmente fuera de la curva correspondiente. Y si este problema fuera sistemático y tendencioso (v gr. un fuero que resuelve constantemente de modo erróneo acorde a designios de un gobierno dictatorial) se podría terminar creando virtualmente toda una nueva curva ficticia fuera de la racional.

En definitiva, si disminuye la calidad epistémica de nuestras inferencias en un sistema en relación con la valoración de la prueba, la consecuencia es que el sistema estará creando estimaciones de culpabilidad ficticias, distintas de las que realmente derivarían de una valoración racional de la prueba. Cuando inferimos erróneamente a partir de la prueba más cosas, o menos cosas, de las que deberíamos inferir, o cuando directamente inferimos un absurdo, la apariencia de culpabilidad resultante para quien está valorando, se ubicará en un punto distinto del que debería. Está, en definitiva, elaborando una nueva apariencia de culpabilidad fuera de los márgenes de lo racionalmente aceptable.

Veamos gráficamente un ejemplo —de los infinitos que pueden darse en la práctica—:



Aquí se representa un sistema con mala calidad de las inferencias de modo sistemático, en las que se da a los conjuntos probatorios menos fuerza probatoria de

la que poseen. También podría ser al revés y darles más fuerza de la que en realidad poseen, lo que colocaría los puntos o curvas ficticias hacia la derecha de las reales. Los ejemplos son infinitos, pues quien estima la apariencia de culpabilidad, una vez que se aparta de los márgenes de la valoración racional, podría colocarse en cualquier punto del eje horizontal. Nótese que este fenómeno es diferente del ensanchamiento (segmento) que se produce en las curvas que representan correctamente el margen de resultado probatorio alcanzado con un conjunto de prueba (ese segmento está contenido en el grosor de la curva «real»).

En efecto, la mala calidad consiste en realizar inferencias que se apartan del segmento de conclusiones logradas a partir de una valoración racional de la prueba. (V. gr. en el ejemplo de la diferencia de edad en la pareja y su capacidad de influencia, dentro de la curva azul se encontrarían dos o más de las inferencias moderadas del tipo la persona mayor «suele influir un poco» o «suele influir mucho» -al ser varias las plausibles se advierte claramente que se trata de un segmento-. En la curva ficticia verde que se ha representado a la izquierda, sin embargo, tendríamos inferencias absurdas del tipo «la persona mayor jamás influye sobre la menor» (y si hubiera representado la curva ficticia a la derecha, en ella se encontrarían inferencias irracionales del tipo «la persona mayor siempre influye de modo absoluto sobre la menor, también falsa pero de sobrecarga de valor probatorio, en lugar de subestimación).

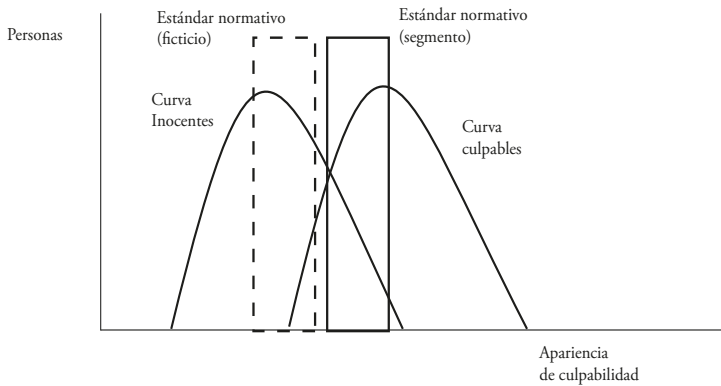
Cabe concluir entonces que la mala calidad de las inferencias en el sistema, en cuanto a su potencial para incidir sobre la distribución de errores y su número total, es muy relevante. Ello se advierte con claridad, por ejemplo, en todo el tramo de la curva de culpables ficticia (verde) que se encuentra a la izquierda del estándar normativo, comparada para cada valor igual del eje Y pero de la curva real (azul). En todos esos casos una buena calidad del sistema -curva azul- arrojaría condena de culpables (o al menos la posibilidad de que se dicte condena), mientras que el sistema de mala calidad -curva verde- arroja absolución (segura) de todos esos culpables. Es decir, todo ese tramo es un problema de mayor cantidad de errores, no de su distribución.

Por otro lado, también la mala calidad epistémica podría acarrear errores de interpretación del estándar normativo, es decir, independientemente de si el resultado de la valoración ha permitido que el juez haga inferencias correctas en cuanto a la apariencia de culpabilidad resultante, podría quizás equivocarse el razonamiento relativo a si ese nivel de apariencia de culpabilidad logrado satisface o no el estándar en su dimensión de pauta.

En los casos de defectos de interpretación de los términos del estándar, los jueces podrán estar resolviendo directamente fuera de él con base en la evidencia disponible. Serán casos de arbitrariedad de las decisiones a la luz de la pauta normativa que establece el estándar. En estos casos, el juez en su interpretación de la formulación del estándar podrá entonces operar con un estándar ficticio, más amplio, y lo que estará creando no será una curva ficticia sino un estándar ficticio

Gráficamente:

Delito modelo. Errónea interpretación del estándar normativo



El gráfico ilustra el efecto de un sistema de mala calidad con errores de subsunción en el estándar normativo. Aquí se ejemplifica un caso donde el error se comete aplicando un estándar menos exigente, pero podría suceder lo contrario y el estándar ficticio se encontraría a la derecha.

III. TERCERA PARTE

1. LA DIMENSIÓN PRÁCTICA DEL ESTÁNDAR Y LOS DOS PROCEDIMIENTOS TRADICIONALES PARA DECIDIR: JUECES PROFESIONALES Y EL JUICIO POR JURADOS LEGOS

Hasta aquí he sostenido que el establecimiento del estándar de prueba en el plano normativo opera como una dimensión lingüística orientadora del nivel de exigencia probatoria que debe cumplirse para justificar válidamente una decisión.

En este terreno FERRER BELTRÁN (2013) ha hecho avances significativos al mostrar lo vagas e indeterminadas que suelen ser las formulaciones de estándares a nivel normativo y ha avanzado en formulaciones mucho más precisas e intersubjetivamente revisables¹⁸.

¹⁸ En cambio, para DEI VECCHI (2020), esta dimensión lingüística solo podría alcanzar una exigencia sobre la índole de las razones que deben usarse para tener por probado un hecho y en todo caso imponer la necesidad de valerse cuanto menos de algunos criterios epistémicos determinados para justificar las decisiones probatorias. Sostiene, sin embargo, que ello no es suficiente para considerar que el estándar de prueba puede cumplir la función de *umbral* —p. 31—.

Ahora bien, he mencionado también en la introducción que luego esa dimensión de «pauta» del estándar de prueba que se expresa por fórmulas lingüísticas tiene una dimensión práctica u operativa que presenta particularidades adicionales y he ejemplificado su funcionamiento con los conjuntos probatorios de referencia y los contextos de desdicha de distintos delitos (secuestro extorsivo, estupefacientes, abuso sexual, corrupción), pero sin detenerme hasta ahora en cómo se fija esta dimensión del estándar.

Creo que esos ejemplos muestran que, en el plano práctico, la aceptación de que algo se encuentra demostrado o probado (siempre *prima facie* y condicionado por el nivel de conocimientos disponibles en un momento y lugar determinados), se vincula con procedimientos o modos de llegar a esas conclusiones que son considerados aptos para ese fin, operando el estándar normativo como una pauta del nivel de rigor, pero no agotando con ella los pasos necesarios para completar el significado del estándar.

Dentro de esta misma dimensión, nuevas observaciones, nuevos análisis, etc. suelen modificar el estado de situación respecto de las exigencias necesarias para aceptar como probada una conclusión, que puede ser descartada, reemplazada por otra, etc.

Presenciamos este fenómeno en el ámbito judicial a menudo. De acuerdo con la idea de los contextos probatorios de referencia y cómo se trabaja con ellos en los sistemas judiciales, la presencia de ciertos elementos de prueba con ciertos grados de fidelidad y autenticidad, la ausencia de otros elementos que los contrarrestarían y la realización de ciertas inferencias a partir de ellos, habilitan ciertas conclusiones (p.ej. culpabilidad) para determinados tipos de hechos.

Pero la estabilidad de ese punto de estándar de exigencia en la práctica suele tener modificaciones en base a nueva información relevante. A veces, gracias a nuevas tecnologías se advierte, con la reiteración de casos, que probablemente se estén filtrando más condenas de inocentes que las aceptables, etc. Y ello sucede no solo sin que sea necesario modificar la formulación normativa del estándar de prueba, sino justamente en su nombre. Es decir, en la dimensión práctica, a partir de los procedimientos aplicables, se advierte que las exigencias que se usaban para aceptar como probado un hecho a la luz del estándar normativo, no eran correctas y corresponde utilizar otras, o extraer conclusiones distintas.

Creo que los casos de abusos sexuales y de violencia doméstica son un buen ejemplo de lo que pretendo expresar aquí. Entiendo que en estos casos se ha producido en los últimos años una revisión sobre las generalizaciones que fundamentan las inferencias sobre la veracidad de las declaraciones de las víctimas de estos delitos. En este sentido, era habitual en muchos sistemas de justicia que, en virtud de ciertas perspectivas de valoración social, a partir del dato de que la declaración proviniera de una mujer se hacía una inferencia sobre su baja credibilidad, basada en una generalización que les restaba en abstracto valor a sus declaraciones. Esas generalizaciones eran tributarias de una concepción machista a partir de la cual era habitual que las

mujeres pudieran estar exagerando o que existían *provocaciones* que podrían justificar reacciones violentas de los hombres, etc.

Como ha señalado la Corte Interamericana de Derechos Humanos en el caso V.R.P., V.P.C. y otros vs. Nicaragua (sentencia de 8 de marzo de 2018, excepciones preliminares, fondo, reparaciones y costas):

A ello se suman los prejuicios e ideas preconcebidas y estereotipadas propias del sistema patriarcal que existen en el imaginario social en torno a la violencia sexual. Los jurados son susceptibles de trasladar al procedimiento tales prejuicios e ideas y ser influenciados por ellos al momento de valorar la credibilidad de la víctima y la culpabilidad del acusado, condicionando de modo especial a quienes no poseen una capacitación especial en este tipo de delitos

En los últimos tiempos, se ha tomado conciencia de lo errado de esas generalizaciones, no solo desde el punto de vista ético o valorativo, sino desde el punto de vista fáctico y de la valoración de la prueba. Es decir, de la falsedad de esas generalizaciones.

Desde ya, esto no implica afirmar que en todo caso concreto, donde haya una contradicción, será la víctima la que resulta veraz y el imputado el que resulta falaz. Esto debe ser una materia de debate y prueba en cada caso concreto, solo que las generalizaciones sobre la base de las cuales elaboramos nuestra reflexión probatoria deben ser adecuadas a la realidad y no sesgadas por prejuicios infundados que derivan de un sistema moral cuyas fallas están en la actualidad bastante claras.

Asimismo, la experiencia secuencial del sistema de justicia testeará la fortaleza del nuevo esquema de exigencias. Sabemos que es difícil que explícitamente se documenten y reconozcan los errores del propio sistema porque no hay mecanismos externos sistemáticos para escrutar si hubo errores. El sistema además, es probablemente más permeable a corregir razonamientos para el futuro que a reconocer el error dentro del mismo caso.

Pero todo esto sucede sin modificar el estándar como pauta, sino por respeto a él.

Y aquí deseo entrar al núcleo de este apartado: Cómo lo hace cada sistema.

Parece claro que en el sistema de jueces profesionales y múltiples instancias el procedimiento funciona por reflexión, fundamentación, repetición y revisión. El mismo juez enfrentará una y otra vez situaciones similares sobre las que podrá analizar y reflexionar muchas veces y allí irá refinando sus razonamientos, los elementos de prueba que busca, cómo los produce, qué exige de ellos en los detalles, etc. A su vez, su operatoria será revisada, corregida y mejorada (o empeorada) por instancias de revisión. Las instancias más altas de los conjuntos de casos de distintos territorios además cristalizarán algunos principios más fuertes y consolidados del razonamiento y las conclusiones que pueden obtenerse.

En los períodos en que el contexto probatorio de referencia está estable, y sin perjuicio de las singularidades de los casos, el control de satisfacción del estándar se limitará a verificar que los componentes del conjunto probatorio de referencia estén

allí o no y afirmarán si el estándar está satisfecho. Pero realizarán esto esencialmente contra la comparación de otros casos, su repetición y revisión y no por oposición a una formulación lingüística del estándar.

Es claro que en este esquema la motivación, como expresión de estos recorridos y guía de puntos para los casos venideros es esencial, es de hecho el terreno de disputa y discusión para la fijación del estándar en su dimensión práctica.

Así lo expresa FERRER BELTRÁN (2020. En el mismo sentido ya desde 2007: 65). Siendo el medio por el que se da cuenta de la justificación probatoria de una aserción, además de ser indispensable para poder articular las vías recursivas¹⁹, es el único modo de saber si la decisión ha satisfecho un mecanismo de valoración racional.

En caso de que el juez o Tribunal no la exprese, no solo los demás no podrán verificar que verdaderamente existió y tampoco podrán recurrir adecuadamente, sino que, dado que contiene esencialmente exigencias procedimentales, abiertas a una aplicación de criterios con mayor o menor destreza, en caso de no ser expresada es imposible saber si la decisión está correctamente justificada²⁰.

En un sistema de jueces profesionales por medio de la motivación se ponen de manifiesto los mecanismos y razonamientos por los que se considera corroborada una hipótesis. Es una actividad que, en un estado dado del conocimiento general, permite el control intersubjetivo.²¹ Se trata además de una exigencia avalada por la jurisprudencia de los tribunales internacionales de derechos humanos²².

Aunque podrían establecerse mecanismos procedimentales mixtos y algo así ocurre con los Tribunales de juicio oral de muchos sistemas (que exigen motivación más concurrencia de cierta cantidad de miembros -por lo general una mayoría en un cuerpo de tres o cinco-) lo cierto es que en este esquema de valoración el estándar consiste en la aplicación de un método para llegar a las conclusiones y justificarlas, que será más o menos exigente y que queda expresado y puede ser escrutado por otros luego, que podrán eventualmente advertir falencias del procedimiento seguido por el juzgador anterior.

En definitiva, es una garantía del procedimiento en un sistema de jueces profesionales y valoración racional de la prueba para satisfacer un estándar que esté sujeta a un escrutinio posterior de terceros, que aplicando el mismo método pueden encontrar falencias o confirmar el acierto de la decisión en un estado dado del conocimiento.

¹⁹ Sobre la función endo-procesal de la motivación para permitir las vías recursivas, FERRAJOLI, 1990: 623, SARRABAYROUSE, 2007: 286 y TARUFFO: 2009: 516).

²⁰ Sobre la exigencia de expresar la motivación, GASCÓN ABELLÁN, 1999: 201.

²¹ Exigen racionalidad para la motivación, NIETO, 2000: 157 y COLOMER, 2003: 32.

²² STEDH Ruiz Torrija v. España, de 9 de diciembre de 1994 y CorteIDH de 5 de agosto de 2008, caso Apitz Barbera y otros (Corte Primera de lo Contencioso Administrativo) vs. Venezuela, Serie C, nº 182.

Ahora bien, aunque en las etapas iniciales e intermedias del procedimiento los sistemas de juicio por jurado pueden tener este tipo de metodología o procedimiento, en la instancia definitiva del veredicto por jurado que delibera de modo secreto, es evidente, según creo, que el sistema no opera del mismo modo ni confía la decisión a la clase de procedimiento desarrollada aquí para el sistema de jueces profesionales, en absoluto.

Recientemente FERRER BELTRÁN (2020) ha analizado el fallo de la Corte Interamericana citado. Allí la Corte pretendió fundamentar la tesis de que el juicio por jurados satisface, desde una consideración global, la exigencia de motivación, aunque esa motivación no sea expresada.

FERRER ha mostrado el error de esa tesis de la Corte. La forma en la que la Corte establece su presunción sobre el mecanismo de toma de decisiones de los jurados no es sostenible porque, aunque acierte de hecho, el único recurso que tiene para establecerla es presuponerla. Y sobre todo porque la Corte asocia al sistema de juicio por jurados con una valoración de la prueba bajo la íntima convicción, que al apelar a estados subjetivos no exige ningún método específico para el arribo a la decisión. Es evidente que en ese sistema el núcleo va por otro carril.

El inconveniente principal de la tesis de la motivación en los jurados, creo, es que no se advierte que la dimensión práctica del estándar en este sistema es cualitativamente diferente a la de los jueces profesionales. El aspecto central del estándar en su dimensión práctica en los sistemas de juicio por jurados es esencialmente *cuantitativo*, no uno de justificación de la corroboración de la hipótesis luego sometido a repetición y revisión como el de jueces profesionales.

En el juicio por jurados, en efecto, es habitual que las normas dispongan restricciones a las veces o la frecuencia con que una persona puede ser jurado, así que los individuos que lo componen mal podrán refinar su técnica por repetición y revisión.

A su vez, en este sistema, en principio, se renuncia a la posibilidad de escrutinio posterior, no se apuesta a que la conclusión sea obtenida por la aplicación de un método específico, más pautado, impugnabile por detección de errores en la ejecución del razonamiento de sus integrantes.

Sí es necesario presumir que los miembros del jurado actuarán de buena fe y que se guiarán por la prueba que se produce en el juicio (para eso se lleva adelante el juicio antes de que el jurado pase a deliberar, es decir, esta no es una presunción aventurada como la de la Corte Interamericana, simplemente es consecuencia de cómo está concebido el juicio, pues si no se espera que el jurado base su decisión en lo que sucede en el juicio, no tendría sentido realizarlo).

Además de ello, también existe el estándar normativo en su dimensión de pauta o guía, que suele ser «más allá de toda duda razonable» (BARD) cuyo oscuro significado ha sido muy claramente expuesto por LAUDAN (2006, cap. II).

Pero en cuanto al procedimiento de valoración, en lugar de pautas más o menos rigurosas para llegar a conclusiones, la garantía del sistema está dada por el hecho de que los jurados estén compuestos por numerosos integrantes.

Es también un estándar procedimental, al igual que el otro, pero de otras características: este es cuantitativo.

En este punto considero que el BARD es solo una exhortación, y así lo avala lo señalado por Ferrer Beltrán, que al igual que Laudan muestra la indefinición del concepto de BARD, incluso en los fallos de la Corte Suprema de Estados Unidos, que terminan por dar referencias por las que ese estándar no se diferencia de la íntima convicción (FERRER BELTRÁN, 2013: 6).

Quisiera explicar que el aspecto central del estándar es por confirmación cuantitativa con un ejemplo. Si usted ingresa a una sala de juicio injustamente acusado de haber cometido un hecho criminal y en los estrados se encuentran el juez, el fiscal, su defensor y en el atril del jurado hay doce personas y le indican que esas personas, todas ellas, deben decidir que usted es culpable para que usted sea declarado culpable, seguramente se sentirá más algo más esperanzado que si en las mismas circunstancias en el atril hay una sola persona y le indican que basta con que esa persona sola piense que usted es culpable para que usted sea declarado culpable. Ese sujeto único podría ser un inepto, un *odiador*, etc. Pero entre 12 personas, las posibilidades de que al menos uno comprenda su caso como usted desea es ciertamente mayor.

Allí se advierte que la exigencia es cuantitativa. Ahora doce personas deben convencerse de que usted es culpable, no una. Es cierto que hay ciertas circunstancias que generan interdependencia en el error. Por ejemplo, si en este caso el problema es que el reconocimiento ocular de la víctima lo sindicó erróneamente a usted. Aquí se trata de una clase de error que tiende a afectar a todos los integrantes del jurado. Aun así, seguramente usted preferirá que haya doce integrantes y no uno, sus chances de que alguno de los jurados desconfíe de la fiabilidad del reconocimiento es mayor con muchos integrantes que si solo hay uno.

El método de valoración que cada uno emplea sobre la prueba, al no tener que expresarlo ni estar sujeto a control posterior, no reviste ni por aproximación la importancia que tiene el número de integrantes.

Hay una expectativa razonable de buena fe y de que los jurados se basen en las pruebas, como ya se indicó, y también las instrucciones previas y la capacitación serán bienvenidas. Pero la forma en la que decidan estará condicionada seguramente mucho más por las generalizaciones que generaron por su historia de vida, su personalidad, etc. que por tales instrucciones. Que allí se les diga que resuelvan más allá de toda duda razonable si usted es culpable será un aliciente, pero uno que usted seguramente estará dispuesto a negociar a cambio de que haya doce jurados en lugar de uno solo.

El origen histórico del jurado, aunque no avala explícitamente esta idea, permite argumentar a favor de la relevancia que tuvo la dimensión cuantitativa, junto con la

pretensión de imparcialidad de sus miembros. Si su origen en el sistema anglosajón fue la exigencia de que los acusados no serían juzgados sino por un conjunto de iguales; y con ello se logró la incorporación de la cláusula 39 en la Carta Magna de 1215 (CONTI GÓMEZ, TOLEDO, 2012) que estableció ese sistema, parece claro que lo que importaba es que el juzgador de los hechos no estuviera en defensa de los intereses de alguien más allá de lo sucedido y que fueran *numerosos iguales* los que decidieran, pero no hay nada allí que indique que la pretensión de que los jurados decidieran bajo cierto método de reflexión probatoria.

SCHIAVO (2016: 96) repasa los orígenes históricos, y la legislación y jurisprudencia de Estados Unidos sobre el número de doce (12) integrantes para la composición del jurado. Tras señalar que no está resuelta la controversia doctrinaria sobre el origen del emblemático número 12 (probablemente místico, con teorías que lo vinculan con la cantidad de apóstoles, con las tribus de Israel y con el número de oficiales del Rey Salomón-), destaca que su persistencia a lo largo de la historia es explicada también por motivos funcionales (cantidad de personas susceptible de ser reunida sin tantas dificultades y al mismo tiempo suficiente para confiar en sus afirmaciones juramentadas).

Luego realiza un repaso minucioso de las regulaciones de los estados de Estados Unidos, especialmente de las disposiciones que establecen supuestos de juzgamiento por jurados con un número menor a 12 integrantes y de los casos en los que la Corte Suprema abordó la cuestión²³. En el primero de esos casos, Thompson, de 1898, la Corte Suprema de Estados Unidos sostuvo que no podía tener una cantidad inferior a 12 integrantes. En el segundo, Williams, de 1970, revisó su posición y convalidó la regulación estatal de Florida para los casos en los que establecía una integración de 6 miembros. En el último, Balew, de 1978, rechazó la posibilidad de que el jurado pudiera estar integrado por cinco miembros.

Más allá del modo en que se explicitan las razones de las exigencias en cuanto a cantidad de miembros (fomentar la deliberación grupal, evitar intentos de intimidación externa, que se proporcione la posibilidad de que el jurado sea una muestra representativa de la sociedad), tanto las regulaciones legales como los fallos confirman la importancia del factor cuantitativo como garantía de una exigencia mayor para la condena (es decir, la elevación del estándar de prueba por medio del número de integrantes).

En este sentido, todas las regulaciones que establecen distintas composiciones en cuanto a cantidad de integrantes del jurado, siempre permiten un número menor de integrantes para delitos con amenazas de pena menos graves y números mayores para delitos más graves. Lo que es consistente con una mayor exigencia de estándar cuando el hecho y la condena es más grave.

²³ «Thompson v. Utah», 170 US 343, 1898, «Williams v. Florida», 399 US 78, 1970 y «Balew v. Georgia», 435 US 223, 1978.

Pero particularmente, como señala SCHIAVO (2016: 110), en *Balew* la Corte recurre a estudios empíricos que muestran que cuanto menor es el tamaño del grupo mayor es el riesgo de condenar a un inocente mientras que cuanto mayor es el tamaño, lo que aumenta es el riesgo de no condenar a un culpable. En línea con ello, y siendo el primer riesgo más grave que el segundo, la Corte termina fallando en contra de la posibilidad de conformar un jurado con solo cinco integrantes como lo había hecho el estado de Georgia en ese caso.

Creo que desde el punto de vista epistemológico, este estándar puede fundamentarse sobre la misma base que NINO (1989) basó su idea sobre el valor epistémico de la democracia.

De acuerdo con el teorema de Condorcet, si aceptamos de modo general que es más probable que una persona acierte en una elección a que cometa un error, cuantas más personas se inclinen por la misma elección, aumentará la probabilidad de que esa elección sea correcta (salvo los problemas señalados de interdependencia del error), tal como lo señala la Corte Suprema de Estados Unidos en el caso *Balew*.

En efecto, como señala Sarrabayrouse (2020: 4), el propio Condorcet publicó en 1785 su *Ensayo sobre la aplicación del análisis a la probabilidad de las decisiones sometidas a la pluralidad de votos* a raíz de la condena que implicó la tortura, decapitación y hoguera de un hombre por no sacarse el sombrero frente a una procesión religiosa. Allí sostuvo que la seguridad acerca del acierto en las decisiones aumenta con el incremento de las mayorías necesarias para adoptarlas.

Entonces, cuantos más jurados incorporemos, más alto será el estándar de exigencia para declarar a alguien culpable, en la medida que se exige que todos ellos estén de acuerdo en cuanto a la culpabilidad del acusado; o cuanto menos una mayoría calificada.

Los supuestos de mayorías calificada para condenar, si bien en principio disminuyen la exigencia del estándar a paridad de miembros del jurado (el efecto debería estimarse según el total de miembros del jurado), no afectan el sustrato del argumento, sino que reconduce la cuestión a un problema de cálculo del impacto que un voto negativo tiene sobre el nivel de exigencia de acuerdo con el mismo teorema.²⁴

Ahora bien, tratar de transpolar hacia este mecanismo de establecimiento del estándar, cuya deliberación por regla es secreta, una exigencia metodológica (la motivación) propia de otro sistema (el de jueces profesionales) simplemente no encaja, salvo que estemos dispuestos a forzar presunciones como hace la Corte Interamericana.

En definitiva, el juicio por jurados exige un número razonable de miembros imparciales, que actúen de buena fe y que se comprometan a resolver en función de la producción de la prueba que se hace frente a ellos y de la pauta de rigor que se les

²⁴ «Johnson v. Louisiana», 406 US 356, 1972, que convalida los veredictos por mayoría calificada no sostuvo la asociación probabilística que sostengo aquí como lo hizo con el caso *Balew*.

da (más allá de toda duda razonable), junto a un conjunto de instrucciones, pero nada más, pues la exigencia esencial en cuanto a estándar viene dada por la necesidad de que una pluralidad de miembros llegue a la misma conclusión, no porque lo hagan por un mecanismo de justificación conforme a exigencias aceptadas dentro del sistema correspondiente.

Lo que hemos visto entonces hasta aquí es que el procedimiento de jueces profesionales motivando a la luz de los conjuntos probatorios de referencia y repitiendo la tarea y sujetándola a revisión por el superior, refinando y precisando esos conjuntos y las inferencias que hacen con sus elementos es parte constitutiva del estándar junto con su formulación lingüística que le sirve de pauta de escrutinio sobre los niveles de rigor que se está exigiendo de tales conjuntos probatorios.

En otro orden, que esos conjuntos probatorios de referencia (construidos por las autoridades de investigación y acusación) superen el convencimiento simultáneo de 12 personas -más o menos exigentes- es parte constitutiva de la dimensión práctica del estándar de otro sistema, que también se complementa con la formulación lingüística que se les brinda (v gr. BARD). Aquí investigadores y fiscales serán probablemente quienes estén atentos a si sus conjuntos de prueba habituales están satisfaciendo a los jurados o si necesitan ajustes. Quizás aquello que antes era suficiente ya no lo es. Deberán entonces aumentar el nivel de exigencia probatoria (aunque el estándar en su formulación lingüística seguramente no cambiará).

Pero algo es seguro, la dimensión práctica del estándar de prueba, la garantía sobre cómo se llega a satisfacer su dimensión lingüística es cualitativamente distinta en ambos sistemas. Lo que sí es válido para ambos es que esa dimensión práctica es tan constitutiva del concepto de estándar de prueba como lo es su dimensión lingüística, e igual de importante.

El estándar de prueba es entonces una expresión o formulación lingüística -dimensión pauta- asociada con un modo determinado de establecer si se la ha satisfecho -dimensión procedimental-, que permite fundamentar de modo legalmente válido si ciertos hechos están suficientemente acreditados a los fines de tomar distintas decisiones en el proceso.

IV. CUARTA PARTE

1. UNA BREVE REFLEXIÓN FINAL SOBRE LA VAGUEDAD DEL ESTÁNDAR Y LA POLÍTICA CRIMINAL

Evidentemente un estándar -pauta- muy vago genera niveles de incertidumbre jurídica inaceptables, y podría llevar a la arbitrariedad y la irracionalidad. De todos

modos, su definición de un modo estricto (un punto en el eje x) no es posible de lograr y entonces siempre será un segmento. Como señalé antes, Jordi Ferrer ha hecho progresos notables para reducir esta vaguedad, más allá de que eliminarla por completo probablemente no sea posible.

Pero ello no debe desesperar. Tenemos, así lo señala LAUDAN (2006: 133), prácticamente un estándar específico formulado para cada decisión relevante del proceso penal y nadie se alarma por ello. Sin embargo, asumimos que el estándar de condena debe ser el mismo para todos los delitos. ¿Por qué?

¿Es conveniente que sea el mismo para todos los delitos sin importar el modo en el que las curvas de inocentes y culpables se emplazan en el gráfico según el delito específico de que se trata? ¿Y ello para todo tiempo y conjunto de valores sociales? En el plano de la formulación lingüística puede ser, pero en la dimensión práctica no me parece.

Al contrario, creo que es la existencia de un cierto margen en la fijación práctica y concreta del estándar dentro del segmento normativo (es decir, el subsegmento que se establece en la práctica judicial por medio de su dimensión procedimental-), el indispensable espacio para que la sociedad pueda darse una política criminal en la práctica consistente con los objetivos que se ha fijado en abstracto (es decir, diferente según la problemática social distinta que cada delito aborda y en cada momento histórico) y que a su vez sea coherente con otros valores exigidos por la sociedad.

ALLEN (2014: 211), en línea con su posición sobre el carácter empírico de las implicancias del estándar de prueba y partiendo de supuestos que conducen a un conjunto de casos cuantitativamente inferior de los que componen la curva que en este trabajo denominamos de inocentes en el sistema penal (deserving defendants) y con poca intersección con la de culpables (deserving plaintiffs) se pregunta por la movilidad del estándar en función de una configuración de ese tipo (similar en intersección a la que aquí he ejemplificado como delitos fáciles pero con la curva de inocentes bastante más pequeña que la de culpables, en razón de que serían muchos menos los inocentes que son efectivamente sometidos a juicio en comparación con los culpables).

Él señala inmediatamente una cuestión muy relevante, que es que debería atenderse, al modificar un estándar de prueba, a cómo se afecta también la configuración de los conjuntos de inocentes y culpables, pues existe una interdependencia entre ellos. La naturaleza empírica de la cuestión, hace que al modificar el estándar muy probablemente se afectará también los subconjuntos de inocentes y culpables sometidos a proceso en función de los cambios que producirá en los incentivos de las partes involucradas y por tanto también se verá afectada la forma y distribución de las curvas. Por ello, habrá que verificar luego del cambio del estándar cuál es el efecto final en la distribución de los subconjuntos de casos y de los errores esperados del sistema. Sin embargo, Allen no rechaza en ningún momento la idea de la modulación

de los estándares de prueba e incluso afirma que puede pensarse en que ciertas clases de casos son diferentes de otros y merecen un tratamiento especial.

Imaginemos por un lado conductas de contaminación ambiental atomizadas. Se trata de conductas agregadamente muy nocivas pero individualmente percibidas como poco graves. Este delito generalmente está amenazado con penas bajas y quizás tengan las curvas de inocentes y culpables razonablemente separadas. A su vez, la demanda social contra este problema ha crecido mucho en el último tiempo. En estos casos podríamos pensar que un estándar de prueba algo más bajo, siempre dentro del margen de la formulación lingüística que asegura el respecto de las garantías, de modo que se produzca una disuasión general más adecuada dada la gravedad de la infracción agregada aunque ello podría incrementar el riesgo de error por condena de inocentes, pero que de todos modos serán pocos y con sanciones bajas.

Cuando se trata de un delito en el que las curvas de inocentes y culpables son cuantitativamente similares y tienen mucha superposición, y partiendo de la idea de que condenar inocentes es mucho más grave que absolver culpables, correspondería extremar el estándar aunque eso implique más riesgo de absolución de culpables. También los delitos con penas graves ameritan estándares bien rigurosos.

Estas diferencias y la incorporación de otros valores es posible si el estándar en el nivel normativo, que se formula para condenar en general sin distinción de delitos, permite este margen a partir de algún nivel aceptable de vaguedad -amplitud- para desplazarse en el ámbito de la dimensión práctica del estándar.

En efecto, seguramente esto es así en los hechos, si es cierto que, volviendo al ejemplo de los abusos sexuales, la condena en casos de violencia contra la mujer e intrafamiliar, por explotación sexual y trata de personas ha aumentado sin que el estándar en su formulación normativa haya variado y sin que hayan prosperado planteos de arbitrariedad contra la nueva tendencia jurisprudencial (si es que lo demás se mantuvo constante, como creo que es el caso).

2. CONCLUSIÓN

En este trabajo, se ha argumentado en defensa de las siguientes tesis:

1) Que la vaguedad en la formulación normativa del estándar y la imprecisión de los resultados probatorios son insalvables e influyen en el estándar de prueba que en la práctica se puede establecer, quedando márgenes en los cuales los casos estarán válidamente decididos tanto si resultan en condena como si resultan en absolución.

2) Que los sistemas penales operan en la práctica a partir de conjuntos probatorios de referencia y contextos de desdicha y que ellos, junto al diseño normativo del sistema, son configurativos de la forma de las curvas de inocentes y culpables, lo que a su vez es determinante para la cantidad total de errores del sistema en cada delito

y en el menú de opciones distributivas del error que estarán disponibles para fijar el estándar de prueba.

3) Que la mala calidad epistémica del sistema judicial, desde una perspectiva práctica, puede modificar significativamente tanto la cantidad como la distribución de los errores que se producirían en comparación con una adecuada reunión de la prueba y su valoración racional (manteniendo constante el diseño del sistema).

4) Que el estándar de prueba, a la par de su dimensión lingüística, posee una dimensión práctica. Que esa dimensión práctica es cualitativamente diferente en los sistemas de jueces profesionales y en los sistemas de juicio por jurados legos.

5) Que esa dimensión práctica del estándar permite una implementación adecuada de la política criminal.

BIBLIOGRAFÍA

- AGUILERA, E., 2008: «El concepto de estándar de prueba», en *Revista del Posgrado en Derecho de la Unam*, núm 6 (Vol. 4), pp. 91-114.
- ALLEN, R. J., 2014, «Burdens of Proof», en *Law, Probability and Risk*, 13: 195-219.
- ANDERSON, T., SCHUM, D. y TWINING, W., 1991: *Analysis of Evidence*, 2nd edition, Cambridge: Cambridge University Press, 2005. Citado por la traducción al castellano de CARBONELL, F. y AGÜERO, C.: ANDERSON, T., SCHUM, D. y TWINING, W., *Análisis de la prueba*, Madrid: Marcial Pons, 2015.
- CALLEN, C., 2008: «Cognitive Strategies and Models of Fact Finding», en JACKSON, J., LANGER, M., TILLERS, P., *Crime, Procedure and Evidence in a comparative and international context. Essays in honour of Professor Mirjan Damaška*, Oxford: Hart Publishing.
- COLOMER HERNÁNDEZ, I., 2003: *La motivación de las sentencias: sus exigencias constitucionales y legales*, Valencia: Tirant lo Blanch.
- CONTI GÓMEZ, M. E., TOLEDO, A. C., 2012: «El juicio por jurados como democratización de la justicia», en *Revista de Derecho Penal*, N° 3: 89-111.
- DEI VECCHI, D., 2020: «Prueba Libre, justificación epistémica y el noble sueño de los estándares de prueba», en *Revista de Derecho (Valdivia)*, 33 (2): 25-58 -consultado en prepublicación de www.academia.edu.
- DIGES, M., PÉREZ MATA, N., 2014: «La prueba de identificación desde la psicología del testimonio», en *AAVV Identificaciones fotográficas y en rueda de reconocimiento. Un análisis desde el derecho procesal penal y la psicología del testimonio*, Madrid: Marcial Pons.
- DUCE, M., 2013: «Errores del sistema y condena de inocentes: nuevos desafíos para nuestra justicia penal acusatoria», en *El Modelo Adversarial en Chile*, Santiago: Legal Publishing: 1-65.
- DUCE, M. 2017: «Los reconocimientos oculares: una aproximación empírica a su funcionamiento y algunas recomendaciones para su mejora», en *Política criminal* [online], N. 23 (vol.12): 291-379.
- FERRAJOLI, L., 1990: *Diritto e ragione. Teoria del galantismo penale*, Roma: Laterza, citado por la traducción al castellano *Derecho y Razón. Teoría del Garantismo Penal*, 9 edición, Madrid: Trotta, 2009.
- FERRER BELTRÁN, J., 2003: *Prueba y Verdad en el Derecho*, 2ª edición, Madrid-Barcelona: Marcial Pons, 2005.
- 2007: *La valoración racional de la prueba*, Madrid-Barcelona-Buenos Aires: Marcial Pons.
- 2013: «Prolegómenos para una teoría sobre los estándares de prueba. El *test case* de la responsabilidad del Estado por prisión preventiva errónea», en PAPAYANNIS, D., PEREIRA FREDES, E. (eds.): *Filosofía del derecho privado*, Madrid: Marcial Pons.

- 2020, «Sobre el deber de motivación de las decisiones probatorias y el juicio por jurados. La sentencia v.r.p., v.p.c. y otros vs. Nicaragua de la Corteidh», en *Quaestio facti. Revista Internacional sobre Razonamiento Probatorio*, nº 1: 359-382.
- GASCÓN ABELLÁN, M., 1999: *Los hechos en el derecho. Bases argumentales de la prueba*, 2ª edición, Madrid-Barcelona: Marcial Pons, 2004.
- GONZÁLEZ LAGIER, D., 2003: «Hechos y Argumentos (racionalidad epistemológica y prueba de los hechos)», en *Jueces para la Democracia*, n. 47: 35-50.
- 2005: *Quaestio Facti. Ensayos sobre prueba, causalidad y acción*, Lima-Bogotá: Palestra-Temis,
- 2020: «¿Es posible formular un estándar de prueba preciso y objetivo? Algunas dudas desde un enfoque argumentativo de la prueba», en *Revista Telmática de Filosofía del Derecho*, n. 23: 79-97.
- KAPLOW, L., 2012: «Burden of Proof», en *Yale Law Journal*, 121:738.
- LAUDAN, L. 2006: *Truth, Error, and Criminal Law. An Essay in Legal Epistemology*, Cambridge: Cambridge University Press. Citado por la traducción castellana de VÁZQUEZ, C. y AGUILERA, E.: *Verdad, error y proceso penal. Un ensayo de epistemología jurídica*, Madrid, Buenos Aires, Barcelona, Sao Paulo: Marcial Pons, 2013.
- 2011: *El estándar de prueba y las garantías en el proceso penal*, Buenos Aires: Hammurabi.
- NIETO, A., 2000: *El arbitrio judicial*, Barcelona: Ariel.
- NINO, CARLOS. S.: 1989, *Ética y derechos humanos. Un ensayo de fundamentación* (2º ed.) Buenos Aires: Astrea.
- RAMÍREZ ORTIZ, J. L., 2020: «El testimonio único de la víctima en el proceso penal desde la perspectiva de género», en *Quaestio Facti. Revista Internacional Sobre Razonamiento Probatorio*, n. 1: 201-246.
- SARRABAYROUSE, E., 2007: *Responsabilidad Penal por el Producto*, Buenos Aires: Ad-Hoc.
- 2020: «La deliberación de los jueces en el CPPF, ley 27.482. Problemas Generales y Particulares», en proceso de publicación. El artículo es una ampliación de SARRABAYROUSE, E., 2019 «La deliberación de los jueces y su forma de votar tras el juicio oral y público en el CPPN, ley 27.063. Problemas generales y particulares», en «*Proceso penal adversarial*», n. 2:191- 212).
- SCHIAVO, N., 2016: *El Juicio por Jurados*, Hammurabi, Buenos Aires: 96 y ss.
- TARUFFO, M., 2009: *Páginas sobre justicia civil*, Madrid: Marcial Pons.
- 2010: *Simplemente la verdad. El Juez y la Confirmación de los hechos*. Madrid-Barcelona-Buenos Aires: Marcial Pons.
- TOULMIN, S., 1964: *The uses of argument*, Cambridge: Cambridge University Press.

A PRIVACIDADE DIGITAL POSTA À PROVA NO PROCESSO PENAL

Paulo de Sousa Mendes*

Professor Catedrático da Faculdade de Direito
da Universidade de Lisboa
paulosousamendes@yahoo.com

RESUMO: o artigo 8.º da Convenção Europeia dos Direitos Humanos trata do respeito pela privacidade. O Quarto Aditamento à Constituição dos EUA trata igualmente do respeito pela privacidade. O processo penal deve assegurar a proteção da privacidade, na medida do possível, e o direito em ação deve respeitar os limites da cópia de dados eletrónicos e as restrições impostas à análise externa do acervo recolhido. Entre os aspetos críticos da análise externa de dados eletrónicos, avulta a questão do procedimento a adotar pelo investigador criminal diante dos conhecimentos fortuitos, uma questão que é analisada detalhadamente neste artigo. A jurisprudência do Tribunal Europeu dos Direitos Humanos caracteriza-se por alguma ineficácia na criação de remédios para a violação da privacidade no processo penal, designadamente no tocante à cópia de dados eletrónicos e à análise externa do acervo recolhido, desde logo porque não comina a exclusão das evidências produzidas por computador que tenham sido obtidas ilicitamente, o que deveria ser o caso, à luz do princípio do processo equitativo. O conhecimento das diretrizes e do direito jurisprudencial norte-americano

* O presente texto resultou da passagem a escrito de conferências proferidas pelo autor nos seguintes eventos: «The seizure of computer files and e-mails», no International Forum the Europeanization of Evidence Law in Transnational and Domestic Criminal Justice, Messina e Siracusa, Itália, 28 e 29 de maio de 2019; «La búsqueda de archivos informáticos y correos electrónicos», no Curso EJ – 1901, Las Garantías de Investigados y Acusados desde la Dimensión Europea (Buenas Prácticas Procesales), Barcelona, Espanha, 13, 14 e 15 de novembro de 2019; «Die Beschlagnahme von Computerdateien und E-Mails», organizada pelo Institut für Deutsches und Europäisches Strafprozessrecht und Polizeirecht der Universität Trier (ISP), na antiga Schwurgerichtssaal (Arbeits- und Sozialgericht, Dietrichstraße 13, Eingang Justizstraße), Trier, Alemanha, em 15 de janeiro de 2020.

no representa um contributo valioso para o aprofundamento da jurisprudência de Estrasburgo, na sua dupla função decisória e nomofilática, assim como para o aperfeiçoamento dos ordenamentos jurídicos nacionais europeus ao nível legislativo e ao nível da prática jurisprudencial.

PALAVRAS-CHAVE: conhecimentos fortuitos; doutrina jurisprudencial da visibilidade imediata; evidências produzidas por computador; mandado de busca digital; pesquisa externa; privacidade digital.

THE DIGITAL PRIVACY AT STAKE IN CRIMINAL JUSTICE

ABSTRACT: The article 8 of the European Convention on Human Rights deals with the need of respect for privacy. The 4th Amendment to the US Constitution addresses the same issue. The criminal procedure ensures the protection of privacy, as far as possible, and the praxis must recognize that, for privacy reasons, there should be limits to the seizure of computer records and to its off-site analysis. Among the critical aspects of the off-site analysis of electronic data, the approach to be adopted by the law enforcement officers in the face of serendipity findings largely remains an unanswered question. The doctrine of the European Court of Human Rights is characterized by some ineffectiveness in the creation of remedies for the violation of privacy in criminal proceedings, namely with regard to the search of data contained on computers and its off-site analysis, mainly because it does not order the exclusion of computer-generated evidence that has been illegally gathered, which should be the case, in the light of the fair trial principle. The knowledge of the North American case law and guidelines could represent a valuable contribution to the improvement of Strasbourg jurisprudence, in its decision-making role and also in its nomophylactic function, as well as to the improvement of European national legal systems at the legislative level and at the level of decision-making processes.

KEYWORDS: computer-generated evidence; digital search warrant; off-site search; plain view doctrine; privacy; seizure by chance.

SUMARIO: 1. INTRODUÇÃO.— 2. O DIREITO À PRIVACIDADE DIGITAL NA JURISPRUDÊNCIA DE ESTRASBURGO: 2.1. O caso *Sérvulo & Associados – Sociedade de Advogados, RL e Outros vs. Portugal* (2015). 2.2. A violação do direito à privacidade e o seu remédio ao nível do processo equitativo.— 3. O DIREITO À PRIVACIDADE DIGITAL NO DIREITO JURISPRUDENCIAL DOS ESTADOS UNIDOS DA AMÉRICA: 3.1. *As Diretrizes Federais para Busca e Apreensão de Computadores* (2009). 3.2. *Pesquise antes de apreender ou...* 3.3. *Apreenda primeiro e depois logo se vê.* 3.4. *A pesquisa externa.* 3.5. *A doutrina jurisprudencial da visibilidade imediata.* 3.6. *O caso Estados Unidos vs. Carey* (1999). 3.7. *O caso State vs. Schroeder* (2000).— 4. CONCLUSÕES.— 5. BIBLIOGRAFIA.

RECOMMENDED CITATION: SOUSA MENDES, PAULO DE, 2020: «A privacidade digital posta à prova no processo penal», in *Quaestio facti*, 2: 225-250. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22487

1. INTRODUÇÃO

A migração para o ciberespaço origina novas ameaças à privacidade, desde logo porque todas as facetas da vida ficam expostas de forma nunca antes vista no mundo físico.

O processo penal acompanha a migração para o ciberespaço, desde a desmaterialização dos autos até à audição de testemunhas por videoconferência eletrónica. No presente texto, interessa-nos sobremaneira a recolha, a custódia e a análise da prova digital. Em processos por delitos económicos e financeiros a prova é quase totalmente digital. Os documentos que servem de meios de prova são digitais. Dada a imensa informação potencialmente acumulada em bits e bytes por comparação com o mundo físico, cabe então perguntar qual é a proteção da privacidade que resta no domínio da prova digital?

A pergunta é feita tanto à luz do artigo 8.º da Convenção Europeia dos Direitos Humanos e da jurisprudência de Estrasburgo como diante do Quarto Aditamento à Constituição dos Estados Unidos da América e do correspondente direito jurisprudencial (*case law*). As soluções aí encontradas ajudarão a resolver o mesmo problema jurídico ao nível de cada ordenamento nacional, quanto mais não seja porque assistimos a uma hibridização do processo penal, ademais incrementada na ciberrealidade em que vivemos.

2. O DIREITO À PRIVACIDADE DIGITAL NA JURISPRUDÊNCIA DE ESTRASBURGO

O artigo 8.º da Convenção Europeia dos Direitos Humanos (doravante, a Convenção) impõe o respeito pela privacidade. Mais exatamente, o artigo 8.º, n.º 1, da Convenção protege o direito ao respeito da vida privada e familiar, do domicílio e da correspondência.

A privacidade é um conceito mais vasto do que parece. Realmente, o Tribunal Europeu dos Direitos Humanos (doravante, o TEDH) tem vindo a fazer uma interpretação extensiva da Convenção, aplicando o artigo 8.º à proteção da informação guardada em servidores, computadores, ficheiros informáticos e e-mails, como aconteceu nos casos *Leander v Sweden* (1987)¹, *Amann v Switzerland* (2000)², *Rotaru v Romania* (2000)³, *Copland v United Kingdom* (2007)⁴ e *Wieser and Bicos Beteiligungen GmbH v Austria* (2007)⁵. Além de que tem alargado o conceito de privacidade

¹ *Leander v Sweden* (queixa n.º 9248/81), de 26 de março de 1987, § 48.

² *Amann v Switzerland* (queixa n.º 27798/95), de 16 de fevereiro de 2000, § 65.

³ *Rotaru v Romania* (queixa n.º 28341/95), de 4 de maio de 2000, §§ 42-43.

⁴ *Copland v United Kingdom* (queixa n.º 62617/00), de 3 de abril de 2007.

⁵ *Wieser and Bicos Beteiligungen GmbH v Austria* (queixa n.º 74336/01), de 16 de janeiro de 2007, § 45.

à vida profissional não só dos trabalhadores, mas também das empresas, de modo que o ambiente informático do local de trabalho acaba por estar incluído na proteção da privacidade, como aconteceu no caso *Société Colas Est and other v France* (2002)⁶. Geralmente, o TEDH basta-se com uma ou duas frases de fundamentação para declarar essa proteção, como aconteceu no caso *Copland v United Kingdom* (2007): «§ 41. De acordo com a jurisprudência do Tribunal, as ligações telefónicas de estabelecimentos comerciais são *prima facie* cobertas pelas noções de “vida privada” e “correspondência” para os fins do artigo 8.º, n.º 1 (ver *Halford*, já referido, § 44 e *Amann v Suíça* [GC], n.º 27798/95, § 43, CEDH 2000 II). Segue-se logicamente que os e-mails enviados do trabalho devem igualmente ser protegidos pelo artigo 8.º, assim como as informações recolhidas através de monitoramento do uso pessoal da Internet»⁷.

Seja como for, a proteção da privacidade não é absoluta. Existem situações em que a autoridade pública pode interferir no direito ao respeito pela vida privada e familiar, pelo lar e pela correspondência. Nos termos do n.º 2 do artigo 8.º da Convenção, a ingerência só é permitida, porém, quando estiver prevista na lei e constituir uma providência que, numa sociedade democrática, seja necessária para a segurança nacional, a segurança pública, o bem-estar económico do país, a defesa da ordem e a prevenção das infrações penais, a proteção da saúde ou da moral, ou a proteção dos direitos e das liberdades de terceiros. Quando é chamado a pronunciar-se sobre se o artigo 8.º da Convenção foi ou não violado, o TEDH verifica, primeiro, se houve ingerência em algum dos direitos atrás elencados; segundo, se tal ingerência estava coberta por uma norma habilitante da jurisdição em que foi praticada; e, terceiro, se a medida realizada era proporcional em ordem a satisfazer uma necessidade social premente numa sociedade democrática⁸.

Ainda não há muitos acórdãos do TEDH que tratem da aplicação do artigo 8.º da Convenção à prova digital no processo penal, mas já há os suficientes para se tirar um retrato da jurisprudência de Estrasburgo a tal respeito⁹. Seleccionámos um que consideramos especialmente representativo, a saber: o caso *Sérvulo v Portugal* (2015)¹⁰.

⁶ *Société Colas Est and other v France* (queixa n.º 37971/97), de 16 de abril de 2002, § 40.

⁷ Em língua original (*Copland v United Kingdom* (2007)): «§ 41. According to the Court's case-law, telephone calls from business premises are *prima facie* covered by the notions of 'private life' and 'correspondence' for the purposes of Article 8 § 1 (see *Halford*, cited above, § 44 and *Amann v. Switzerland* [GC], no. 27798/95, § 43, ECHR 2000 II). It follows logically that e-mails sent from work should be similarly protected under Article 8, as should information derived from the monitoring of personal internet usage».

⁸ De Hert / Gutwirth, 2009: 15-17.

⁹ A título de exemplo, recomenda-se a leitura dos seguintes acórdãos do TEDH: *Robathin v Austria* (queixa n.º 30457/06), de 3 de julho de 2012; *Bernh Larsen Holding As v Norway* (queixa n.º 24117/08), de 14 de março de 2013; *Trabajo Rueda v Spain* (queixa n.º 32600/12), de 30 de maio de 2017; e *Ivashchenko v Russia* (queixa n.º 61064/10), de 13 de maio de 2018.

¹⁰ *Sérvulo & Associados – Sociedade de Advogados, Rl v Portugal* (queixa n.º 27013/10), de 3 de setembro de 2015.

2.1. O caso *Sérvulo & Associados – Sociedade de Advogados, RL e Outros vs. Portugal* (2015)

No caso *Sérvulo v Portugal* (2015), o TEDH considerou, por maioria (6 votos contra 1), que o artigo 8.º da Convenção não foi violado. O caso dizia respeito a uma busca em escritório de advogados e à apreensão de ficheiros e mensagens alojados nos servidores da sociedade e nos computadores de secretária usados por alguns dos seus advogados, no âmbito de dois processos penais, instaurados pelo Ministério Público do Departamento Central de Investigação e Ação Penal (DCIAP), em 2006, relativamente a crimes de corrupção, participação económica em negócio, branqueamento e prevaricação conexos com a compra, em 2004, de dois submarinos pelo Governo português a um consórcio alemão¹¹. Nem a sociedade, nem os respetivos advogados eram arguidos nos mencionados processos, embora um antigo advogado dessa sociedade tenha sido, de facto, constituído arguido em 29 de setembro de 2009, anteriormente às diligências de busca e apreensão¹². O TEDH concluiu que a apreensão de ficheiros informáticos e mensagens de correio eletrónico no escritório de advocacia havia sido efetuada com garantias processuais suficientes contra eventuais abusos, arbitrariedades ou violações do sigilo profissional de advogado¹³.

O despacho do juiz de instrução e respetivos mandados autorizavam o acesso aos servidores e computadores em que pudessem encontrar-se os documentos eletrónicos

¹¹ Cabe aqui antecipar que o primeiro processo terminaria arquivado pelo Ministério Público do DCIAP, em 17 de dezembro de 2014, e o segundo processo terminaria com a absolvição de todos os arguidos pelo Tribunal da Relação de Lisboa (TRL), em 19 de março de 2015, confirmando a decisão de primeira instância, de 14 de fevereiro de 2014. Veja-se, a propósito, *Sérvulo v Portugal* (queixa n.º 27013/10), de 3 de setembro de 2015, §§ 45-48.

¹² Em 6 de outubro de 2009, um dos advogados da sociedade acabaria sendo constituído arguido, após a busca e precisamente por causa desta, por se ter descoberto que também teria acompanhado as negociações que constituíam o alvo dos inquéritos criminais então em curso. Em 19 de outubro de 2009, esse mesmo advogado requereu ao juiz de instrução do Tribunal Central de Investigação Criminal (TCIC) a declaração de nulidade da sua constituição como arguido, alegando que não fora ouvido nessa qualidade, nem fora informado dos factos que lhe eram imputados, sendo que o único móbil da sua constituição como arguido teria sido justificar a apreensão dos dados armazenados no seu computador. Por despacho de 3 de novembro de 2009, o juiz de instrução do TCIC indeferiu o pedido de declaração de nulidade da constituição de arguido, tendo o advogado, na sequência, interposto recurso para o TRL. Por acórdão de 15 de abril de 2010, o TRL revogou o despacho recorrido, considerando que a constituição do advogado como arguido tinha sido ilegal, anulando essa constituição e ordenando a restituição de toda a correspondência referente a ele que havia sido apreendida. Cabe aqui reproduzir o trecho mais representativo da decisão: «[...] não é a apreensão de documentação profissional num escritório de advogados que permite fundamentar a constituição do advogado como arguido, antes sendo a constituição de um advogado como arguido que abre a possibilidade de apreensão de correspondência profissional do mesmo» (Ac. TRL 15/04/2010, proc. n.º 56/06.2TELSB-B.L1-9, Relator: Fátima Mata-Mouros). Online: <http://www.dgsi.pt/jtrl.nsf/e6e1f17fa82712ff80257583004e3ddc/fdc745090a69eaa18025770b003dd2eb?OpenDocument> (consultado em 03/08/2020). Este acórdão também é citado e, ademais, elogiado por Costa Ramos / Pinto de Abreu / Cordeiro, 2020: 257, n. 59.

¹³ Press Release issued by the Register of the Court – ECHR 267 (2015), 03/09/2015.

relacionados com as negociações de compra dos submarinos. Os documentos eletrónicos deveriam ser selecionados através de pesquisa por meio de 35 palavras-chave indicadas na promoção do Ministério Público, tais como, por exemplo, nomes de bancos e personalidades ou palavras genéricas como *spread* ou *swap*. Os documentos eletrónicos sinalizados pelo emprego das palavras-chave deveriam ser copiados para um suporte digital autónomo e apresentados ao juiz de instrução para levantamento do sigilo profissional e validação da apreensão, se fossem relevantes para a investigação em curso. No final das contas, o juiz de instrução ordenou que se fizesse a juntada aos autos de 89.000 ficheiros de computador e 29.000 mensagens de correio eletrónico¹⁴.

O TEDH decidiu que não ocorrera qualquer violação do artigo 8.º da Convenção, considerando que: *a)* a ingerência no direito ao respeito da vida privada ocorrida no caso concreto estava prevista por lei, ainda que nenhuma lei previsse especificamente as pesquisas e as apreensões em sistemas informáticos, o que só veio a suceder posteriormente, com a entrada em vigor da Lei n.º 109/2009, de 15 de setembro¹⁵; *b)* as buscas e apreensões visavam um fim legítimo, pois foram autorizadas no âmbito de um inquérito relativo à prática de crimes de corrupção, participação económica em negócio, branqueamento e prevaricação e, por conseguinte, tinham em vista a repressão de infrações penais¹⁶; *c)* a ingerência era necessária numa sociedade democrática, enquanto necessidade social imperiosa e proporcional ao fim legítimo visado, dado que as buscas e apreensões em escritório de advogados foram autorizadas e presididas por um juiz, foram sujeitas a controlo adicional por via da intervenção do Vice-Presidente do TRL e contaram com a presença de observadores independentes, designadamente o representante do Ordem dos Advogados (OA)¹⁷; *d)* apesar de entre as palavras-chave utilizadas na pesquisa informática estarem termos correntes num escritório de advocacia especializado em áreas de negócio, incluindo contratação pública, as mesmas eram proporcionais ao fim legítimo visado e a lei portuguesa contém garantias processuais suficientes contra eventuais abusos, arbitrariedades e restrições do sigilo profissional de advogado, que foram observadas no caso concreto através da autorização e intervenção do juiz de instrução e da ulterior reclamação para o Vice-Presidente do TRL¹⁸; *e)* ainda que apenas existisse um juiz no TCIC, o mesmo interveio apenas na qualidade de garante das liberdades individuais, para controlar a legalidade das buscas e apreensões e proteger o sigilo profissional, não tendo competência para iniciar uma investigação, sendo o alegado a este respeito pelos queixosos insuficiente para gerar dúvidas fundadas acerca da eficácia do controlo por ele exercido¹⁹.

¹⁴ *Sérvulo v Portugal* (2015), *cit.*, § 13.

¹⁵ *Sérvulo v Portugal* (2015), *cit.*, § 96.

¹⁶ *Sérvulo v Portugal* (2015), *cit.*, § 97.

¹⁷ *Sérvulo v Portugal* (2015), *cit.*, § 100.

¹⁸ *Sérvulo v Portugal* (2015), *cit.*, § 111.

¹⁹ *Sérvulo v Portugal* (2015), *cit.*, § 119.

O juiz *ad-hoc* português (Saragoça da Matta), discordando da maioria, considerou que ocorrera uma violação do artigo 8.º da Convenção, por várias razões, que disse não se basearem na avaliação da lei portuguesa, mas resultarem antes da aplicação da Convenção aos factos do caso, a saber, entre outras: *a)* não ser admissível que, para recuperar os documentos desaparecidos do Ministério da Defesa, as autoridades judiciárias escolhessem uma diligência intrusiva, ademais muito ampla, contra um escritório de advocacia e sem agirem contra os responsáveis pela irregularidade praticada dentro do próprio Ministério²⁰; *b)* não ser admissível a realização de uma pesquisa informática com base numa lista de 35 palavras-chave que incluía palavras que fazem parte do jargão de escritórios de advocacia, para além de nomes de bancos e personalidades, pois o objeto da pesquisa de dados informáticos deveria ter sido delimitado de uma forma muito mais rigorosa, ponderando as necessidades da investigação e os contrapostos direitos dos advogados e dos seus clientes²¹; *c)* ainda que as garantias processuais contra eventuais arbitrariedades e restrições excessivas do sigilo profissional de advogado previstas pela lei portuguesa sejam adequadas, em abstrato, apenas ficou demonstrado que o procedimento previsto na lei foi seguido, mas não que as garantias tenham sido efetivas (designadamente quanto à salvaguarda do sigilo profissional) quanto a outros clientes do escritório cujos documentos, ficheiros informáticos e e-mails foram apreendidos e posteriormente transferidos para utilização noutro processo²²; *d)* tendo em conta o elevado número de mensagens de correio eletrónico apreendidas, só se poderia concluir que a pesquisa fora, de facto, desproporcional em relação ao fim prosseguido²³; *e)* não se poderia dizer que, no caso concreto, os receios manifestados pelos queixosos contra o juiz de instrução fossem infundados, dado que a sua intervenção teria sido, na melhor das hipóteses, meramente formal²⁴; *f)* os elementos apreendidos irrelevantes não foram devolvidos aos queixosos, mas foram abusivamente emprestados para valoração noutros processos²⁵. Assim sendo, o juiz dissidente concluiu que o TEDH deveria ter reconhecido a violação do artigo 8.º da Convenção.

O caso *Sérvulo v Portugal* tem características que o tornam notado, mas não necessariamente por a decisão de Estrasburgo concitar o aplauso da doutrina académica, antes pelo contrário. Alguma doutrina refere que a presente decisão de Estrasburgo é ilustrativa da necessidade de estabelecer padrões e proteções adequados à salvaguarda da confidencialidade entre o advogado e o cliente (*lawyer-client confidentiality*)²⁶. O privilégio da confidencialidade entre o advogado e o seu cliente é um domínio especial do direito à privacidade, à luz da jurisprudência de Estrasburgo. Por isso

²⁰ *Sérvulo v Portugal* (2015), *cit.*, Opinião Dissidente, III/1.

²¹ *Sérvulo v Portugal* (2015), *cit.*, Opinião Dissidente, III/2.

²² *Sérvulo v Portugal* (2015), *cit.*, Opinião Dissidente, III/3.

²³ *Sérvulo v Portugal* (2015), *cit.*, Opinião Dissidente, III/3.

²⁴ *Sérvulo v Portugal* (2015), *cit.*, Opinião Dissidente, III/6.

²⁵ *Sérvulo v Portugal* (2015), *cit.*, Opinião Dissidente, III/7.

²⁶ BACHMAIER WINTER / THAMAN, 2020: 57.

mesmo, estranhamos que a decisão do caso *Sérvulo vs. Portugal* (2015) espelhe uma tão grande tolerância do TEDH com a devassa em larga escala de ficheiros informáticos e mensagens de correio eletrónico cobertos pelo segredo profissional da advocacia. A propósito, cabe aqui a comparação com o procedimento adotado no mundo anglo-saxónico. Em Inglaterra e Gales, assim como nos Estados Unidos da América (doravante, os EUA), existe um procedimento especial para buscas em escritórios de advocacia quando o objetivo seja a apreensão de documentos protegidos por segredo profissional²⁷. Nestes casos, a preferência das autoridades vai para a solicitação dos documentos sob cominação de desobediência (*subpoena*)²⁸. Tão-somente em situações de emergência, seja porque o advogado recuse o envio dos documentos requisitados ou porque atue em colusão com o cliente na prática de factos puníveis, é legítimo promover a obtenção de um mandado de busca (*search warrant*) junto de um juiz (*Magistrate Judge*) com vista à entrada no escritório de advocacia. Ainda assim, os inspetores devem começar por informar o advogado de que poderá entregar os documentos antes de se iniciarem as diligências de busca e apreensão no seu escritório²⁹.

À parte a questão do segredo profissional do advogado, a decisão de Estrasburgo no caso *Sérvulo v Portugal* surpreende também pela tolerância com a utilização da lista de 35 palavras-chave de uso corrente num escritório de advocacia especializado em áreas de negócio e contratação pública. Parece que a mera utilização de palavras-chave é considerada pelo TEDH como uma garantia contra abusos e arbitrariedades, ao reduzir o universo da pesquisa informática. Mas a utilização de palavras-chave garante sobretudo a eficientização da mineração de dados do ponto de vista das autoridades de investigação criminal. Seja como for, a utilização de palavras-chave não permite, só por si, joeirar as falsas correspondências, trazendo dados irrelevantes para o acervo probatório acolhido nos autos do processo. Acresce que a possibilidade de falsas correspondências aumenta na razão direta da maior vagueza das palavras-chave escolhidas. Não se deve, pois, perder de vista que os dados irrelevantes para a investigação configuram, do mesmo passo, uma devassa abusiva e arbitrária para os visados. Tão-pouco a utilização de palavras-chave elimina a possibilidade de conhecimentos fortuitos. Neste contexto, os conhecimentos fortuitos acabam sendo acreditados ilimitadamente para utilização em outros processos, em curso ou a instaurar. Foi, aliás, o que aconteceu no caso *Sérvulo v Portugal*, em que —tal como foi destacado no voto dissidente— os elementos apreendidos que eram irrelevantes para a investigação que justificara a diligência intrusiva no escritório da advocacia não foram devolvidos aos queixosos, mas foram emprestados para valoração noutros processos. Cabe aqui perguntar se não devem existir regras estritas quanto à possibilidade do aproveitamento de conhecimentos fortuitos por parte das autoridades de investigação criminal. O requisito da previsão legal do aproveitamento de conhecimentos fortuitos, enquanto ingerência das autoridades de investigação criminal na esfera do direito à privacidade

²⁷ BACHMAIER WINTER / THAMAN, 2020: 57-58.

²⁸ BACHMAIER WINTER / THAMAN, 2020: 57.

²⁹ BACHMAIER WINTER / THAMAN, 2020: 57-58.

de qualquer pessoa, nos termos do n.º 2 do artigo 8.º da Convenção, não se verificava à data das buscas no escritório de advocacia, uma vez que o regime legal das buscas e apreensões é omissivo a tal respeito³⁰. O TEDH revela insensibilidade quanto à excepcionalidade do aproveitamento de conhecimentos fortuitos.

2.2. A violação do direito à privacidade e o seu remédio ao nível do processo equitativo

A jurisprudência do TEDH tem contribuído, como é sua função, para a edificação de um menor denominador comum garantístico na Europa, à luz da Convenção³¹. Mas também é verdade que a Convenção é uma carta de direitos mínimos, pois tem de abrigar ordenamentos jurídicos nacionais muito diversos entre si e tem de dar respostas para todos³². Na verdade, os ordenamentos jurídicos nacionais abrangidos ainda estão longe de partilhar na prática os mesmos princípios e garantias penais. Sempre que é chamado a decidir quaisquer casos de violação da CEDH, o TEDH usa de alguma contenção, dado que tem de lidar, à vez, com os diferentes ordenamentos jurídicos nacionais. Não admira, pois, que as decisões e os remédios sejam minimalistas. Tal poderá ser dececionante para quem for atrás de soluções jurídicas vanguardistas na jurisprudência do TEDH, já que estaria a procurá-las no lugar errado. Paradoxalmente, a jurisprudência do TEDH ganha, afinal, uma importância acrescida por causa do seu carácter moderado. Podemos, assim, dar por adquirido que, onde o TEDH viu uma violação à CEDH, é difícil de dizer o contrário. Mas, onde o TEDH deixou passar uma eventual violação de direitos humanos, é sempre possível discordar. Neste caso, a fundamentação do acórdão em causa é tão importante como os votos dissidentes. Um voto dissidente de hoje pode ser a jurisprudência de amanhã, como todos sabemos desde que o Chief Justice Oliver Wendell Holmes Jr. ficou famoso como protagonista de votos dissidentes³³.

No tocante à violação do artigo 8.º da Convenção, a jurisprudência do TEDH, valendo-se de um raciocínio de ponderação de interesses (*balancing approach*), acaba não extraindo consequências dessa violação para o funcionamento do processo equitativo como um todo (*fair as a whole*)³⁴, à luz do artigo 6.º da Convenção, desde que ao acusado, no caso concreto, tenham sido dadas oportunidades de contestar a prova

³⁰ Veja-se os artigos 174.º a 186.º do Código de Processo Penal (CPP) português. A possibilidade do aproveitamento de conhecimentos fortuitos só está contemplada para as escutas telefónicas e repetidas extensões, nos termos do n.º 7 do artigo 187.º do CPP, na redação da Lei n.º 48/2007, de 29 de agosto. Na atual Lei n.º 109/2009, de 15 de setembro (vulgo: Lei do Cibercrime), também não há previsão do aproveitamento de conhecimentos fortuitos fortuitos, salvo por remissão do n.º 4 do artigo 18.º para o regime das escutas telefónicas.

³¹ KOSTORIS, 2014: 44-60.

³² Os atuais 47 Estados-Membros do Conselho da Europa.

³³ SOUSA MENDES, 2020: 2365.

³⁴ COSTA RAMOS, 2017: 742.

em questão, tenham sido respeitados os seus outros direitos de defesa e não haja dúvidas sobre a fiabilidade da prova – o que é, genericamente, o caso para as provas obtidas em violação do artigo 8.º da Convenção³⁵. Sendo assim, a jurisprudência do TEDH parece não fornecer quaisquer regras de exclusão da prova, as quais, enquanto critérios operativos a nível nacional, possam constituir remédios efetivos contra a utilização de provas obtidas em violação do artigo 8.º da Convenção³⁶. Mas, dizemos nós, a utilização de provas obtidas através da lesão do direito à privacidade não deveria ser indiferente para a noção do processo equitativo como um todo. Pelo contrário, deveríamos esperar do TEDH que impusesse aos Estados, quando fosse o caso, remédios efetivos contra a violação da equidade processual (*fairness*), devolvendo assim o acusado à posição processual em que se encontraria se não fosse a lesão da sua privacidade. Muito embora a jurisdição do TEDH não funcione como última instância face aos ordenamentos jurídicos nacionais, aquele Tribunal até já tem decretado a *restitutio in integrum*, em conformidade com a Recomendação do Comité de Ministros do Conselho Europa N.º R(2000)2, de 19 de janeiro, obrigando à reabertura do processo-crime no ordenamento de origem, sem que isso implique, naturalmente, que o acusado tenha de ser absolvido³⁷. Vamos, pois, esperar que a evolução da jurisprudência do TEDH se dê no sentido de um estreitamento da conexão entre o direito substantivo à privacidade e o direito processual ao processo equitativo, ambos direitos convencionais, ainda que o campo de aplicação do direito à privacidade e a sua violação ocorram muitas vezes fora do processo penal³⁸. Mas essa evolução é improvável enquanto o TEDH se mantiver apegado a um raciocínio de ponderação de interesses que contrapõe, fundamentalmente, o interesse do Estado na preservação de uma prova fiável aos olhos do juiz-julgador para poder produzir uma decisão robusta do ponto de vista factual e o interesse do acusado em defender-se, estando este último interesse assegurado, na perspetiva do TEDH, se o acusado tiver tido a oportunidade no processo penal para contestar essa prova³⁹.

³⁵ COSTA RAMOS, 2017: 756. Segundo a autora citada, o único caso em que o TEDH teria decidido no sentido de ser excluída a prova obtida em violação do artigo 8.º da Convenção (embora este artigo não fosse expressamente mencionado), por a sua utilização ser violadora do artigo 6.º da Convenção, foi o caso *Lisica v Croatia* (queixa n.º 20100/06), de 25 de fevereiro de 2010, §§ 60-62.

³⁶ COSTA RAMOS, 2017: 757.

³⁷ JAHN, 2014: 5-9.

³⁸ COSTA RAMOS, 2017: 760, n. 72.

³⁹ Diversamente, a recolha de prova em processo penal que viole a proibição convencional da tortura (artigo 3.º da Convenção) deve, na ótica do TEDH, implicar a impossibilidade de utilização dessa prova, abstraindo de quaisquer ponderações de interesses. Assim, a utilização de prova ilicitamente obtida em violação do artigo 3.º da Convenção tem levado o TEDH à consideração de que o processo em causa foi desleal (*unfair*). Neste sentido, Costa Ramos, 2017: 760.

3. O DIREITO À PRIVACIDADE DIGITAL NO DIREITO JURISPRUDENCIAL DOS ESTADOS UNIDOS DA AMÉRICA

O Quarto Aditamento à Constituição dos EUA declara que o direito do povo à inviolabilidade de suas pessoas, casas, papéis e haveres contra busca e apreensão arbitrárias não poderá ser infringido; e nenhum Mandado será expedido a não ser mediante indícios de culpabilidade confirmados por Juramento ou declaração, e particularmente com a descrição do local da busca e a indicação das pessoas ou coisas a serem apreendidas⁴⁰.

O Quarto Aditamento à Constituição dos EUA oferece uma proteção substantiva maior do que a do artigo 8.º da Convenção, porquanto limita as possibilidades de ingerência legítima da autoridade pública na privacidade às diligências de busca e apreensão (*search and seizure*), as quais, por sua vez, dependem de haver uma probabilidade razoável de descoberta de provas relacionadas com um crime (*probable cause*) e haver uma promoção formal ou declaração junto de um juiz competente e imparcial (que não tenha tido contacto com o caso em qualquer outro contexto) para obtenção de um mandado de busca indicando qual o fundamento daquela probabilidade razoável⁴¹. O artigo 8.º da Convenção não exige causa provável, nem mandado, mas exige, mais vagamente, que qualquer ingerência da autoridade pública na privacidade venha prevista na lei e constitua uma providência necessária numa sociedade democrática.

O Supremo Tribunal dos EUA (*Supreme Court of the United States*) adotou a regra de exclusão relativa a buscas e apreensões ilegais (*search and seizure exclusionary rule*) no caso *Weeks v United States* (1914)⁴² e tornou-a aplicável não apenas ao nível federal, mas também ao nível estadual no caso *Mapp v Ohio* (1961)^{43/44}. Se quisermos, a regra de exclusão funciona num plano equivalente ao da proibição de valoração de prova⁴⁵. A proibição de produção de prova, por sua vez, encontra-se no Quarto Aditamento à Constituição dos EUA, que, como vimos, proíbe as buscas e apreensões injustificadas (*unreasonable searches and seizures*) e sem mandado judicial. Compare-se com a violação do artigo 8.º da Convenção, relativamente à qual a jurisprudência do TEDH, lamentavelmente, não comina quaisquer regras de exclusão da prova.

⁴⁰ Em língua original (*Fourth Amendment of the U.S. Constitution*): «*The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.*»

⁴¹ Há várias exceções à exigência de mandado judicial, mas não relevam imediatamente para a discussão. Oportunamente voltaremos a este ponto no próprio texto.

⁴² *Weeks v United States*, 232 U.S. 383, 34 S.Ct. 341, 58 L.Ed. 652 (1914).

⁴³ *Mapp v Ohio*, 367 U.S. 643 (1961).

⁴⁴ Sobre os antecedentes, a evolução e o conteúdo da regra de exclusão relativa a buscas e apreensões ilegais, LAFAVE, 2004 (vol. 1): 3-26, LAFAVE, 2011-2012 (Pocket Part): 1-101 e KAMISAR, 2003: 121-140.

⁴⁵ AMBOS, 2010: 19.

3.1. As Diretrizes Federais para Busca e Apreensão de Computadores (2009)

Em 1994, o Departamento de Justiça publicou as Diretrizes Federais para Busca e Apreensão de Computadores (*Federal Guidelines for Searching and Seizing Computers*)⁴⁶. As Diretrizes foram atualizadas através de Suplementos (*Supplements*), em 1997 e 1999, e foram sujeitas a uma ampla reformulação em forma de Manual de Busca e Apreensão de Computadores e Recolha de Prova Digital em Investigações Criminais (*Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations*), publicado em 2001 e amplamente revisto em 2009⁴⁷.

De acordo com as Diretrizes de 2009, as pesquisas informáticas em disco rígido (*hard-drive*) ou em outros ambientes informáticos (*computer media*) carecem de autorização judicial que contemple, entre outros aspetos, o modo específico de realização dessa diligência de obtenção de provas. Na grande maioria dos casos, a análise forense (*forensic analysis*) de um disco rígido (ou outro ambiente informático) leva tempo de mais para poder ser realizada no local (*on-site*) durante a execução inicial do mandado de busca e apreensão (*search warrant*)⁴⁸. A decisão mais importante que deve constar do mandado é se podem ser apreendidos computadores e demais equipamentos ou apenas as informações que o *hardware* contém. Na primeira hipótese, o mandado deve descrever o próprio *hardware*⁴⁹. Se a causa provável que justificou a diligência estiver relacionada apenas com certas informações, então o mandado deve descrever as informações a ser apreendidas e autorizar a sua apreensão em qualquer suporte em que possam ser armazenadas, seja eletrónico ou não^{50/51}. Neste caso, os inspetores (*officers*) devem recolher as informações que se enquadram no escopo do mandado através de cópia eletrónica de todo o dispositivo de armazenamento (*image copy*)⁵²,

⁴⁶ Online: https://epic.org/security/computer_search_guidelines.txt (consultado em 11/08/2020).

⁴⁷ Online: <https://www.justice.gov/criminal-ccips/ccips-documents-and-reports> (consultado em 11/08/2020).

⁴⁸ *Guidelines* 2009, p. 72.

⁴⁹ Naturalmente, o mandado judicial não precisa de ser excessivamente técnico (*overly technical*), podendo seguir uma visão de senso comum (*commonsensical*) e uma abordagem prática (*practical*) para descrever genericamente o tipo de equipamentos que podem ser apreendidos. Veja-se *United States v Ventresca*, 380 U.S. 102, 108 (1965).

⁵⁰ As anteriores Diretrizes de 2001 recomendavam, como um dos modos possíveis de executar pesquisas em computadores (II.B.1. *Basic Strategies for Executing Computer Searches*), a pesquisa no local e a impressão de ficheiros específicos nesse momento («[s]earch the computer and print out a hard copy of particular files at that time»), mas as Diretrizes de 2009 já não mencionam expressamente esta possibilidade.

⁵¹ *Guidelines* 2009, p. 70.

⁵² «[A copy that] duplicates every bit and byte on the target drive including all files, the slack space, Master File Table, and metadata in exactly the order they appear on the original». Veja-se *United States v Vilar*, 2007 WL 1075041, *35 n.22 (S.D.N.Y. Apr. 4, 2007) e *United States v Stierhoff*, 477 F. Supp. 2d 423, 439 & n.8 (D.R.I. 2007). Do ponto de vista técnico, o curso de ação recomendável implica que se comece por preservar uma cópia forense do ambiente informático antes de se iniciar qualquer pesquisa, a fim de prevenir a contaminação ou destruição de evidências. As ferramentas forenses convencionais

feita ou não no local⁵³, e posteriormente devem produzir uma cópia de trabalho para poderem realizar o exame fora do local (*off-site examination*) por meio de programas de mineração de dados que lhes permitam segregar aqueles registros que correspondam (*responsive records*) aos crimes abrangidos pelo mandado judicial em execução⁵⁴. Mas o mandado judicial não pode simplesmente autorizar uma pesquisa e apreensão de «todos os registros» («all records»)⁵⁵, sob pena de ser um mandado genérico inconstitucional (*unconstitutional general warrant*)⁵⁶. Em vez de os inspetores fazerem uma cópia eletrônica de todo o dispositivo de armazenamento, é, pois, admissível que façam uso de técnicas forenses (*forensic techniques*) que, por exemplo, lhes permitam restringir o universo da pesquisa e apreensão apenas aos ficheiros que contenham determinadas palavras-chave (*keywords*) relacionadas com os crimes investigados ao abrigo do mandado judicial, mas esta é uma possibilidade que não deverá ser imposta pela autoridade judicial, embora deva constar da promoção que lhe for dirigida. A restrição do universo de pesquisa e apreensão pode interessar aos titulares da investigação criminal, sobretudo se pensarmos em processos-crime que envolvam empresas e vastas quantidades de dados informáticos. Seja como for, é inadmissível impor qualquer limitação significativa (como uma restrição a pesquisas por palavras-chave) às técnicas forenses que os inspetores pretendam usar para encontrar as evidências que caibam no escopo de um mandado judicial⁵⁷.

A apreensão do equipamento, em princípio, só é possível se o mesmo for contrabando, prova, instrumento ou produto de um crime, nos termos da Regra 41(c)⁵⁸ do Regulamento Federal de Processo Penal (*Federal Rules of Criminal Procedure*)⁵⁹. Se o equipamento for apenas um dispositivo de armazenamento de evidências, os agentes de autoridade pública podem, a título excecional, apreendê-lo só se não existirem alternativas menos disruptivas⁶⁰.

oferecem a capacidade de criar cópias precisas e infalsáveis (graças à atribuição de valores *hash*), contendo todas as características do ambiente original, incluindo a informação sobre os ficheiros excluídos e o espaço no disco rígido ou no servidor de rede ainda não alocado a qualquer ficheiro.

⁵³ Em alguns casos, os agentes de autoridade pública fazem a cópia de imagem no local; em outros, apreendem o *hardware* do computador e fazem a cópia da imagem fora do local, mas tal é admissível somente se constar da promoção quais os constrangimentos práticos e técnicos que requerem este procedimento, conforme é referido nas *Guidelines* 2009, p. 78.

⁵⁴ *Guidelines* 2009, p. 76.

⁵⁵ *United States v Ford*, 184 F.3d 566, 576 (6th Cir. 1999), citando casos, e *In re Grand Jury Investigation Concerning Solid State Devices, Inc.*, 130 F.3d 853, 857 (9th Cir. 1997). Também relevante: *United States v Fleet Management Ltd.*, 521 F. Supp. 2d 436, 443-44 (E.D. Pa. 2007).

⁵⁶ *Guidelines* 2009, p. 73.

⁵⁷ *Guidelines* 2009, pp. 80-81.

⁵⁸ Em língua original (*Fed. R. Crim. P. 41(c)*). Persons or Property Subject to Search or Seizure): «A warrant may be issued for any of the following: (1) evidence of a crime; (2) contraband, fruits of crime, or other items illegally possessed; (3) property designed for use, intended for use, or used in committing a crime; or (4) a person to be arrested or a person who is unlawfully restrained».

⁵⁹ As *Federal Rules of Criminal Procedure* entraram em vigor em 21 de março de 1946 e foram alteradas pela última vez em 1 de dezembro de 2019.

⁶⁰ *United States v Tamura*, 694 F.2d 591, 595 (9th Cir. 1982).

As restantes possibilidades remetem a pesquisa do equipamento para fora do local da apreensão (*off-site search*). Só a possibilidade cada vez menos utilizada de impressão de ficheiros específicos implica que toda a pesquisa seja realizada no local (*on-site search*). Cabe aqui destacar que não é indiferente que a pesquisa se realize no local ou se faça ou prossiga fora do local. A pesquisa no local tem um tempo limitado de duração, ao passo que a pesquisa fora do local pode durar o tempo que for necessário para ser completada. Tal demonstra, só por si, que a pesquisa fora do local é mais invasiva da privacidade do que a pesquisa no local. Acresce que a quantidade de informação guardada em formato digital pode ser de tal maneira vasta – na ordem dos gibabytes ou mesmo terabytes – que a devassa da privacidade numa pesquisa fora do local pode suplantar largamente a que ocorreria numa simples pesquisa e apreensão no local, de duração limitada a algumas horas ou dias. Daí que seja da máxima relevância garantir que quaisquer pesquisas informáticas a prosseguir fora do local tenham a devida justificação para poderem ser consideradas legítimas à luz das Diretrizes de 2009.

As Diretrizes de 2009 enfatizam a importância de os agentes da investigação criminal conceberem uma estratégia minuciosa antes de promoverem junto de um juiz a obtenção de um mandado de busca e apreensão em ambiente digital⁶¹. Por consequência, a garantia (*affidavit*) da causa provável que é necessária para a promoção do mandado junto de um juiz deve reportar quais os factos específicos que justificam a indispensabilidade de prossecução da pesquisa informática fora do local⁶². Independentemente de o mandado judicial permitir expressamente a pesquisa fora do local, se forem copiados elementos digitais para posterior pesquisa fora do local, o auto de busca e apreensão deve então pormenorizar os factos e as circunstâncias que impuseram um tal procedimento.

3.2. Pesquise antes de apreender ou...

Se os documentos armazenados em equipamentos informáticos estivessem em formato de papel, a seleção exigiria que os inspetores (*officers*) analisassem milhentos detalhes para determinar quais deles continham afinal informações que justificassem a sua apreensão ao abrigo do mandado judicial. Acresce que os inspetores poderiam precisar da orientação de um promotor de justiça (*prosecutor*) para fazer a seleção dos documentos relevantes. Essa orientação aumentaria o tempo necessário para rever os documentos e seleccionar aqueles que pudessem ser legitimamente apreendidos. Se tudo isso fosse feito no local, os referidos agentes da investigação criminal poderiam ter de ocupar as instalações durante várias horas ou mesmo dias⁶³.

A migração para o ciberespaço veio, sem dúvida, facilitar a vida às autoridades públicas, de tal sorte que a simples possibilidade de fazerem a pesquisa de documentos

⁶¹ *Guidelines* 2009: 61-86.

⁶² BERMAN, 2018: 53.

⁶³ BRENNER / FREDERIKSEN, 2002: 59.

relevantes em equipamentos informáticos em tempo real e no local da busca representa, só por si, um ganho de eficácia e de tempo. As Diretrizes de 2009 não desaconselham que se pesquise apenas uma parte do sistema informático, eventualmente através de palavras-chave ou frases específicas, respeitando os termos do mandado judicial. Tal só é viável, porém, se as autoridades públicas tiverem de antemão uma percepção rigorosa das evidências de que estão à procura, uma espécie de evidência conclusiva —vulgo, a «*smoking gun evidence*»— do delito sob investigação, que, por exemplo, lhes tenha sido revelada com exatidão por um denunciante. Nestes casos, o procedimento de utilização de palavras-chave é preferível para todos. Por um lado, é preferível para as autoridades públicas, pois só trazem consigo aquilo que for essencial para a investigação e evitam trazer documentação irrelevante, a qual acabaria por se transformar em lastro processual, potenciando incidentes e demoras. Por outro lado, é preferível para os visados, pois a devassa da sua privacidade é restringida ao estritamente necessário para a investigação em curso.

Do ponto de vista técnico, o curso de ação recomendável implica que se comece por preservar uma cópia forense do ambiente informático antes de se iniciar qualquer pesquisa, a fim de prevenir a contaminação ou destruição de evidências. As ferramentas forenses convencionais oferecem a capacidade de criar *back-ups* precisos e infalsificáveis, graças à atribuição de valores *hash*, que contenham todas as características do ambiente original, incluindo a informação sobre os ficheiros excluídos e o espaço no disco rígido ou no servidor de rede ainda não alocado a qualquer ficheiro⁶⁴.

3.3. Aprenda primeiro e depois logo se vê

Não sendo possível realizar a pesquisa de documentos em equipamentos informáticos em tempo real e no local da busca, então resta a alternativa entre usar palavras-chave (*key-words*) para copiar ficheiros individuais ou fazer uma cópia integral (*mirror-image copy*) do disco rígido ou do servidor de rede⁶⁵. Mas as Diretrizes de 2009 enfatizam que fazer uma cópia eletrónica de uma unidade inteira é bem diferente de fazer uma cópia eletrónica de ficheiros individuais. Esta última reduz não só o tempo necessário para a posterior seleção de possíveis evidências, mas também o risco de ultrapassagem do objeto do mandado.

As evidências geradas por computador (*computer-generated evidences*) podem ser recolhidas e custodiadas usando a automação. Os inspetores podem correr um programa para executar, por intermédio de palavras-chave, a apreensão de documentos em computadores independentes, servidores de rede ou armazenamento em nuvem. A utilização de palavras-chave automatizadas tem a vantagem de ser relativamente célere e cirúrgica. Porém, as palavras-chave são insensíveis ao contexto e ficam muito

⁶⁴ BRENNER / FREDERIKSEN, 2002: 63-65.

⁶⁵ BARTHOLOMEW, 2014: 1034-1036 e BERMAN, 2018: 92-93.

aquém da capacidade de discriminação típica de um investigador humano⁶⁶. Acresce que o emprego de palavras ou frases genéricas como palavras-chave pode ajudar a localizar evidências relevantes, mas produz um número elevado de falsos positivos (*false hits*). Os falsos positivos são documentos que contêm o termo procurado, mas não têm valor probatório e escapam ao objeto do mandado judicial⁶⁷.

3.4. A pesquisa externa

A questão agora é saber se o Quarto Aditamento à Constituição dos EUA autoriza a apreensão de evidências geradas por computador para pesquisa externa. Um tal procedimento implica necessariamente apreender documentos que não têm valor probatório e que estão para além do escopo do mandado⁶⁸. Tão-pouco dispensa a revisão de cada ficheiro (*file-by-file*) por um ou mais agentes de investigação criminal, o que significa que estes vão ter acesso a informação que, em princípio, lhes estaria vedada sob a autoridade do mandado judicial. A maneira de lidar com este problema, especialmente se estiver ameaçado o segredo profissional entre advogado e cliente (*attorney-client privilege*), poderá passar pela nomeação de uma equipa provisória de agentes de investigação criminal (*filter team* ou *taint team*) à qual se retira o acompanhamento posterior do caso ou até mesmo a nomeação pelo tribunal de um supervisor especial (*special master*)⁶⁹, que, por certo, oferece mais garantias de independência em relação ao poder executivo^{70/71}. Poderá ainda tornar-se necessária a intervenção de um juiz especial (*special Magistrate Judge*)⁷².

O Departamento de Justiça baseia a alegação de que as pesquisas externas são necessárias em duas premissas diferentes. A primeira é uma variante das exceções tradicionais à exigência de mandado judicial. As exceções podem ser justificadas pela necessidade de impedir a destruição de evidências essenciais. Este é, certamente, um argumento válido, desde que se demonstre que a destruição de evidências estaria, de facto, iminente⁷³. A segunda é a necessidade de envolver técnicos de informática (*computer experts*) na pesquisa para evitar a contaminação ou destruição de evidências essenciais, o que normalmente só é possível fazer com segurança e tranquilidade fora das instalações buscadas⁷⁴.

⁶⁶ MOSHIRNIA, 2010: 624-626.

⁶⁷ BRENNER / FREDERIKSEN, 2002: 60-62.

⁶⁸ BERMAN, 2018: 52.

⁶⁹ Veja-se o Manual do Procurador dos Estados Unidos (*United States Attorney's Manual – USAM*, section 9-13.420).

⁷⁰ Veja-se o recente caso *United States v Gallego*, No. 4:18- cr-01537, 2018 U.S. (Dist. Ariz.).

⁷¹ Assinalando, porém, as dificuldades postas pela criação destas equipas provisórias, designadamente relacionadas com a sua falta de acurácia, veja-se MANTEI, 2011: 1000.

⁷² BRENNER / FREDERIKSEN, 2002: 105.

⁷³ BRENNER / FREDERIKSEN, 2002: 68.

⁷⁴ BRENNER / FREDERIKSEN, 2002: 69.

Nas situações que envolvam a criação no local de uma cópia forense para subseqüente pesquisa externa, a promoção do mandado pelo titular da investigação criminal diante do juiz deve especificar os métodos de recolha, custódia e pesquisa que serão usadas e as precauções que serão tomadas para garantir que a pesquisa respeite o objeto do mandado judicial. O próprio mandado judicial deve indicar os delitos visados, revelando o máximo de pormenores e, se forem necessárias palavras-chave, listando os termos a utilizar para o efeito. A autorização judicial pode estar contida no mandado original ou em mandado complementar (*supplemental warrant*). Os agentes de investigação criminal promovem o mandado complementar quando, após iniciarem a execução do mandado original, chegarem à conclusão de que uma pesquisa no local simplesmente não é viável⁷⁵.

3.5. A doutrina jurisprudencial da visibilidade imediata

À luz do direito jurisprudencial norte-americano, a recolha de evidências geradas por computador rege-se pelas normas aplicáveis às tradicionais diligências de busca e apreensão, nos termos do Quarto Aditamento à Constituição dos EUA⁷⁶. Mas a proibição constitucional das buscas e apreensões injustificadas implica uma reelaboração de conceitos de cada vez que for aplicada ao mandado de busca digital (*digital search warrant*)⁷⁷. Essa reelaboração é sumamente necessária quanto à questão do destino a dar aos conhecimentos fortuitos que ocorram em ambiente digital.

A doutrina jurisprudencial da visibilidade imediata (*plain view doctrine*) é uma exceção ao imperativo constitucional de mandado judicial para a realização de buscas e apreensões⁷⁸. A visibilidade imediata autoriza que sejam utilizados como prova de um delito quaisquer objetos apreendidos por um agente de autoridade que tenha atuado sem ou para além do mandado judicial de busca e apreensão, se forem atendidas as seguintes três condições: «(1) a evidência deve estar imediatamente à vista; (2) o agente de autoridade tem de possuir uma razão justificativa anterior para se encontrar no local a partir do qual consegue visualizar imediatamente a evidência; (3) a evidência “por si mesma ou juntamente com factos conhecidos do agente de autoridade no momento da apreensão, [deve fornecer] uma probabilidade razoável para crer que exista uma conexão entre a evidência e alguma atividade criminosa”»⁷⁹.

⁷⁵ BRENNER / FREDERIKSEN, 2002: 102-104.

⁷⁶ CLANCY, 2005: 208 e KERR, 2005: 532.

⁷⁷ BRENNER / FREDERIKSEN, 2002: 41, CHANG, 2007: 33-34, WARD, 2011: 1170-1179, BARTHOLOMEW, 2014: 1033, SILVA RAMALHO, 2014: 93-146, BERMAN, 2018: 53 e WITTLER CONTARDO, 2020: 305-306.

⁷⁸ Veja-se os casos *Coolidge v New Hampshire*, 403 U.S. 443, 465 (1971) e *Horton v California*, 496 U.S. 128, 134 (1990).

⁷⁹ Em língua original (*State v Guy*, 172 Wis.2d 86, 101-02, 492 N.W.2d 311 (1992), citando *State v Washington*, 134 Wis.2d 108, 121, 396 N.W.2d 156 (1986), que citava *Bies v State*, 76 Wis.2d 457, 464, 251 N.W.2d 461 (1977)): «(1) *the evidence must be in plain view*; (2) *the officer must have a*

A doutrina jurisprudencial da visibilidade imediata baseia-se na experiência empírica da percepção visual no mundo físico. No mundo cibernético, porém, não há, em princípio, analogia com a visão no mundo físico. Um objeto pode ser imediatamente avistado no mundo físico, ao passo que um ficheiro informático só pode ser visto se for aberto. Aquilo que se vê imediatamente é apenas o nome do ficheiro, que pode nem ser revelador do respetivo conteúdo. Esta dificuldade tem dado azo a uma rica casuística no direito jurisprudencial norte-americano⁸⁰. De seguida, oferecemos uma imagem, a voo de pássaro, da referida casuística através de dois casos paradigmáticos.

3.6. O caso Estados Unidos vs. Carey (1999)

O caso *United States v Carey* (1999) é paradigmático das dificuldades de aplicação da doutrina jurisprudencial da visibilidade imediata às evidências geradas por computador⁸¹.

Patrick J. Carey foi acusado de posse de um disco rígido de computador que continha três ou mais imagens de pornografia infantil resultantes de materiais adquiridos no mercado ilícito. Após um reconhecimento parcial de culpabilidade (*conditional plea of guilty*), o réu recorreu da decisão do tribunal distrital que indeferira a sua moção para exclusão do material retirado do seu computador, alegando que o mesmo fora apreendido como resultado de uma busca geral sem mandado. O Tribunal de Apelação do Décimo Circuito dos Estados Unidos da América (*United States Court of Appeals, Tenth Circuit*) decidiu, em 14 de abril de 1999, que a moção do réu devia ser deferida e reverteu a decisão do tribunal recorrido⁸².

Carey estava sob investigação, havia já algum tempo, por possível venda e posse de cocaína. A polícia obteve um mandado para detê-lo em sua residência. Durante a execução do mandado de detenção, os agentes de autoridade observaram, bem à vista no apartamento do visado, um dispositivo normalmente utilizado para fumar marijuana (*bong*) e o que parecia ser droga. Alertado por esses itens, um dos agentes solicitou a Carey que consentisse na busca domiciliária. O agente disse que conseguiria um mandado de busca se Carey recusasse a permissão. Após uma discussão considerável com o agente, Carey consentiu verbalmente e mais tarde assinou um consentimento formal na esquadra de polícia. Receando que os agentes «destruissem» o seu apartamento durante a busca, Carey deu-lhes indicações sobre onde poderiam encontrar evidências da venda e posse de drogas. Os agentes também apreenderam dois computadores, pois acreditavam que poderiam conter evidências

prior justification for being in the position from which [he or] she discovers the evidence in 'plain view'; (3) the evidence seized 'in itself or itself with facts known to the officer at the time of the seizure, [must provide] probable cause to believe there is a connection between the evidence and criminal activity'».

⁸⁰ BRENNER / FREDERIKSEN, 2002: 94, MOSHIRNIA, 2010: 612, BERMAN, 2018: 59-60 e WITTLER CARTARDO, 2020: 302-306.

⁸¹ *United States v Carey*, 172 F.3d 1268 (10th Cir. 1999).

⁸² *United States v Carey* (1999), *cit.*, § I.

de tráfico de drogas. Os computadores foram levados para a esquadra da polícia e os agentes obtiveram um mandado judicial que lhes permitia pesquisar ficheiros nos computadores em busca de nomes, números de telefone, recibos de contabilidade, endereços e outras evidências documentais relativas à venda e distribuição de substâncias proibidas. Nos diretórios estavam vários ficheiros com títulos sexualmente sugestivos, em formato JPG. O inspetor Lewis inseriu os discos em outro computador e começou a pesquisar os ficheiros copiados dos computadores de Carey. O seu método consistia em inserir palavras-chave como «dinheiro, contas, pessoas, etc.» na função de pesquisa para encontrar ficheiros baseados em texto (*text-based*) contendo essas palavras. Esta pesquisa não produziu qualquer correspondência relativa a drogas (*related to drugs*). O inspetor Lewis continuou a explorar os diretórios e encontrou alguns ficheiros com os quais, disse, «não estava familiarizado». Incapaz de visualizá-los no computador que estava usando, copiou-os para um disco externo e colocou-os em outro computador. Então foi «imediatamente» capaz de visualizar o que mais tarde descreveria como um ficheiro em formato JPG (*JPG file*). Ao abrir este ficheiro, descobriu que continha pornografia infantil. O inspetor Lewis abriu aproximadamente duzentos e quarenta e quatro ficheiros JPG ou de imagem⁸³.

Carey entrou com uma moção requerendo a exclusão dos ficheiros que continham pornografia infantil. Durante a audiência judicial sobre a moção, o inspetor Lewis afirmou que, embora a descoberta dos ficheiros JPG tenha sido completamente fortuita, logo que vira a primeira foto contendo pornografia infantil desenvolveu uma causa provável para acreditar que o mesmo tipo de material estaria presente nos restantes ficheiros de imagem. Quando perguntado por que razão não promovera, afinal, um mandado complementar para pesquisa dos demais ficheiros de imagem, o inspetor Lewis respondeu que a questão fora equacionada e o seu superior hierárquico, supostamente, tratara disso através do gabinete do promotor de justiça local (*county attorney's office*). Não foi, porém, obtido qualquer mandado complementar, mas o inspetor Lewis, mesmo assim, continuou a explorar os ficheiros de imagem. Na opinião do governo, «uma pesquisa em computador como a realizada neste caso equivale a procurar documentos em um arquivo físico, de acordo com um mandado de busca e apreensão válido, acabando por se encontrar pornografia infantil»⁸⁴.

Na sua decisão, o Tribunal de Apelação recordou que o Supremo Tribunal (*Supreme Court*) fixara o entendimento de que «a doutrina jurisprudencial da visibilidade imediata não pode ser usada para estender uma pesquisa exploratória geral de um objeto para outro até que algo de incriminador finalmente apareça» (*Coolidge v New Hampshire*, 403 U.S. at 466, 91 S.Ct. 2022)⁸⁵. O mandado obtido para um propó-

⁸³ *United States v Carey* (1999), *cit.*, § I.

⁸⁴ Em língua original (*United States v Carey* (1999), § II): «*a computer search such as the one undertaken in this case is tantamount to looking for documents in a file cabinet, pursuant to a valid search warrant, and instead finding child pornography*».

⁸⁵ Em língua original (*United States v Carey* (1999), § II): «*the plain view doctrine may not be used to extend a general exploratory search from one object to another until something incriminating at last emerges*».

sito específico de pesquisa nos computadores do visado permitiria apenas a busca de ficheiros por nomes, números de telefone, livros, recibos, endereços e outras provas documentais relativas à venda e distribuição de substâncias proibidas. O escopo da busca foi, portanto, circunscrito a evidências pertencentes ao tráfico de drogas. Lewis havia abandonado temporariamente essa pesquisa para procurar mais pornografia infantil e apenas voltou a pesquisar documentos relativos a drogas depois de conduzir uma pesquisa de cinco horas nos ficheiros de pornografia infantil⁸⁶.

O Tribunal de Apelação não aceitou a alegação do governo de que a imagem sexualmente sugestiva que repentinamente surgiu à «vista de todos» (*plain view*) na tela do computador tornara os ficheiros acessíveis como se fosse um «jogo justo» (*fair game*) baseado numa busca consensual, desde logo porque o consentimento do visado referia-se apenas à busca no respetivo apartamento e nada mais. A consideração crítica a este respeito é que os agentes nunca anunciaram, antes que o visado desse o seu consentimento, que estavam investigando uma suspeita de posse de pornografia infantil⁸⁷.

O Tribunal de Apelação concluiu, por unanimidade (com voto de concordância separado do juiz Porfilio), que o inspetor Lewis excedera o escopo do mandado no caso em apreço. A apreensão de evidências em que se baseara a acusação foi consequência de uma busca geral inconstitucional e, por conseguinte, o tribunal distrital cometera um erro ao recusar-se a excluir tais evidências⁸⁸.

O caso *United States v Carey* (1999) apresenta uma característica singular que ajuda a explicar a decisão do Tribunal de Apelação, a saber: o mandado judicial autorizava a pesquisa nos computadores do visado através de palavras-chave, o que tornava os ficheiros de imagem inacessíveis ao tipo de pesquisa autorizado, a menos que os títulos desses ficheiros correspondessem a alguma das palavras-chave, o que não era o caso. Uma correta pesquisa através de palavras-chave não consentiria sequer que o inspetor Lewis vasculhasse livremente os diretórios dos computadores do visado⁸⁹. Neste contexto, é compreensível que a doutrina jurisprudencial da visibilidade imediata não tenha aplicação, tanto mais que o acesso ao conteúdo dos ficheiros de imagem só ocorreu após sucessivas transferências do material informático apreendido para distintos computadores de trabalho. Que a doutrina jurisprudencial da visibilidade imediata não tenha aplicação neste contexto não significa, porém, que nunca tenha aplicação às evidências geradas por computador.

⁸⁶ *United States v Carey* (1999), *cit.*, § II.

⁸⁷ *United States v Carey* (1999), *cit.*, § II.

⁸⁸ *United States v Carey* (1999), *cit.*, § III.

⁸⁹ MOSHIRNIA, 2010: 623-624 e MANTEL, 2011: 993-996.

3.7. O caso *State vs. Schroeder* (2000)

O caso *State v Schroeder* (2000)⁹⁰ é, por sua vez, paradigmático da possibilidade de aplicação da doutrina jurisprudencial da visibilidade imediata às evidências geradas por computador.

Keith Schroeder recorreu da sentença que o condenara por dezoito acusações de posse de pornografia infantil, assim como da decisão de recusa de atenuação do regime de cumprimento da pena. Entre outras, suscitou a questão da violação do Quarto Aditamento à Constituição dos EUA durante a pesquisa ao seu computador pelo laboratório de polícia científica. Durante a pesquisa foram descobertos fortuitamente ficheiros de pornografia infantil, embora o mandado judicial fosse destinado à pesquisa de evidências relativas a crimes de assédio online (*harassment*) e conduta desordeira (*disorderly conduct*). Logo que deparou com os primeiros ficheiros pornográficos, o analista interrompeu a pesquisa e aguardou a obtenção de novo mandado judicial de busca digital, o qual viria, de facto, a ser emitido. Mas Schroeder manteve a alegação de que faltava um mandado judicial de busca e apreensão válido e alegou que as imagens de pornografia infantil deveriam ser eliminadas como fruto da árvore venenosa (*fruit of the poisonous tree*). O Tribunal de Apelação de Wisconsin (*Court of Appeals of Wisconsin*), em 24 de maio de 2000, rejeitou este (e os demais) argumento(s) trazido(s) pelo recurso⁹¹.

Tudo começou com a execução do primeiro mandado judicial de busca e apreensão. O delegado de polícia (*County Sheriff's Sergeant*) Harry Sokel procurou instalar no computador de Schroeder um programa para preservar (*to freeze*) o sistema de modo a que os respetivos conteúdos não pudessem ser alterados. Enquanto procurava os equipamentos periféricos do computador, Sokel deparou com um disco compacto que parecia pela respetiva capa de invólucro conter pornografia. Sokel perguntou a Schroeder se guardava pornografia adulta no computador e ele respondeu afirmativamente. Sokel perguntou ainda a Schroeder se guardava pornografia infantil no computador, ao que este retorquiu não saber o que era tal coisa. Sokel acrescentou que se referia a fotografias de crianças despidas com idade inferior a dezasseis anos. Schroeder respondeu que deveria haver algumas⁹².

O computador foi enviado ao laboratório de polícia científica para análise. Sokel disse ao inspetor (*investigator*) que levou o computador para o laboratório, Steve Malchow, que poderia haver pornografia infantil no computador. Malchow, por sua vez, disse a Marty Koch, o analista forense (*crime lab analyst*), que se encontrasse qualquer pornografia infantil no computador deveria interromper a pesquisa e ligar para Malchow. Em sua busca de evidências de assédio online, Koch encontrou, de facto, algumas fotos pornográficas de crianças⁹³.

⁹⁰ *State v Schroeder*, 613N.W.2d911 (Wis. App. 2000).

⁹¹ *State v Schroeder* (2000), *cit.*, § 1.

⁹² *State v Schroeder* (2000), *cit.*, § 2.

⁹³ *State v Schroeder* (2000), *cit.*, § 4.

Em recurso, Schroeder alegou que Koch abrisse indevidamente alguns ficheiros cujos nomes sugeriam conteúdos de pornografia infantil. Segundo Schroeder: «[e]ste passo adicional de abrir e ver a pasta para verificar se continha pornografia infantil torna a pesquisa ilegal»⁹⁴. Schroeder parece, portanto, desafiar a aplicação da doutrina jurisprudencial da visibilidade imediata, ao referir que os ficheiros não estavam à vista⁹⁵.

O Tribunal de Apelação, porém, não aceitou a alegação de Schroeder. Koch testemunhara que, ao pesquisar um computador, costumava vasculhar sistematicamente e abrir os ficheiros independentemente dos seus nomes. Tal faz sentido, pois um utilizador é livre de nomear um ficheiro com qualquer nome. Se Koch limitasse sua pesquisa a ficheiros cujos nomes sugerissem o tipo de evidência que ele buscava, seria muito fácil para os visados esconderem as evidências no computador: – Nomeie o seu ficheiro de pornografia infantil como «1986.taxreturn» e ninguém poderá abri-lo! Ao abrir sistematicamente todos os ficheiros criados pelo utilizador, Koch deparou com um que continha imagens que ele considerava serem de pornografia infantil. Neste momento, Koch interrompeu a pesquisa e chamou Malchow. Não retomou a pesquisa, nem encontrou o resto das imagens de pornografia infantil antes de obter o segundo mandado de busca digital. Assim, a sua descoberta inicial de pornografia infantil ocorreu quando ele abriu um ficheiro e viu a foto de uma criança nua aparecendo na tela. Estava bem à vista. Segundo o Tribunal de Apelação: «[i]sto não era diferente de um investigador abrir uma gaveta enquanto procurava por drogas e ver a foto de uma criança nua em cima de uma pilha de meias»⁹⁶. O primeiro requisito do teste da visibilidade imediata está satisfeito. Com relação ao segundo e terceiro requisitos, é indiscutível que Koch tinha um mandado para pesquisar o computador em busca de evidências de assédio online e que a primeira imagem que Koch encontrou poderia ser razoavelmente vista, de caras, como pornografia infantil. Por consequência, a doutrina jurisprudencial da visibilidade imediata, segundo o Tribunal de Apelação, aplicava-se ao caso⁹⁷.

A situação é diferente da ocorrida no caso *United States v Carey* (1999), onde o tribunal de recurso considerara a pesquisa irrazoável a partir do momento em que o inspetor abandonara a sua procura original por evidências de tráfico de drogas e passara a procurar pornografia infantil, após a descoberta inadvertida da primeira imagem de pornografia infantil. A diferença significativa é que, em *Carey*, o investigador abandonara a sua pesquisa original e começara a procurar apenas mais pornografia infantil, sem obter um novo mandado de busca digital. Assim, nenhuma das imagens pornográficas descobertas posteriormente caiu sob a alçada da doutrina jurisper-

⁹⁴ Em língua original (*State v Schroeder* (2000), § 13): «*This additional step of opening and reviewing the folder to verify it contained child porn makes the search illegal*».

⁹⁵ *State v Schroeder* (2000), *cit.*, § 13.

⁹⁶ Em língua original (*State v Schroeder* (2000), § 14): «*This was no different than an investigator opening a drawer while searching for drugs and seeing a nude picture of a child on top of a pile of socks*».

⁹⁷ *State v Schroeder* (2000), *cit.*, § 14.

dencial da visibilidade imediata. No presente caso, pelo contrário, o analista parou de pesquisar e obteve um segundo mandado. A pesquisa por pornografia infantil só prosseguiu depois de autorizada pelo segundo mandado⁹⁸.

Em suma, a doutrina jurisprudencial da visibilidade imediata justifica apenas apreensões, não buscas⁹⁹. O investigador criminal deve, por isso, ignorar as descobertas fortuitas? A resposta é negativa. Se tiver uma causa provável para acreditar que os ficheiros que sobram contêm evidências de atividade criminosa não abrangida pelo mandado de busca digital, então deverá solicitar um mandado de busca digital complementar que o autorize a prosseguir a pesquisa.

4. CONCLUSÕES

O artigo 8.º da Convenção Europeia dos Direitos Humanos trata do respeito pela privacidade. O Quarto Aditamento à Constituição dos EUA trata igualmente do respeito pela privacidade.

A migração da vida contemporânea para o ciberespaço deve preservar o respeito pela privacidade, inclusivamente ao nível do processo penal. Para assegurar este desiderato, o direito em ação deve definir com mais rigor os limites da cópia de dados eletrónicos e as restrições impostas à análise externa do acervo recolhido.

Entre os aspetos críticos da análise externa de dados eletrónicos, avulta a questão do procedimento a adotar pelo investigador criminal diante dos conhecimentos fortuitos. A descoberta inadvertida de evidências produzidas por computador relacionadas com uma atividade criminosa não abrangida pelo mandado de busca digital primitivo obriga o investigador criminal a interromper a visualização dos ficheiros eletrónicos e a promover a obtenção de um mandado de busca digital complementar com base em indícios razoáveis da prática do crime detetado. A falta de mandado de busca digital complementar determina a exclusão das evidências porventura descobertas através da indevida continuação da análise externa do acervo recolhido.

Consoante a vastidão do acervo eletrónico recolhido, a análise externa pode implicar a necessidade de nomeação *ad hoc* de uma equipa de promotores de justiça, inspetores de polícia criminal e analistas informáticos cuja função se limite à triagem das evidências relevantes para o processo criminal, em curso ou a instaurar, sinalizando os documentos eletrónicos cobertos por privilégios de sigilo profissional ou segredo de negócio e eliminando os documentos eletrónicos irrelevantes. A equipa *ad hoc* deve ser afastada da prossecução da investigação criminal logo que termine a tarefa para que foi designada, dado que ficou contaminada pelo conhecimento de

⁹⁸ *State v Schroeder* (2000), *cit.*, § 16.

⁹⁹ Mas cada nova visualização de dados na tela do computador é uma busca, não uma apreensão, segundo KERR, 2005: 534.

informação que, por definição, não poderia conhecer sob a autoridade do(s) mandado(s) de busca digital.

Não basta a nomeação *ad hoc* de uma equipa de investigação criminal, à qual faltará sempre a suficiente imparcialidade para assegurar a proteção da privacidade do visado. A única maneira eficaz de mitigar, na prática, a manipulação abusiva da doutrina da visibilidade imediata quando os investigadores criminais tenham de analisar grandes acervos de documentos eletrónicos é através da intervenção de um juiz especial (dependendo dos ordenamentos jurídicos que o prevejam, essa é função do juiz de instrução). O juiz especial deve examinar todas as evidências produzidas por computador e confirmar a respetiva correspondência com o mandado de busca digital.

A jurisprudência do Tribunal Europeu dos Direitos Humanos caracteriza-se por alguma ineficácia na criação de remédios para a violação da privacidade no processo penal, designadamente no tocante à cópia de dados eletrónicos e à análise externa do acervo recolhido, desde logo porque não comina a exclusão das evidências produzidas por computador que tenham sido obtidas ilicitamente, o que deveria ser o caso, à luz do princípio do processo equitativo.

Os ordenamentos jurídicos nacionais europeus caracterizam-se por um regime legal geral e abstrato em matéria de prova digital. A distância entre as normas legais habilitantes da cópia de dados eletrónicos e respetiva análise externa, por um lado, e a atuação no terreno das autoridades de investigação criminal, por outro, não é colmatada pela existência de diretrizes mandatórias para tais autoridades, ademais divulgadas publicamente, que contenham a descrição detalhada das boas práticas. Naturalmente, a mera existência de manuais de procedimentos internos, indicativos e secretos não é garantia suficiente das boas práticas.

O conhecimento das diretrizes e do direito jurisprudencial norte-americano representa um contributo valioso para o aprofundamento da jurisprudência de Estrasburgo, na sua dupla função decisória e nomofilática, assim como para o aperfeiçoamento dos ordenamentos jurídicos nacionais europeus ao nível legislativo e ao nível da prática jurisprudencial.

Alguns dirão que as conclusões deste artigo são contraproducentes porque não anunciam formas de efficientização da recolha das evidências produzidas por computador em ordem ao deslindamento dos factos puníveis investigados, antes estabelecem limites à atividade de investigação criminal no ciberespaço. A crítica é pertinente, mas a sociedade democrática tem de pagar um preço para que a vida dos indivíduos e das empresas não se torne um livro aberto à absoluta devassa, nem mesmo quando as autoridades públicas andem legitimamente à procura das evidências dos crimes investigados.

5. BIBLIOGRAFIA

- AMBOS, Kai, (2010: *Beweisverwertungsverbote: Grundlagen und Kasuistik – internationale Bezüge – ausgewählte Probleme*, Berlin: Duncker & Humblot.
- BACHMAIER WINTER, Lorena, THAMAN, Stephen C., 2020: «A Comparative View of the Right to Counsel and the Protection of Attorney-Client Communications», in: Lorena Bachmaier Winter, Stephen C. Thaman e Veronica Lynn (eds.), *The Right to Counsel and the Protection of Attorney-Client Privilege in Criminal Proceedings – A Comparative View*, Cham: Springer, pp. 7-73.
- BARTHOLOMEW, Paige, 2014: «Seize First, Search Later: The Hunt for Digital Evidence», *30 Touro Law Review* 4, pp. 1027-1052.
- BERMAN, Emily, 2018: «Digital Searches, the Fourth Amendment, and the Magistrates' Revolt», *68 Emory Law Journal*, pp. 49-94.
- BRENNER, S. W., FREDERIKSEN, B. A., 2002: «Computer Searches and Seizures: Some Unresolved Issues», *8 Mich. Telecomm. Tech. L. Rev.* 39, pp. 39-114.
- CHANG, RayMing, 2007: «Why the Plain View Doctrine Should Not Apply to Digital Evidence», *Suffolk Journal of Trial and Appellate Advocacy* 12, pp. 31-67.
- CLANCY, Thomas K., 2005: «The Fourth Amendment Aspects of Computer Searches and Seizures: A Perspective and a Primer», *Mississippi Law Journal* 75, pp. 193-286.
- COSTA RAMOS, Vânia, 2017: «Os problemas em matéria de proibições de prova—uma dimensão internacional—regras de exclusão da prova obtida em violação da Convenção Europeia dos Direitos Humanos na jurisprudência do Tribunal Europeu dos Direitos Humanos», in: J. FARIA COSTA et al. (eds.), *Estudos em Homenagem ao Prof. Doutor Manuel da Costa Andrade*, vol. II (Direito Penal, Direito Processual Penal), Coimbra: Universidade de Coimbra, pp. 740-773.
- COSTA RAMOS, Vânia, PINTO DE ABREU, Carlos, CORDEIRO, João Valente, 2020: «Confidentiality of Correspondence with Counsel as a Requirement of a Fair Trial in Portugal», in: Lorena Bachmaier Winter, Stephen C. Thaman e Veronica Lynn (eds.), *The Right to Counsel and the Protection of Attorney-Client Privilege in Criminal Proceedings – A Comparative View*, Cham: Springer, pp. 235-271.
- DE HERT, Paul, GUTWIRTH, Serge, 2009: «Data Protection in the Case Law of Strasbourg and Luxembourg: Constitutionalisation in Action», in: Serge Gutwirth, Yves Pouillet, Paul de Hert, Cécile de Terwangne e Sjaak Nouwt (eds.), *Reinventing Data Protection?* Brussel / Namur / Utrecht: Springer, p. 3-44.
- JAHN, Jannika, 2014: «Ruling (In)directly through Individual Measures? Effect and Legitimacy of the ECtHR's New Remedial Power», *ZaöRV* 74, pp. 1-39.
- KAMISAR, Yale, 2003: «In Defense of the Search and Seizure Exclusionary Rule», *26 Harv. J. L. & Pub. Pol'y* 119, pp. 119-140.
- KERR, Orin S., 2005: «Searches and Seizures in a Digital World», *119 Harvard Law Review* 2, pp. 531-569.
- KOSTORIS, Roberto E., 2014: «Diritto europeo e giustizia penale», in: Roberto E. Kostoris (ed.), *Manuale di Procedura Penale Europea*, Milano: Giuffrè, pp. 1-62.
- LAFAVE, Wayne R., 2004: *Search and Seizure – A Treatise on the Fourth Amendment*, vol. 1 (Sections 1.1 through 2.7: The Exclusionary Rule and Other Remedies & Protected Areas and Interests), 4.^a ed., New York: Thomson Reuters, (1.^a ed., 1978).
- , 2011-2012: *Search and Seizure – A Treatise on the Fourth Amendment*, vol. 1 (Sections 1.1 through 2.7), 4.^a ed., New York: Thomson Reuters, Pocket Part (1.^a ed., 1978).
- MANTEI, Corey J., 2011: «Pornography and Privacy in Plain View: Applying the Plain View Doctrine to Computer Searches», *53 Arizona Law Review*, pp. 984-1012.
- MOSHIRNIA, Andrew Vahid, 2010: «Separating Hard Fact from Hard Drive: A Solution for Plain View Doctrine in the Digital Domain», *23 Harvard Journal of Law & Technology* 2, pp. 609-634.
- SILVA RAMALHO, David, 2017: *Métodos Ocultos de Investigação Criminal em Ambiente Digital*, Coimbra: Almedina.

- SOSA MENDES, Paulo de, 2020: «O princípio do processo equitativo na jurisprudência do TEDH», in: Anabela Antunes *et al.* (eds.), *Estudos de Homenagem ao Professor Doutor Germano Marques da Silva*, vol. II, Lisboa: Universidade Católica Editora, pp. 2325-2339.
- WARD, Kate Brueggemann, 2011: «The Plain (or Not So Plain) View Doctrine: Applying the Plain View Doctrine to Digital Seizures», *University of Cincinnati Law Review* 79, pp. 1163-1187.
- WITTLER CONTARDO, Ricardo, 2020: «Apreensão de correio eletrónico em Portugal: Presente e futuro de uma questão de “manifesta simplicidade”», in: Paulo de Sousa Mendes e Rui Soares Pereira (eds.), *Novos Desafios da Prova Penal*, Coimbra: Almedina, pp. 277-313.

JURISPRUDÊNCIA DO TRIBUNAL EUROPEU DOS DIREITOS HUMANOS

- Leander v Sweden* (queixa n.º 9248/81), de 26 de março de 1987.
- Amann v Switzerland* (queixa n.º 27798/95), de 16 de fevereiro de 2000.
- Rotaru v Romania* (queixa n.º 28341/95), de 4 de maio de 2000.
- Société Colas Est and other v France* (queixa n.º 37971/97), de 16 de abril de 2002.
- Wieser and Bicos Beteiligungen GmbH v Austria* (queixa n.º 74336/01), de 16 de janeiro de 2007.
- Copland v United Kingdom* (queixa n.º 62617/00), de 3 de abril de 2007.
- Robathin v Austria* (queixa n.º 30457/06), de 3 de julho de 2012.
- Bernh Larsen Holding As v Norway* (queixa n.º 24117/08), de 14 de março de 2013.
- Sérvulo & Associados – Sociedade de Advogados, Rl v Portugal* (queixa n.º 27013/10), de 3 de setembro de 2015.
- Trabajo Rueda v Spain* (queixa n.º 32600/12), de 30 de maio de 2017.
- Ivashchenko v Russia* (queixa n.º 61064/10), de 13 de maio de 2018.

DIRETRIZES E JURISPRUDÊNCIA DOS ESTADOS UNIDOS DA AMÉRICA

- Federal Rules of Criminal Procedure* (1946, 2019).
- Federal Guidelines for Searching and Seizing Computers* (1994) e *Supplements* (1997, 1999).
- Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations* (2001, 2009).
- Weeks v United States*, 232 U.S. 383, 34 S.Ct. 341, 58 L.Ed. 652 (1914).
- Mapp v Ohio*, 367 U.S. 643 (1961).
- Coolidge v New Hampshire*, 403 U.S. 443 (1971).
- Bies v State*, 76 Wis.2d 457, 464, 251 N.W.2d 461 (1977).
- State v Washington*, 134 Wis.2d 108, 121, 396 N.W.2d 156 (1986).
- Horton v California*, 496 U.S. 128 (1990).
- State v Guy*, 172 Wis. 2d 86, 492 N.W.2d 311 (1992).
- United States v Carey*, 172 F.3d 1268 (10th Cir. 1999).
- State v Schroeder*, 613 N.W.2d 911 (Wis. App. 2000).

CONJETURAS
Y REFUTACIONES

NATURALIZED EPISTEMOLOGY AND THE LAW OF EVIDENCE REVISITED

Ronald J. Allen*

ABSTRACT: This article revisits «Naturalized Epistemology and the Law of Evidence», published twenty years ago. The evolution of the relative plausibility theory of juridical proof is offered as evidence of the advantage of a naturalized approach to the study of the field and law of evidence. Various alternative explanations of juridical proof from other disciplines are examined and their shortcomings described. These competing explanations are similar in their reductive, a priori approaches that are at odds with an empirically oriented naturalized approach. The shortcomings of the various a priori approaches are driven by their common methodological commitments to employing weird hypotheticals to engage in intuition mining about the American legal systems that in turn ignore important aspects of the actual legal systems and consistently make impossible epistemological demands. As a result, both their descriptions of and prescriptions for American legal systems are implausible, unlike the relative plausibility theory.

KEYWORDS: naturalized epistemology, juridical proof, jurisprudence, probability, relative plausibility

RECOMMENDED CITATION: ALLEN, R. J., 2020: «Naturalized Epistemology and the Law of Evidence Revisited», in *Quaestio facti*, 2: 253-284. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22446

* John Henry Wigmore Professor, Northwestern University Pritzker School of Law, President, International Association of Evidence Science; Fellow, The Forensic Science Institute, China University of Political Science and Law. I am indebted to William Lawrence for excellent research assistance and to the Julius Rosenthal Foundation Fund for supporting this research. I am also indebted to Marcello Di Bello for helpful comments on a draft of this article, and to the Visiting Scholars and SJD students studying with me this year at Northwestern: from Brazil Vitor de Paula Ramos and from China Zhuhao Wang, Siyun Liao, Kaixin Zheng, and Xiaobiao Xiong. I am particularly indebted to Regina M. Hurley, a Ph.D. student at Northwestern University for very helpful discussions on various aspects of this paper, although all of the views expressed herein are my own.

Twenty years ago, Brian Leiter and I published an article examining the relationship between naturalized epistemology and investigations into the field and law of evidence. In brief, our argument was that the naturalized turn in philosophy inspired by Quine and others contained valuable lessons for scholars interested in the field of evidence and the various rules for trials that emerge from that field. «Naturalized epistemology» can mean different things to different people. The way we used the phrase was, first, to denote the difference between a priori conceptual/normative analysis on the one hand and on the other empirical inquiry into the actual state of some phenomena (in our case, the law of evidence and the nature of litigation), and second, to emphasize that, no matter what else, «should implies can». For conceptual/normative work, and any other kind for that matter, to have value for understanding or theorizing about the objects under inquiry, it must be responsive to what is in fact possible given the human condition. We then employed that methodological approach to demonstrate that an emerging theory of juridical proof—the relative plausibility theory—was systematically superior to the then reigning probabilistic conception¹.

Much has happened since then. On the theoretical side, philosophical disputations over «naturalized epistemology» continue, as has the implications of that debate for law in general and jurisprudence in particular². On the applied side, there is good reason to believe that relative plausibility has supplanted probabilism as the dominant paradigm for understanding juridical proof³. This has been aided by the shoring up of the foundations of relative plausibility that began in the original Allen/Leiter article through incorporation of aspects of inference to the best explanation informed by the reality of the American trial process and the manner in which people reason in real life⁴. Together these developments are evidence of the significance of naturalized epistemology for the field of evidence.

Perhaps in part responsive to these developments, the editors of *Quaestio Facti* suggested that I revisit the implications of naturalized epistemology for the field of evidence in the form of a contribution to the journal's «Conjectures and Refutations» section, and this article is the result. Before turning to substance, a clarification is in order concerning the meaning of «naturalized epistemology», as I would express it in somewhat different terms (but not obviously with significant change of meaning). As I am using the phrase, it refers to inquiries—analytical or empirical—into how the human mind engages with its environment and the implications of that form of engagement for Western legal systems—in particular the legal systems within the United States. More particular issues include how such legal systems structure dispute

¹ ALLEN and LEITER, 2001.

² See the symposium on Brian Leiter's book in *Law and Philosophy*, 2011, 30: 377-516.. Also see LEITER, 2011, 2020.

³ See the symposium on Relative Plausibility and Its Critics, *International Journal of Evidence and Proof*, 2019, 23 (1-2): 1-217.

⁴ See, e.g. PARDO and ALLEN, 2008.

resolution, the relationship between that structure and the field of evidence with its resultant rules, and how all of this is influenced by cognitive capacities.

To be clear, the objective here is not to canvas such a sprawling field but instead provide a relatively brief but hopefully provocative essay to stimulate «refutations», and in doing so to stress two things. First, I will briefly review the successes of the naturalized approach. Second, in an admittedly solipsistic vein, I will lament some of the shortcomings of alternative approaches to theorizing about the field of evidence if the objective is to understand the actual state of the subject of the inquiry—to help us jurists in other words. This last point is critical. Different disciplines do different things, are motivated differently, and have differing objectives. For whatever reasons, a cottage industry focusing on juridical proof has arisen from various disciplines. Using what are taken as examples from the legal arena and subjecting them to scrutiny from the perspective of some other discipline may very well be valuable to that discipline, a matter beyond my competence. However, the utility of such efforts for understanding juridical proof may be another matter, and it often is for three reasons. The first is the weird hypothetical problem. Hypotheticals are posited that are supposed to reflect some aspect of the legal system or the field of evidence, but do not, and then subjected to scrutiny. Because such hypotheticals do not reflect critical aspects of any existing legal system, the scrutiny they are subject to yields little of value for understanding of or prescribing for American legal systems. Second, and relatedly, some theorizing neglects critical aspects of the actual state of affairs in the object being theorized about. Neither of these necessarily violates the «should implies can» canon, but tying analysis to either realistic hypotheticals or the actual state of affairs might be a good idea if one is trying to understand those states of affairs. The third problem more intimately focuses on «should implies can» and that is that much of the theorizing about the field of evidence from the cognate disciplines makes essentially impossible epistemological demands, and to the extent that is true the resulting analysis is of little use for the jurist.

Collectively, these three problems may be driven by a factor that is at least contingently if not necessarily highly correlated with a priori analysis, which is the tendency to attempt to reduce problems to a small set of variables—one if possible—in order to identify «the» solution or key to a problem or answer to a question. At a very fundamental level, this effort to reduce things to various sorts of algorithms misconceives the nature of the inquiry (if again one is trying to explain juridical proof). The American legal systems are not simple and static but instead complex and dynamic, and to be understood must be approached from that perspective⁵. Legal systems, like markets and democratic governments, are complex adaptive systems, and their «parts» (for example, the field and rules of evidence or procedure, conventional trial practices and so on) are part of the construction of that complex adaptive system. To

⁵ See, e.g. ALLEN, 2013. For an argument about the misdirection of «analytic jurisprudence» that is sympathetic to the central argument of this paper, see PRIEL, 2020.

approach the field of evidence as though it were a simple and static matter usefully subjected to a priori, top-down analysis would be like asking what are the simple rules governing the Amazon River's relationship to the rain forest, or asking what the point of either the Amazon River or the rain forest is. Answers to such questions will be uninformative at best. So too will be algorithmic answers to the nature of juridical proof.

That, at least, is the burden of this essay. It is not to give advice to other disciplines how to do their work; rather, it is to explain why what they do often cannot be used by the jurist—pure solipsism in other words, but from a perspective «to highlight the importance of continued dialogue between legal epistemology and formal epistemology»⁶ at least for the jurists. I start with an examination of the success of relative plausibility in supplanting probabilism as the best explanation of juridical proof, which includes a discussion of the quite odd debate over probabilistic/statistical evidence, an odd debate stimulated in large measure by its failure to attend to the lessons of naturalized epistemology. I then turn to some possible causes of that oddness: weird hypotheticals, neglecting the actual state of legal systems, and impossible epistemological demands.

The interrelationship of law and probability is ancient⁷, but sustained scholarly attention was not focused on it until the seminal article by John Kaplan in 1968⁸. Probability theory in general and decision theory as its companion seemed to provide a compelling framework for understanding much of juridical proof. Indeed, it is not too much to say that the idea swept the field that conventional probability theory applied directly to legal decision-making, and that there remained only idiosyncratic problems like naked statistical evidence⁹. Cracks in the façade appeared when L. Jonathan Cohen identified a set of proof paradoxes latent in the probability approach¹⁰. He thought he could resolve them by replacing one formalism—probability theory—with another—a formalized Baconian inductive logic¹¹. Shortly thereafter, the paradoxes were extended and another somewhat formal solution offered to solve the problem—a reconceptualization of trials rather than of probability¹². The problems were not solved, and deeper anomalies were discovered¹³.

The anomalies and the failure to resolve them were driven by the variables noted above: weird hypotheticals that distract rather than facilitate inquiry, failure to attend

⁶ HEDDEN and COLYVAN, 2019. For an example of the general proposition of this article implemented, see ALLEN, 2020.

⁷ See HACKING, 2006. For criticisms of the first edition of Hacking's book, see GARBER & ZABELL, 1979.

⁸ KAPLAN, 1968.

⁹ See e.g., KAYE, 1979a, 1979b, 1982.

¹⁰ COHEN, 1977.

¹¹ *Ibid.*

¹² ALLEN, 1986.

¹³ ALLEN, 1997.

to the actual conditions of juridical proof, and impossible epistemological demands, all driven by the reductive tendency of a priori analysis. Take weird hypotheticals and the related problem of not attending to the actual conditions of juridical proof first. Much discussion has focused on hypotheticals that are designed to elucidate the issues surrounding statistical evidence in particular and probability theory more generally. Two in particular dominate: the Gatecrasher hypothetical of Cohen, and the Blue Bus example of just about everybody. Rather than being enlightening about juridical proof, the attention focused on these (and similar) hypotheticals may have simply impeded progress.

The Gatecrasher involves a rodeo with 1000 attendees but only 499 have paid. The rodeo operator picks one at random to sue for the entry fee, and those are the only facts before the court. Who should win? According to a version of probabilism, the rodeo operator, as the probability is greater than .5 that the randomly selected person did not pay. This, in turn, is said to violate our intuitions that purely statistical evidence is not sufficient for liability, and thus counts against a probability interpretation of juridical proof¹⁴. The Blue Bus hypothetical is similar and is said to be inspired by a real case¹⁵. Imagine an accident where an injured party says she was injured by a bus, that the Blue Bus Company operates 70% of the buses in the city and the Red Bus Company operates 30%, and that is the only evidence. Should the Blue Bus Company be liable? A number of legal academics as well as those from other disciplines apparently think no and that that is the result the legal systems would reach. How does that compare to a witness who is 70% reliable and testifies that a Blue Bus hit her? By comparison to the directly statistical, those same academics often shift to yes, the answer they assert that would also be given by the legal system¹⁶.

¹⁴ Enoch et al. assert that «the powerful, uncompromising intuitions of pretty much all of [evidence] scholars» is that there is a critical difference between statistical and non-statistical evidence and that the law is consistent with that intuition. ENOCH, et. al, 2012. For what it is worth, that is neither my intuition nor more importantly my reasoned judgment and as discussed below it also grievously misstates the «law». Perhaps reflecting a similar impatience with weird hypotheticals, Christian Piller, «criticize[s] safety without participating in the production and discussion of ingenious counterexamples». PILLER, 2019: 307. Jonathan Weinberg, Shaun Nichols, and Stephen Stich provide some evidence that Gettier intuitions are culturally dependent, and from that suggest that intuitions should not be relied upon. WEINBERG et al., 2001. A recent article, GARDINER 2020, argues that many of the intuitions concerning these various hypotheticals are unjustifiable in large measure due to the failure to account for differences in the scenarios. From this, she questions whether a unified explanation of various hypotheticals can be provided. This recognizes the complexity being addressed, and to that extent is consistent with my argument; she does not take into account the implications of adaptability, which makes the matter of complexity all that more complex.

¹⁵ Smith v. Rapid Transit, 58 N.E.2d 754 (Mass. 1945).

¹⁶ See, e.g., ENOCH et al., 2019 for a version of the Blue Bus hypothetical that concludes that if there is no other evidence, and even if «the Blue Bus Company controls 70 percent of the market in the relevant area, the law [...] refuses to find the Blue Bus Company liable». This is usually equivalent to the statement that statistics are an inadequate basis for liability. Even though repeated ad nauseam in the literature, the statement is false, as the text elaborates. There is some empirical support for the proposition that people behave consistently with these intuitions. WELLS, 1992.

Employing hypotheticals such as these is a popular strategy in philosophy and it can be a helpful way of abstracting away from unnecessary details of real-life cases that are orthogonal to the questions at hand. However, for these hypotheticals to be helpful (to jurists), they must be crafted without losing sight of the actual object of inquiry, namely, the legal system. Both fail to do so, and thus neither of these hypotheticals is enlightening about juridical proof. In the Gatecrasher hypothetical, the obvious problem is that in the real world there are numerous sources of evidence that would be proffered by the defendant, such as their own testimony, the ticket stub, a receipt, the testimony of someone who observed the person buy the ticket, to name just a few¹⁷. Cohen assumed no such evidence could be forthcoming, but the reason given demonstrates the problem with the hypothetical: «Perhaps the defendant was prevented by death from giving evidence on his own behalf and the relentless management pursued the case against his estate»¹⁸. In that case, in many states the issue would never arise because the dead man statute would prohibit interested parties from testifying to interactions or transactions with the dead person¹⁹. But what if you were not in a dead man state? The death of the defendant would not preclude any of the sources of evidence noted above, and there are many more in addition. A plethora of hearsay exceptions would apply that allow testimony from third parties about the person's intentions, and thus what the person likely did²⁰, and in any event who was the person sitting next to? Did anyone see him or her at the rodeo, and did they discuss the price of tickets? The hypothetical does not abstract away from unnecessary complications in order to distill the essence of some question facing American legal systems; it abstracts away from many aspects of the legal system that connect the hypothetical to reality²¹.

Perhaps focusing attention on limited and attenuated aspects of the formalities of proof and evidence, rather than on the structure of actual legal systems and dispute resolution, led to the neglect that trials pursue many objectives in addition to truth, one very important one being a governance function that creates incentives for ap-

¹⁷ The next few paragraphs are little more than an elaboration on KAYE, 1982, who argued that it was sensible to construct incentives to produce more evidence when it is available, but also to decide cases on the basis of what he called naked statistical evidence when that is all the evidence that there is. That is exactly right, although the structure of those incentive programs can be complicated.

¹⁸ COHEN, 1981a.

¹⁹ See, e.g., *Spencer v. Wayne*, 80 N.E.3d 764 (2017 IL App (2d)); *Matter of Estate of Owen*, No. 35879-8-III, 2019 WL 6876791, at *6 (Wash. Ct. App. Dec. 17, 2019) («[Washington's dead man's statute] provides that in an action against an estate, an interested party may not offer testimony adverse to the estate regarding transactions made with the decedent». (Citing WA ST 5.60.030 (West 1977))).

²⁰ For a famous example, see *Mutual Life Insurance Co. of New York v. Hillmon*, 145 U.S. 285 (1892).

²¹ Responding to a version of the unrealism point in the text, David Enoch and Talia Fisher write: «This is an empirical matter lying outside the scope of this Article». That captures the distance between the opposing approaches to these problems that being addressed. ENOCH and FISHER, 2015: n. 17. The hypothetical also involves strange epistemological demands. Some critical facts are taken as absolutely certain, with no explanation as to how that could be so. For further on this, see n. 57, *supra*.

propriate behavior²². Allowing recovery in cases like the Gatecrasher would create incentives for liability traps—hold events, and start suing people randomly for entry fees. That is absurd, of course, but what is not absurd is for the law to discourage people like rodeo operators from creating anything analogous to liability traps by disallowing recovery when the real problem is the very conditions that the plaintiff created rather than the asserted behavior of others. In more concrete terms, it is perfectly appropriate for the law to conclude that it is the rodeo operator's problem if he or she does not fence off the rodeo adequately, have security or a controlled entrance point, fails to install video cameras, or whatever²³.

When more general incentive considerations did enter the picture, the result was equally strange. For example, one recent work argued that a serious problem with statistical evidence of the Gatecrasher sort—and a potential basis for distinguishing it from testimonial evidence—is that it eliminates an individual's incentive to comply with the law²⁴. If one knows that the mere fact of being selected will result in liability, why bother in the Gatecrasher context to buy a ticket? Just crash the rodeo and take your chances. You cannot be worse off as the worst-case scenario is you get picked randomly and end up having to pay the entry fee, whereas if you actually buy a ticket you might end up having to pay twice. The answer is parallel to the discussion above about incentives for rodeo owners: if you buy the ticket and keep the stub, get a receipt, have a witness observe you, or speak the language of the country the trial is in so that you can testify yourself, you will not suffer the inconvenience of being sued. It seems like an easy trade-off. Only if your selection is random and unknown forces of nature prohibit you from any of the escape hatches will you be liable, but the existence of such a world is fanciful²⁵.

The Blue Bus hypothetical may be somewhat less ridiculous than the Gatecrasher hypothetical, but it, too, contributes little of interest regarding the nature of juridical proof. It is simply an example of lawyerly ineptitude, which seems to be what caused the court in the *Smith v. Rapid Transit* case to do what it did. In the actual

²² ALLEN, 2015. Prof. Nance has extended the governance function to include incentives to produce the best evidence, which may serve the epistemological aspirations of the American legal systems. See NANCE, 2016. His book is the best and most systematic examination of the implications of a general best evidence principle that animates a significant portion of the law of evidence.

²³ Prof. Kaye long ago noted a possible incentive effect concerning the production of evidence. KAYE, 1979c. The difficulty, which made the debate at the time «odd», ALLEN, 1986: 411, is that such an incentive effect operates, or can operate, on all parties to litigation. There is no reason to limit the analysis to plaintiffs or defendants.

²⁴ ENOCH and SPECTRE, 2019. Also, ENOCH et al., 2012.

²⁵ There are other arguments in the literature about incentives that neglect the almost infinite array of choices a person can make to influence their likelihood of liability. See SANCHIRICO, 2001. Character evidence is said to be an example of an incentive killing evidence rule but in fact it just creates different incentives—like don't go to bars where fights occur or jewelry stores where robberies happen. Or if you do go, go with a responsible companion. For a discussion of such matters, see PARK & SAKS, 2006. The incentive arguments are discussed further below.

case, no evidence was adduced as to the prevalence of private buses in the area, of the safety record of the company being sued, of the likelihood of a bus being in the area of the accident at the appropriate time given the bus schedules of the company, of interrogations of the bus drivers that might have been present, of any effort to search for other eyewitnesses, and so on²⁶. If the case stands for anything, it is to vindicate Craig Callen's point that an explanation of juridical evidence should contain an escape hatch for cases that are just woefully insufficient, regardless of any other circumstances²⁷. And as David Kaye argued with respect to the Gatecrasher hypothetical, to give lawyers and parties cost/benefit justified incentives to do their jobs properly²⁸.

The problems with the debates over probability theory extend beyond weird hypotheticals disconnected to reality to the third problem—impossible epistemological demands. Probabilities have to come from somewhere and in a typical case (unlike employment discrimination or DNA cases, or stylized hypotheticals), they would come from the massive amounts of discrete pieces of qualitative evidence adduced at trial. There are numerous debilitating difficulties for probabilistic theories of juridical proof resulting from the manner in which juridical proof actually unfolds²⁹. For example, any effort to deal with a mass of evidence in a probabilistic manner immediately confronts computational intractability³⁰ and without computation there is little point to probabilistic approaches (otherwise people are just making up numbers)³¹.

There is a still deeper problem here. For probability theory to be explanatory of evidentiary processing at trials, people would need to know the necessary probabilistic data to appraise evidence, such as accurate base rates and likelihoods. But to do that, they must already know the outcome of the case being tried. Evidence at trial is obviously contingent in the sense that the implications of any particular piece of evidence is a function of all the other data relevant to the case. Litigated cases are not like pari-mutuel betting in which fully specified statistical rules determine precisely the effect of each new bet. Evidence of the existence of a dead body does not mean that a murder occurred, for example, because subsequently admitted evidence may show the death to have been accidental or in self-defense. The evidence that the de-

²⁶ There was evidence that indicated the accident occurred on the defendant's bus routes at a time within the bus schedule's route, but little else. No effort to show that, given the bus schedule, a bus should or could have been close to the scene of the accident at the asserted time it occurred, and as the court said, «While the defendant had the sole franchise for operating a bus line on Main Street, Winthrop, this did not preclude private or chartered buses from using this street; the bus in question could very well have been one operated by someone other than the defendant». *Smith v. Rapid Transit, Inc.*, 58 N.E.2d 754 (Mass. 1945).

²⁷ First articulated in CALLEN, 1991.

²⁸ KAYE, 1979c.

²⁹ ALLEN, 1997.

³⁰ CALLEN, 1982.

³¹ For an example of the futility of making up numbers, see HEDDEN and COLYVAN, 2019, and ALLEN, 2020.

defendant was in the vicinity can be quite exculpating if there is further evidence that he always visits his sick mother at that particular time. If evidence is not contingent in this sense from a fact finder's perspective, then the fact finder already must know that there was no accident or self-defensive act or that the defendant did commit the crime, but that is exactly the point. To implement a probabilistic theory of evidence such as Bayesian updating requires that the outcome already be known, which of course makes the trial itself superfluous³².

The astute observer will notice, and possibly complain, that analytical points have been met with pragmatic considerations. For example, a primary concern about the Gatecrasher hypothetical is that it is ridiculous on its face and would never arise in an American legal system, and the significant point about the Blue Bus hypothetical is that it is an example of lawyerly ineptitude. Perhaps these responses really are "fighting the hypothetical", but even if so perhaps it is justified. One should not fight the hypothetical when it may actually be a productive object of reflection, but it is a different matter when that is not true. The hypotheticals do not test intuitions about an actual legal system, though they purport to do so. Instead, the assumptions that they rely upon misconceive the object of inquiry, and thus reflecting upon them may obscure rather than enlighten about an actual legal system. Furthermore, as discussed below, when a real case analogous to the hypotheticals actually arises—such as cold hit DNA cases—the results are precisely what they should be if the point is error allocation. And again, there may be other «points» than error allocation justifying differing rules. One must actually look at a legal system to see if this is true.

By comparison to the formalistic probability approaches, what one sees when one actually looks at the American legal systems, and the scope of the project of understanding juridical proof, was recently succinctly explained:

Its scope is [...] about the entire process of proof, including (1) the form, securing, and presentation of evidence, (2) the forms of argumentation employed at trial, (3) the manner in which humans process and deliberate on evidence, (4) the trial structure created by the rules of evidence and procedure, (5) the structure of litigation before and after trial, (6) the manner in which judges and juries, on the one hand, and trial and appellate judges, on the other hand, interact, and (7) to some extent about the meaning and nature of rationality. In other words, all of these features comprise the sprawling entity that we refer to as "juridical proof", and relative plausibility is an attempt to explain that entity in all of its aspects, from beginning to end³³. How any part of the process operates is thus relevant to the project...

³² For a discussion, see ALLEN, 1997. By contrast, statistical evidence plainly can be probative in the context of deciding the relative plausibility of competing explanations of events. This emphasizes why it is important to keep in mind what one is trying to explain. As noted in the text, relative plausibility is attempting to explain the proof process from beginning to end, and I am taking the propounders of the weird hypotheticals discussed herein to be engaged in an analogous effort. It's quite a different matter if one is simply trying to explain idiosyncratic cases or evidentiary proffers.

³³ At the beginning of litigation, see, for example, the U.S. Supreme Court's insistence that a plaintiff's allegations in a complaint must be «plausible», and that plausibility is assessed by comparing the allegations with other possible alternative explanations. See *Bell Atlantic Corp v. Twombly*, 550

Relative plausibility explains this sprawling entity and its interconnected parts better than conventional probabilism. Relative plausibility explains the behavior of lawyers in preparing for trial, judicial involvement in pre-trial hearings, and the formulation of theories of the case. It explains virtually all of the procedural context at trial from the opening statements through disjointed presentations of evidence to closing arguments to the meaning of relevance (where by contrast probabilism cannot explain the massive overlap of evidence so that it predicts that most of the evidence actually admitted at trial is irrelevant)³⁴. Relative plausibility explains the behavior of trial and appellate judges. It is explained by (the relationships are bilateral)³⁵ the basic requirements of rationality and the potential and limits of cognition. Most of which the conventional probability explanation fails utterly to accommodate. Indeed, to the extent that relative plausibility relies on what decision-makers do, it exploits the necessity of knowing what the alternatives are and how they are examined for truth value in the process of rational decision-making. The primary message of relative plausibility is that from beginning to end the legal system pushes the parties to provide competing explanations, and these explanations structure the decision that is subsequently made (even if the decision is based on an explanation not advanced by the parties)³⁶.

Relative plausibility has received sustained attention in the literature, and need not be belabored here beyond, simply noting that its success as an explanatory theory stands in stark contrast to the consistent pattern of failure of probabilistic explanations. It also bearing noting that the development of relative plausibility was not influenced by weird hypotheticals; it embraced rather than neglected the actual state of the systems under inquiry; and it exploited rather than assumed away human cognitive capacities. I, of course, see causation operating here rather than correlation.

Other reductionist approaches to juridical proof that have followed in the wake of the demise of the probabilistic explanation exhibit similar attributes of invoking weird hypotheticals, not attending to the actual condition of legal systems, and making impossible epistemological demands. I will discuss two examples: the sensitivity

U.S. 544, 567-78 (2007); *Ashcroft v. Iqbal*, 556 U.S. 662, 682 (2009). Toward the end of the process, see, for example, the fact that appellate courts routinely evaluate the sufficiency of evidence to support a verdict by comparing alternative explanations. See *Anderson v. Griffin*, 397 F.3d 515, 521-22 (7th Cir. 2005); *Bammerlin v. Navistar*, 30 F.3d 898, 902 (7th Cir. 1994); *Yeschick v. Mineta*, 675 F.3d 622, 627 (6th Cir. 2012); *O’Laughlin v. O’Brien*, 568 F.3d 287, 304-08 (1st Cir. 2009); *United States v. Beard*, 354 F.3d 691, 692-93 (7th Cir. 2004); *United States v. Newell*, 239 F.2d 917, 920 (7th Cir. 2001). See also *United States v. Sever*, 603 F.3d 747, 753 (9th Cir. 2010) (abuse of discretion in denying defendant’s discovery request for evidence to support defendant’s «alternative explanation»).

³⁴ See *DIAMOND et al.*, 2012; *VIDMAR and HANS*, 2007; *BENNETT and FELDMAN*, 1981; *WAGENAAR et al.*, 1993. On overlapping evidence see, *PARDO* 2013: 576-89.

³⁵ See, e.g., Ribeiro’s elegant presentation of this point: «What I am suggesting is that the testimony is relevant because it makes the hypothesis that the employee stole the necklace a better explanation, not because the evidence is explained by this hypothesis». *RIBEIRO*, 2019: 109.

³⁶ *ALLEN and PARDO*, 2019b: 207-208. The Supreme Court of the United States recently commented on this fundamental aspect of the American legal systems in *United States v. Sineneng-Smith*, *** U.S. *** (decided May 5, 2020): «In our adversarial system of adjudication, we follow the principle of party presentation. As this Court stated in *Greenlaw v. United States*, 554 U.S. 237 (2008), “in both civil and criminal cases, in the first instance and on appeal [...], we rely on the parties to frame the issues for decision and assign to courts the role of neutral arbiter of matters the parties present”. *Id.*, at 243».

and incentive proponents on the one hand and the relatively new idea of «normic» support on the other³⁷.

The sensitivity theorists, perhaps more influenced by the realities of the legal systems, are more cautious in their conclusions than were the probability theorists³⁸, but one sees analogous problems popping up in their discussions of hypotheticals designed to elucidate the nature of juridical proof generally and specifically the difference between statistical and non-statistical. Here is a recent bus color version:

Eyewitness: A bus causes harm, an eyewitness recognizes the bus as belonging to the Blue Bus Company. The witness, however, is imperfectly reliable; let us say that she is roughly 70 percent reliable in matters such as this one. The law has no qualms about accepting the eyewitness testimony as evidence and indeed basing a positive finding that the bus was a Blue Bus bus (and perhaps also that the Blue Bus Company is liable) on the testimony³⁹.

Statistic: A bus causes harm. There is no eyewitness, but we have uncontested data regarding the distribution of buses in the relevant area; in particular, the Blue Bus Company owns roughly 70 percent of the buses there. Here, though, the law typically will not be willing to base a positive finding of fact—and certainly not liability—on just this kind of evidence, sometimes called statistical evidence⁴⁰.

David Enoch, Talia Fisher, and Levi Spectre take this hypothetical as a starting point for analyzing sensitivity. A verdict, v , is «sensitive» if and only if, had v been false, fact finders would (most probably) not have found that v ⁴¹. Verdicts based on the market share statistic are insensitive because, had it not been a blue bus that caused the harm, the statistical evidence nonetheless would lead the jury to falsely find the Blue Bus Company liable for the harm, given that the market share data would remain the same in those cases. The verdict based on the eyewitness testimony, by contrast, is sensitive. Had it not been a blue bus that caused the harm, Enoch et al. assert, the witness would have «probably» testified accordingly and, thereby, the jury probably would have come to correctly find that the Blue Bus Company is not liable for the harm⁴².

The sensitivity theorists are careful to note that their claims about sensitivity are not meant to be definitional of statistical evidence but, rather, are meant to provide a general «instrumental, incentive-based» explanation for why the law generally prefers

³⁷ Also addressed in passing is Sarah Moss' work in *Probabilistic Knowledge*, Moss (2018), which possesses analogous attributes. See n. 68, *infra*.

³⁸ On probabilistic imperialism, see EDWARDS, 1986. See also many of the responses to COHEN, 1981b, in *Behavioral Brain Sciences*, 4(3).

³⁹ ENOCH et al., 2019: 197.

⁴⁰ *Ibid.*

⁴¹ Their formulation of sensitivity is framed in terms of beliefs as follows: «*Sensitivity*: S's belief the p is sensitive =df. Had it not been the case that p , S would (most probably) not have believed that p ». *Ibid.*: 204. I have taken the liberty of changing the epistemic notion of «belief» to the legal notion of «verdict» given that the legal notion is what is of central interest.

⁴² ENOCH et al., 2019: 197, 206, n. 13.

non-statistical to statistical evidence⁴³. They reiterate the argument against statistical evidence that an agent's reasons to comply with the law are «seriously compromised unless the agent knows, *ex ante*, that her action will affect how the law treats her; that is, unless she has very good reason to expect the law's treatment of her to be sensitive to her behavior»⁴⁴ and that «treating statistical evidence as the law treats individual, direct evidence would diminish the incentives supplied by substantive law in sufficiently many central cases»⁴⁵. Although they recognize that sensitivity alone may not be of central concern to the law, they see it to be intimately bound up with their incentive argument.

Regardless of how one gets to their destination, the general point about disincentivizing agents seems particularly peculiar for analogous reasons given above with respect to the Gatecrasher. There are two possible applications of the incentive argument, although Enoch et al. do not sort them out. One is the effect on the most immediately critical agents, the bus drivers, and the other is the effect on their employers, the bus companies. The possible effect on bus drivers is discussed first because it is so obviously wrong, followed by a somewhat more extended discussion of the effect on bus companies.

According to one application of the argument of Enoch et al., the law does not hold the Blue Bus Company liable on the basis of statistical evidence because doing otherwise would disincentivize bus drivers from causing such harm. This would require that bus drivers of both red and blue buses ponder the risks of harm to them or their companies from being held liable on the basis of statistics regarding their company's market share. It would be quite surprising if bus drivers even knew about let alone had these statistics in mind given that there are far more important things to be mindful of in their line of work, and far more incentivizing reasons to avoid causing harm, such as losing their jobs, hurting pedestrians, or being shunned by their colleagues, just to name a few. Moreover, the bus drivers would have to think that they would be in the category where the exclusive evidence is statistical—only if they run over pedestrians where no other probative evidence is available will the statistical evidence be of importance. That seems a bit implausible.

Enoch et al. recognize the limitations of the incentive argument and mention what one might call the larger concerns of the legal system contained within the governance function of evidence law⁴⁶. For the law to be sensible and for the incentive argument to get any traction, the various options and the costs and benefits of all the relevant participants' behavior have to be taken into account, which includes

⁴³ ENOCH et al., 2012: 200; ENOCH and SPECTRE, 2019: 181.

⁴⁴ ENOCH and SPECTRE, 2019: 181, 183.

⁴⁵ *Ibid.*: 182. ENOCH et al. also invoke normic support for the distinction that they are drawing, which is addressed below.

⁴⁶ ENOCH et al., 2019: 219. On the governance function, see *supra.*: n. 22. For the most part, this is just another aspect of socially useful incentives.

the activity of bus companies in addition to their drivers. Brushing aside the larger systemic effects, Enoch et al. conclude that the testimonial evidence is still to be preferred to the statistical «because there is no similar incentive-corrupting effect to relying on individual evidence—even individual evidence that is probabilistically indistinguishable from the relevant piece of statistical evidence»⁴⁷.

This is wrong on both counts. On the one hand, the argument that the testimonial evidence is probabilistically equivalent but somehow qualitatively different from the statistical is another problematic inference from a weird hypothetical. On the other hand, bus companies as well as rodeo operators and attendees have many options in life and the risk of this form of liability is likely to cause them to take precautionary action. I discuss these two points in turn.

The hypothetical demonstrates the subtle fashion in which impossible or outlandish epistemological demands can be smuggled into weird hypotheticals⁴⁸. Recall that according to Enoch et al., what makes the witness testimony sensitive is that «had it not been a Blue Bus, she would have probably not testified that it was»⁴⁹. Well, actually, there is by hypothesis a 30% chance that she would have, but what does that mean? The authors are implicitly asking the reader to assume that the 30% error rate is composed of a random variable over all decisions in the (undefined) reference class, and therefore, the witness would «probably not» identify a red bus as a blue bus in the incident in question (the errors being distributed over all colors of buses). The point based on this implicit distribution is that the probability of any particular kind of error in the class of all bus identifications would be small, which

⁴⁷ ENOCH et al., 2019: 219.

⁴⁸ A sin that seems ubiquitous in this literature. My frequent co-author Michael Pardo falls prey to the weird-hypotheticals-that-make-peculiar-epistemological-demands dynamic in his otherwise quite sensible tracking of the relationship and pitfalls between safety and sensitivity on the one hand and reliability on the other. Here is one example from PARDO, 2018: 66:

Drug Weight: A criminal defendant is tried for possessing a large amount of an illegal substance. According to the criminal statute at issue, the prosecution must prove that the defendant possessed 500 grams or more of the substance. A chemist who sampled and tested the substance will testify as an expert that the amount seized was over a kilogram. The expert is well qualified, employed an acceptable and reliable methodology, and otherwise satisfies the criteria for admitting expert testimony. The chemist's process, however, tends to overestimate weight by a very small amount (say, by one gram or less).

Pardo uses this to illustrate safety through the use of possible worlds. The testimony is safe when it is nowhere close to whatever the element of the crime is, and otherwise it is not because even though the true amount is somewhat less than the statutory requirement the chemist would testify that it exceeds that amount. This seems to me entirely unnecessary. If one «knows» that the process «overestimate[s] by a» gram, then one «knows» what the true amount is. If the chemist's mistake is within a range, then again you know the range of the true weight. This hypothetical makes sense only if you already know what you need to know, and if you do, then one simply presents it to the fact finder. If one does not «know» the error, then the argument about safety does not run through. Notwithstanding this weird hypothetical, I doubt that his views are much at odds with the general thrust of this article. See Pardo 2005.

⁴⁹ *Ibid.*: 206.

is why they say the witness would «probably» not misidentify the bus. But note the smuggling in of knowledge. How do they know what kind of errors the witness is going to make? How do they know that this witness is making random errors (actually uniformly distributed errors) and simply not able to identify red buses under the conditions present, so that in fact most and maybe all of her error rate is the product of misidentifying red buses as blue?

This implicit smuggling in of an answer to a critical question of «knowledge» that assists the analysis is highlighted by considering the unrealism surrounding the claim that the witness is «70 percent reliable». It is all but impossible to decipher the meaning of that statement. Is the reference class the testimony of all witness under all conditions and the generalization is being applied to the witness? Does it refer to bus accidents at night? In what kind of roadway, weather conditions, and lighting due to the phases of the moon? Did someone do a study of this witness? Under what conditions and relating to what kind of incident? Just like 70% market share does not entail causing 70% of the accidents, 70% reliability does not entail that the witness in this particular case with respect to this particular harm is 70% reliable in reporting the truth about what she saw (again, whatever that might mean).

In any event, presumably Enoch et al. believe that we all live in a causal universe, and whatever causes the witness' errors will cause them every time, just like the statistic will always point to the Blue Bus Company. Summed over all possible cases, these two scenarios are structurally identical which means that they are identical from the point of errors even if not from the point of distribution of those errors. The witness' testimony is not more «sensitive»; rather, one simply does not know what causes her mistakes and thus one does not know how they will be distributed. One form of evidence is being preferred for no other reason than ignorance of its distributional effects⁵⁰. That is peculiar.

In comparing the statistical evidence of bus distributions to testimonial evidence, the bus distribution data (assuming it maps perfectly onto accident rates) is at least as good as the testimonial evidence, as the incentive argument properly understood makes clear. Just like rodeo operators can easily protect their income, bus companies can protect against statistical anomalies based on factors under their control⁵¹. These are complex and adaptive systems, in other words, and not simple

⁵⁰ ENOCH et al. (2012) may agree with this, although the matter is unclear. They suggest at times that perhaps there is a qualitative pertinent difference between sensitive and insensitive evidence, yet they also take the position that the law should not discard evidence even if statistical that contributes to reliable results. Well, how «reliable» is the statistical proffer as compared to the testimony, in their view? If what they mean is that there really is no interesting difference between the two forms of evidence, it would have been helpful, and saved a lot of time, for them to have just said that. For sustained treatment of the lack of pertinent epistemological distinctions between «statistical» and «nonstatistical» evidence, see PUNDIK, 2008, 2011.

⁵¹ Another demonstration (see *supra.*: n. 21) of the distance between some forms of theorizing and the naturalist approach is the response of David Enoch and Talia Fisher to a point analogous to the one

and static. For example, bus companies facing liability because of market share can engage in enhanced safety practices to reduce accidents and thus errors overall, and they can reduce errors at trial through such means as tracking the movement of their buses or installing video cameras. That, in turn, will affect the position and thus incentives of other bus companies, the combined effect of which almost surely will increase social welfare (fewer accidents in society and errors at trial)⁵². The real distinction between the bus and witness evidence has little to do with the incentive to engage in socially useful conduct but, once again, is simply that the statistical form carries its ambiguity on its face whereas the testimonial form does not. If preferring the latter to the former has any impact on overall social welfare, it is probably to decrease it⁵³.

in the text. They say: «Our answer to this critique is that while this effect may take place, as long as there are instances when it is counterbalanced by the chilling effect suggested by Sanchirico, both should be taken into account. And of course, one must also bear in mind that overdeterrence is also a problematic result» ENOCH and FISHER, 2015: 593, n. 130. That is exactly right and emphasizes the lack of utility for the jurist of the formal theorizing. The problem for the jurist is trying to identify in the real world what might maximize total social welfare. See ALLEN, 2015.

⁵² I may be accused here of engaging in my own a priori reasoning, but my defense is that it attempts to take into account real features of the world that bear on the construction of legal systems.

⁵³ To bolster their a priori incentive argument, they summarize Richard Posner's ridiculous a priori assessment of the Blue Bus hypothetical from the perspective of micro-economics. ENOCH and FISHER, 2015: 583, n. 83, say:

In addition to these precautionary distortions, the admissibility of statistical evidence may also impair market competition: Each of the bus companies will be incentivized to hold less than fifty-one percent of the market share so as not to suffer from the evidentiary disadvantage that a larger market share imposes. Moreover, the company holding the larger market share will absorb higher liability costs, which may lead to a decrease in its market share and to a possible exit from the market.

Here is the argument from POSNER, 1999: 1510:

There is still another objection to allowing the bus case to go to the jury. If B, though responsible in fact for almost half the accidents, is never held liable and A is always held liable, A will have a big incentive to be careful and B little or no incentive to be careful. As a result, over time, more than half the accidents will be caused by B, increasing the error rate resulting from allowing juries to base decisions on the ratio of the companies' buses on the route in question. Eventually, A, having higher liability costs, will probably withdraw from the route; the rule on burden of proof will have created a monopoly!

The absurdity of this argument because of its neglect of the adaptive nature of the real world was pointed out by ALLEN and LEITER, 2001: 1526:

Our doubts about the value of Posner's form of a priori reasoning for explicating the law of evidence are well captured by this paragraph. Taking the argument on its own terms-terms which we reject, of course-the consequence Posner fears will occur only if A is remarkably stupid. In the world Posner is hypothesizing, all A need do is take at most three buses out of service. If A takes three buses out of service, the ratio of buses now favors A (it changes from 51/49 to 48/49), and B will henceforth be held liable for all accidents. Of course, another economic fear might now arise: B, also not being massively stupid, might respond by taking two of its buses out of service. Perhaps the prediction would now be that «bus reduction» would become recursive, eventually resulting in only one bus company with only one bus driven by a very careful driver, which would surely not serve the needs of the community! Therefore, the logic would go, allowing verdicts based on naked statistics is not economically sensible.

In any of the Blue Bus contexts, that «the law typically will not be willing to base a positive finding of fact—and certainly not liability—on just this kind of evidence, sometimes called statistical evidence»⁵⁴, may be right but for the wrong reason. The reason that the law would not be willing to hold the Blue Bus Company liable on the basis of this statistic has nothing to do with the claim that the statistic is insensitive or incentive crushing. Rather, it is that the statistic is just another example of lousy evidence. The fact that the Blue Bus Company owns 70% of the buses does not entail that there is a 70% chance that the Blue Bus Company caused the harm in question. The Blue Bus Company may own 70% of the buses but have extremely effective safety measures in place; they may have well-seasoned drivers with impeccable driving records; they may not even have buses that frequently visit the location of the harm at the time in question. Actual fact finders could easily regard this statistic as being insufficient to ground liability because these relevant possibilities remain unaccounted for by the statistic about ownership.

Once modified with an element of realism, the hypothetical no longer demonstrates any difficulties with statistical or market share evidence that distinguish it from testimonial evidence. A litigator would know that the relevant statistic would be a risk assessment of how likely it is that a blue bus rather than a red bus caused the harm in question, obviously enough, and construct a case on that basis. That might include market share data but also would include data regarding the records of each company's drivers, statistics about the frequency with which each company's buses visit the location of the harm, statistics about the effectiveness of each one's safety measures, figures about similar harms caused by each company in the relevant location, and so on. The «and so on» is centrally important. If sensitivity theorists are correct, then the number of variables included in the risk assessment should not matter because all statistics will be equally insensitive and incentive eradicating. Including more and more relevant statistical evidence ad infinitum—making this an extremely thorough assessment—should not influence fact finders' judgment about the case, nor evidence theorists' assessment of the evidence.

But does it? Suppose that, instead of the single statistic about market share, a litigator presents an extremely thorough, multi-variable assessment of the risk posed by the Blue Bus Company and Red Bus Company that reveals that there is a robust 70% likelihood that the Blue Bus Company and a 30% chance that the Red Bus

As was also pointed out by ALLEN and LEITER, 2001: 1526, n. 106, taking such a priori thinking seriously and assuming really smart bus companies could lead to the prediction of no liability ever for accidents. The bus companies would simply have to run the same number of buses so that the probability of both would always be .50, which does not meet the preponderance standard, and so the plaintiff would always lose. Such logical outcomes indicate a serious problem with the logic. As Allen and Leiter concluded: «Neither a monopoly nor the essential elimination of the industry would result from allowing probabilistic verdicts. If such verdicts began to accumulate inaccurately against bus companies, they would invest in precautions, many of which are cheaply available».

⁵⁴ ENOCH et al., 2012: 197.

Company caused the harm on the day and at the time in question. Would a sensible legal system still regard the eyewitness evidence as more worthy than the statistical evidence? Of course not. When it comes to the actual trial setting, a complex body of evidence consisting of several pieces of relevant and interrelated statistical information that converges in a way that reveals a 70% likelihood that a blue bus caused the harm is at least as good, if not better, evidence as a body of evidence that includes only the testimony of a single witness who is known to be 70% reliable in identifying the color of the bus that caused the harm. The market share evidence presented is insensitive, while the testimonial evidence is apparently sensitive according to these theorists—but their evidentiary value is either the same or the statistical evidence is better. The relative value of statistical and testimonial evidence does not map on to considerations about sensitivity and deterrence (as the sensitivity theorists conceive it), as the hypothetical is designed to show.

The real difference between statistical evidence and testimony lies in the trial convention of assuming reliability of witnesses until credibility is attacked. This convention obviously does not apply to statistical evidence, as its reliability is immediately apparent (under certain assumptions to make the hypothetical comprehensible). Once a witness' credibility is attacked, a «risk assessment» highly analogous to what might be done of bus companies can occur with the opposing party pointing out the «risks» in the testimony. The risk assessment of the testimony involves identifying reasons why the testimony may be unreliable. In the case at hand, this may mean pointing out that, when the harm occurred, it was dark outside or that the witness was far away from the scene—thus impairing her ability to correctly identify the color of the bus. Or, perhaps, the witness herself points out the risks in her testimony by admitting that she may be misremembering the color of the bus.

It is true that the two forms of evidence come in different frames, and it is up to the parties how to react. The parties are perfectly capable of contextualizing both kinds of evidence. With statistical evidence, it is up to the parties to demonstrate its pertinence to the case at hand while with testimony it is up to the parties to demonstrate its problems. That is the real difference between the types of evidence, and not considerations of sensitivity and incentives. In doing so, they may very well want to argue that certain evidentiary contexts may involve more sensitive conditions, and that may very well be persuasive to a fact finder. Were exegeses of sensitivity directed to advice about the presentation rather than the admissibility or categorical treatments of types of evidence, they would be useful. They are not, however, explanatory of the structure of American legal systems. Once again, weird hypotheticals have led astray.

I will give one last example of unrealistic and weird hypotheticals leading astray. In a recent essay, Martin Smith extends his analysis of «normalcy» from the episte-

mological realm⁵⁵ to legal systems⁵⁶. He, like others, is trying to defend the intuition that statistical evidence is different from non-statistical evidence and doing so by a search for the key distinction⁵⁷. His key is «normic support». The central idea is that «a body of evidence E *normically supports* a proposition P just in case the circumstance in which E is true and P is false would be less normal, in the sense of requiring more explanation, than the circumstances in which E and P are both true...»⁵⁸. Smith is quite cautious in his ultimate conclusions, conceding for example that the DNA cases present a «clash with the normic standard» that «could be seen as a reason for resisting the standard and seeking an alternative solution to the legal puzzle of statistical evidence»⁵⁹. And he concludes in a catholic vein that «If we are open, from the outset, to a range of different ways of managing uncertainty, the legal puzzle of statistical evidence need not strike us as a “puzzle” at all»⁶⁰. Perhaps this means that he accepts that normic support is just one of many tools to use in thinking about evidence, which would not be controversial. All of these different efforts to find the solution or key to the problem of legal proof fail as general explanations but are enlightening perspectives on proof that could very easily play a role in a particular case and are clearly encompassed by the field of evidence. An advocate would find studying all these different approaches useful indeed.

Although the central idea of normic support may be a useful heuristic, what is not terribly useful is the way Smith goes about explicating it, and once again there

⁵⁵ SMITH, 2010.

⁵⁶ SMITH, 2017.

⁵⁷ He worries about the distinction between «purely statistical evidence» and some other form, see *ibid.*: 1195 n. 3 and 1213 n. 19, but that is an unnecessary distraction. It is not clear what «purely statistical evidence» could possibly be, or at least what the alternative is that he worries about, and he does not clarify the matter. He raises a question whether a cold hit DNA case is «purely statistical in nature». *Ibid.*: 1213 n. 19. If it is not, then it is doubtful whether the set of «purely statistical cases» has any members, in which case the analysis of that set is rather pointless. On the other side of the coin, there will always be testimony about an evidentiary proffer that contextualizes it in some way. This is another example of the distracting quality of weird hypotheticals. They are offered as if a certain set of facts were stipulated, and thus incontestable, but that would never be the case in reality. There would have to be evidence, contestable evidence, of the operations of the rodeo for example in the gatecrasher scenario and it would have to be explained why that evidence allowed the inference to certainty that there were exactly 1,000 people in the stands of whom only 499 (or whatever) paid, an explanation that would never carry the day. This applies to DNA as well. See n. 74, *infra*. This, it should be noted, is an example of the general problem of impossible epistemological demands that these hypotheticals make. Certain things are just taken as true rather than as a matter of fallible evidence and inference, yet that is a world that does not exist. In the gatecrasher hypothetical, it would require that the defendant stipulate to all the facts so that they are not then a matter of proof, but then in addition somehow being unable to testify, or offer any other form evidence, that she actually bought a ticket. This combines weird hypotheticals and impossible epistemological demands into a truly remarkable brew the only implication of which is as a testament to the creativity of the human mind.

⁵⁸ *Ibid.*: 1208.

⁵⁹ *Ibid.*: 1214.

⁶⁰ *Ibid.*: 1216.

are the problems of weird hypotheticals, not attending to how trials actually function, and in this case not just the demand for too much knowledge but also subtle shifts in the epistemological demands that bake in the answer he is proposing.

Here is the hypothetical he uses to get the analysis off the ground:

Suppose an electronics store is struck by looters during a riot. 100 people walk out of the store carrying televisions, while the transaction record at the cash register indicates that only one television was paid for, though no receipt was issued. Suppose Joe is apprehended carrying a television from the store, but we have no other information about him. Should Joe be prosecuted for theft? Should he be convicted? Should he be punished? If our only evidence against Joe is that he carried a television from the store and that 99 out of the 100 televisions carried from the store were stolen, while one was legitimately purchased, then, by launching into such actions against Joe, we would do him a serious wrong⁶¹.

Problems begin at the beginning. What could it possibly mean that «the cash register indicates that only one television was paid for, though no receipt was issued»? Who would buy an expensive television set and not get a receipt? What television store would allow such a transaction? And the recurring problem of impossible epistemological demands—without the presentation of fallible evidence, how does anyone know how many people walked out with a stolen television set? Did the cash register that indicated only one television was paid for also keep track of how many were stolen? How do we know it is never wrong? Perhaps all this is fighting the hypothetical, but a scenario involving a riot and looting but with a clerk calmly ringing up one customer for their television purchase while chaos ensues around them is a far cry from the scenarios the legal system ever encounters. As is the requirement of infallible knowledge of the surrounding facts.

Nevertheless, once again let's take it at face value. It is false that the only evidence is what is presented. If one television set was purchased during the riot, someone had to be at the register unless we make the even weirder assumption that stores allow self-checkout with television sets but also don't video the encounter (the video would substitute for identification by a witness). Perhaps now we have to add the additional assumptions that the video wasn't working; the customer was honest enough that even though people were streaming out of the store with television sets, they decided to scan it at the self-checkout line; they were unconcerned with the chaos around them; and they did not care about a receipt, having apparently no concerns about proof of ownership in case something was wrong with the television.

Even if a series of adjustments is made to the hypothetical that make it weirder still but preserve the supposedly purely statistical nature of the case, there still remains further evidence, as there always will be in criminal cases—the accused themselves. The accused can testify to whatever they want, and thus no «serious wrong» would be done by charging them with a crime they almost surely committed and seeing how they defend. And if they did not make a satisfactory response, testify, or

⁶¹ *Ibid.*: 1196.

offer any other defense, a jury or judge would, as they should, convict and no appellate court would reverse it. Regardless of hallowed conventions that say that defendants do not have to defend, in fact they do. Everyone in the courtroom knows that at least one person present knows the truth, and that is the defendant. The facts of this hypothetical do more than provide a plausible story of guilt; it is an overwhelming case. The failure to testify in a hypothetical like this would confirm the truth of the prosecution's case, no matter what instructions are given about drawing no inferences from the exercise of Fifth Amendment rights.

The problems go deeper. How would any evidence of the facts of the hypothetical except that there had been rioting, looting, and Joe was caught carrying a TV set, get before the fact finder? Someone would have to testify to the sale of the TV set, a person who could then be asked if they recognized Joe as the person who purchased it. Ok, maybe this witness would just authenticate records kept by the cash register, or self-checkout machine, or whatever. Still, Joe was arrested, which would have been preceded by an on-the-street encounter between the police and Joe that would have entailed on-the-scene questioning. A refusal to answer would be admissible, as would any obviously concocted story Joe told, all confirming guilt. If Joe asserted he had paid cash (any other form of payment would create a record whether a receipt was issued or not) the police would ask where the money had come from, and other questions seeking to confirm or deny the truth—such as how much the television cost, a question a guilty person could probably not answer. To the jurist (at least to me), it is entirely unclear what to make of these matters precisely because they are so divorced from reality (although I will say that charging the hapless Joe with a crime is about as far from a «serious wrong» as is imaginable).

Like just about everybody, Smith relies on the Blue Bus hypothetical but there is a gap in its construction that illustrates the curious epistemological demands being made that generate the underlying intuition that in his view creates the need for some explanation like the notion of normic support on offer. He hypothesizes «90% of the buses operating in the area on the day in question are Blue-Bus buses»⁶² and a witness whose probability of testifying accurately is unstated⁶³. Usually the hypothetical assumes the same level of fallibility as the proportion of buses. Not asserting an error rate changes the hypothetical from prior uses of it in a subtle way. His point is that «it could just so happen that the bus involved was not a Blue-Bus»⁶⁴ but it could not «just so happen» that «the testimony was wrong» (which is his point in the television set hypothetical as well). There would need to be an explanation of what went wrong with the witness, and thus the witness, unlike the statistic, offers normic support for the verdict.

⁶² *Ibid.*: 1202.

⁶³ *Ibid.*: 1196.

⁶⁴ *Ibid.*: 1196.

By not asserting an error rate, Smith is exploiting the natural inclination and trial conventions discussed earlier, namely, that witnesses are to be believed unless a reason is given not to believe them. In a state of ignorance about the witness's testimonial capacity, a reasonable person has no reason to think that a witness is going to make the rather grievous mistake of identifying a red bus as a blue bus. He is thus probably right that «intuitions» would reflect that there is a difference between the two forms of evidence. Until someone introduces a concern about error, the comparison is between a known rate of error and a nonexistent rate of error, and of course the two look different. That outcome is baked into the hypothetical from the beginning. This is not instructive on the difference between statistical evidence and something else but simply the obvious result of trial conventions (don't litigate what does not need to be litigated).

Smith might object that he has not suppressed the possibility of error because part of his hypothetical involves a general recognition of witness fallibility and he goes so far as to say: «If we were forced to come up with some numerical estimate of how likely it is that the bus really was a Blue-Bus bus, given the witness testimony, it's doubtful that we would go quite as high as 90%—that would seem overly trusting. But 90% is, of course, precisely how likely it is that the bus involved was a Blue-Bus bus, given the statistical evidence about which we were so apprehensive»⁶⁵. But fact finders are not «forced» to come up with a number; nor is someone reading the hypothetical. As the hypothetical stands, it is simply exploiting the unstated assumption of witness credibility, which stands until controverted.

Smith, to the contrary, might say that is exactly his point; one case requires an explanation that another does not, but there is a looming problem here. Once the possibility of error is introduced, the two cases become essentially identical; the possibility of error will have a statistical distribution, even if unknown, structurally identical to the statistical evidence. There could be multiple explanations for the 10% inaccuracy rate, or whatever it is, but for purposes of comparing these cases it does not matter what they are. Suppose there is evidence that a witness lacks credibility and it is asserted that the witness has been known to lie. Still the question remains whether the witness is testifying truthfully in this case, and that is the same type of question as to what color bus hit the plaintiff. Unless we live in an indeterminate world, there will be reasons why the witness is untruthful in different sets of cases. Perhaps at all times like the one in question, after a night out with friends and distracted about thinking of her aging mother, the witness's identification is accurate in 90% of the cases. Only if one can pursue the witness's psychology adequately to get to certainty about what happened—a reference class of one, in other words—would this not be

⁶⁵ *Ibid.*: 1196. Why he thinks that this is overly trusting is a mystery. If a single unimpeached (on any ground) eyewitness testified to the color of a bus, that would not only suffice for a verdict absent evidence indicating that it was virtually impossible for that color of a bus to be in the vicinity, but if asked fact finders would probably indicate that they are close to certainty about what happened.

true, and reference classes of one at trial are rare to say the least. Thus, it «could just happen», indeed it does just happen, that the testimony is wrong 10% of the time. On the assumption that 90% of the accidents are caused by blue buses, the two are identical, which is partially explanatory as to why the «intuition» that these two cases are different does not describe American legal systems (see below).

Curiously, Smith seems to admit as much⁶⁶, but now there is an even larger looming problem. It is a rare case in which the possibility of error is not introduced. If there were no questions of credibility and no conflicting evidence, there would be no trial. Thus, the standard, virtually universal, condition at trial is that witness testimony does *not* provide normic support in his terms—in his view it would be identical to statistical evidence in the relevant ways. What Smith is proposing as a general distinction between types of evidence would seem to apply to virtually no evidence at trial⁶⁷. It begins to sound like a broken record, but the real distinction is not between evidence providing or not providing normic support but instead between evidence that carries its ambiguities on its face and that which does not. It is not that witness testimony provides «normic support»; rather, it is that we pretend that it does until someone wants to argue about it, and in virtually all cases the parties do want to argue about it. That is why they are expending their resources on a trial.

In fact, the similarities between the two types of evidence press even deeper and run in both directions. The hypothetical is designed to have the reader compare an immutable and inscrutable statistic—90% Blue Buses in the area—to a malleable and quite scrutable assertion about witness credibility. But just like the witness credibility evidence might call for further explanation, so too does the statistic—as laid out in detail above. Just like one might probe further into whether or why the witness made a mistake, one could probe further into the variables that predict bus accidents. Neither bus accidents nor witness errors in fact «just happen» in a causal world; there are reasons for both that sometimes can be explored further and sometimes cannot. If a particular witness, or witnesses generally, or . . . (note the frustrating ambiguity in specifying the reference class) give accurate testimony 90% of the time, it is no more abnormal for a witness to be mistaken than it is for a Red Bus to cause one of its 10% share of accidents (if that is what the hypothetical is supposed to be)⁶⁸. What deprives

⁶⁶ *Ibid.*: n. 14.

⁶⁷ For an analogous argument to that in the text, see DI BELLO, 2019. For a critique of Smith's arguments that is at least sympathetic to some of the arguments here, see MICHAEL BLOME-TILLMANN (2020).

⁶⁸ Another recent effort to explain juridical proof among other topics is Sara Moss's *Probabilistic Knowledge*. Her weird hypothetical, another twist on gatecrashers, is:

For instance, suppose you are the factfinder in Gatecrasher. The plaintiff proves that an arbitrary person at the rodeo is more likely than not to have climbed over the fence. Meanwhile, the defendant claims that he is an exception to this statistical generalization. He denies that it is more likely than not that he climbed over the fence, insisting that he did not climb over the fence at all. The statistical evidence of the case may justify your having .501 credence that the defendant is liable. But on the basis of this statistical evidence, your credence does not constitute knowledge. You do not know that the defendant is exactly as likely to have climbed over the fence as

an arbitrary person at the rodeo, because you cannot rule out a certain possibility that is inconsistent with this content—namely, that the defendant is an exception to the statistical generalization presented by the plaintiff. As a result, you do not know that it is more likely than not that the defendant is liable for trespass. According to my account of legal proof, it follows that the plaintiff has failed to prove that the defendant is liable by a preponderance of the evidence. Moss, 2018: 210-211.

Weirdness number one lies in the suggestion that an innovative theory is needed to explain this hypothetical. The defendant has first-hand knowledge and testifies that he bought a ticket (presumably that is what «claims» means). Unless quite seriously impeached (of which there is no evidence in the hypothetical), no one would find the defendant liable on these facts. There is nothing to explain.

Weirdness number two is that this argument implies that any time a defendant testifies to a defense, one cannot have whatever Moss thinks is necessary for a verdict. This does not just apply to cases involving statistical proffers of liability. As laid out in the text, a competent empirical study of blue bus accident causing likelihoods can be just as good or better justification for a verdict than an error prone eyewitness. If the equivalent or stronger evidentiary basis is defeated by a single defensive claim, so too in a rational world should be the weaker. In the nonstatistical case, perhaps evidence undermining the claim could be presented, but then it would have to be explained why equivalent evidence is not available in the statistical case. Once in a while that may be true, but categorically it is plainly false. Impeaching evidence is not limited to witnesses providing substantive evidence.

Weirdness number three lies in the implicit assumption that «knowledge», whether probabilistic or not, is the goal of a trial. It is not, at least not in any sophisticated philosophical sense. Instead, the goal is a common sense, naïve realist, folk psychological best effort to discern what happened under conditions of uncertainty. In an effort to contradict the previous sentence, I searched the pattern jury instructions of all 50 American states for references to knowledge as a criterion for a verdict. No state civil or criminal pattern jury instruction available online asks a jury to come to «knowledge». Instead, they call for the party with the burden to persuade to the appropriate standard. See e.g. RAJI (CIVIL) 6th Standard 2 («On any claim, the party who has the burden of proof must persuade you, by the evidence, that the claim is more probably true than not true.»). Alternatively, they call for the jury to weigh the evidence and decide which proposition is more likely true than the other. See e.g. IDJI 1.20.1 («When I[...]use the expression “if you find” or “if you decide”, I mean you must be persuaded that the proposition is more probably true than not true»). Jury instructions in criminal cases follow the same common sense approach. See e.g. 1 IN Pattern Jury Instructions Criminal Instruction No. 1.15 (2020) («The burden is upon the State to prove beyond a reasonable doubt that the Defendant is guilty of the crime(s) charged. It is a strict and heavy burden. The evidence must overcome any reasonable doubt concerning the Defendant’s guilt. But it does not mean that a Defendant’s guilt must be proved beyond all possible doubt.»). The only reference to knowledge found reminds jurors that «There are very few things in this world that we know with absolute certainty, and in criminal cases the law does not require proof that overcomes every possible doubt». Utah Pattern Jury Instruction.

Moss reiterates her knowledge argument in Moss, S. (2021), where she defends her knowledge account from what many if not all legal observers think is ineluctable uncertainty (including, I would say, uncertainty of the probability of possessing knowledge) by claiming that in criminal cases so long as jurors put aside «unreasonable doubts» then they can have «knowledge» that a defendant is guilty as charged. She operationalizes her account with, of course, another weird hypothetical involving a «wacky juror» who is under the influence of apparently equally wacky conspiracy theories, and other «far-fetched possibilities that jurors ought to set aside, such as possibilities involving evil mind-controlling demons». This is decidedly unhelpful. Surely there is universal agreement, quite independent of any particular epistemological commitments, that it is not a good idea to have wacky people and those who believe in evil mind-controlling demons serve as jurors or judges. More importantly, no mechanism to sort out wacky from not-wacky people is provided, neither is either the extension nor intension of the set with mind controlling demons as a member identified. How this argument advances any particular epistemological approach thus is opaque; rather, it seems to be an ad hoc move designed to insulate

the bus statistics of normic support deprives all evidence of it as well once one looks beneath the surface. It thus is not much of an explanation of the distinction between statistical and non-statistical evidence, nor more generally of juridical proof.

Put weird hypotheticals aside. What if it were true that the only evidence reasonably available, and it was reliable (a key point), is that there is a slightly better than .5 probability that the defendant is liable? The answer in the United States is that it has long been true that if there are no other sources of information, the statistical evidence meets certain conditions of reliability, and deciding the case one way or the other would not create cost/benefit justified incentives for socially useful behavior that outweigh the value of more accurate verdicts, verdict would likely, even if not always, be given for the plaintiff⁶⁹. As has been pointed out for some time⁷⁰, the Smith case is not in fact a case about statistical evidence but instead about lousy evidence⁷¹. Pace this «intuition» that courts are skittish about statistical evidence, based in part on a 75 year old case, in general the courts are receptive to statistical evidence and allow verdicts on that basis. In 2000, the Federal Judicial Center published a manual on the use of statistical evidence. The authors sum up the judicial experience with statistical evidence by saying that «Statistical reasoning can be crucial to the interpretation of psychological tests, toxicological and epidemiological studies, disparate treatment of employees, and DNA fingerprinting», and «this list could easily be extended»⁷². For example:

—Employment discrimination cases are often heavily if not exclusively statistical, as are antitrust cases⁷³.

an epistemological argument from its inherent impossible epistemological demands. Her argument is also inconsistent with the legal system she is trying to explicate. She appears to be arguing that the standard for a criminal conviction requires that the fact finder reach knowledge by putting aside unreasonable doubts and finding that all other possibilities of innocence have been ruled out. This is precisely what the standard was in the United States for convictions based on circumstantial evidence, but not otherwise: «where the Government's evidence is circumstantial it must be such as to exclude every reasonable hypothesis other than that of guilt». The Supreme Court of the United States rejected that standard as too stringent half a century ago in *Holland v. U.S.*, 348 U.S. 121 (1954), and virtually every state has now followed suit. The most recent court of last resort to deal with this issue is the Florida Supreme Court in *Bush v. State*, No. SC18-227, decided May 14, 2020. For a better explanation of reasonable doubt, see ALLEN and PARDO, 2019a.

Recently a number of efforts focusing on knowledge in addition to Moss' have attempted to identify what justifies the «intuition» that something is wrong with naked statistical evidence as exemplified by weird hypotheticals, yet there is no agreement on what that justification is. See, e.g., BLOME-TILLMANN, M., 2017; LITTLEJOHN, C., 2017; LEVANON, L., 2019.

⁶⁹ See, e.g., *Kramer v. Weedhopper of Utah, Inc.*, 490 N.E. 2d 104 (Ill. Ap. Ct. 1986). It would not be a surprise if some courts would want more than a > .5 burden to be met, but others would not. As I say, it is a large, sprawling system.

⁷⁰ ALLEN, 1987: 429, n. 67. See also ALLEN, 1991.

⁷¹ See, e.g., LEMPERT, 1986.

⁷² NATIONAL RESEARCH COUNCIL, 2000: 85.

⁷³ *E.E.O.C. v. O & G Spring & Wire Forms Specialty Co.*, 38 F.3d 872, 876 (7th Cir. 1994) («Reliance on statistical evidence by no means diminishes the plaintiff's obligation to prove discriminatory

—In criminal cases, there are cold hit DNA convictions⁷⁴.

—In *Daubert v. Merrill Dow Pharmaceutical, Inc.*⁷⁵, the Supreme Court included control studies with known error rates (statistical evidence) as part of the foundation for expert testimony, and in the follow-up to the Supreme Court case, the 9th Circuit held that liability could be based on a relative risk of 2/1—a direct statistical measure of liability⁷⁶.

—Class actions often depend on statistical evidence⁷⁷.

—Commercial protection cases⁷⁸.

—Intellectual property cases⁷⁹.

—Real estate cases involving property valuations⁸⁰.

—Tax cases⁸¹.

intent—but in some cases, statistical disparities alone may prove intent.») In re Sulfuric Acid Antitrust Litig., 446 F. Supp. 2d 910, 922 (N.D. Ill. 2006) (Rejecting defendant's motion to reconsider allowing plaintiff's expert testimony based on statistical sulfuric acid price data in an antitrust class action).

⁷⁴ For a discussion see ROTH, 2010. Enoch and Fisher attempt to explain away DNA evidence and cold hit cases on the basis of the enormously high probabilities that can be reached. Standing alone, those are the wrong probabilities. They need to be qualified by among other things the possibility of an innocent explanation of a suspect's DNA being found on the scene, police planting evidence, lab error, and the risk of adulteration of the sample. ENOCH and FISHER, 2015.

⁷⁵ 509 U.S. 579 (1993). See JONES and HAGTVEDT, 2002.

⁷⁶ *Daubert v. Merrell Dow Pharm., Inc.*, 43 F.3d 1311, 1320–21 (9th Cir. 1995) (Daubert II) (Because the background rate of limb reduction defects is one per thousand births, plaintiffs must show that among children of mothers who took Bendectin the incidence of such defects was more than two per thousand). In *Cooper v. Takeda Pharm. Am., Inc.*, 239 Cal. App. 4th 555, 593–94, 191 Cal. Rptr. 3d 67, 98 (2015), the Court made clear the statistical implications of the Daubert II holding:

When the relative risk is 2.0, the alleged cause is responsible for an equal number of cases of the disease as all other background causes present in the control group. Thus, a relative risk of 2.0 implies a 50% probability that the agent at issue was responsible for a particular individual's disease. This means that a relative risk that is greater than 2.0 permits the conclusion that the agent was more likely than not responsible for a particular individual's disease. [Reference Manual on Scientific Evidence (Federal Judicial Center 2d ed. 2000) (Ref. Manual),] Ref. Manual at 384, n. 140 (citing Daubert II).

⁷⁷ See, e.g., *Kurtz v. Kimberly-Clark Corp.*, 414 F. Supp. 3d 317 (E.D.N.Y. 2019) (Allowed hedonic regression analysis to show market wide price premium with respect to allegations of false labeling of products.)

⁷⁸ *United States v. Dish Network LLC*, 75 F. Supp. 3d 916, 939 (C.D. Ill. 2014) (Allowing statistical analysis of random sampling of call records to prove claim that satellite television provider violated the TCPA.)

⁷⁹ *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 423–24 (1984) (Relying on statistical evidence to hold that the Betamax is capable of non-infringing uses and thus acceptable for in-home use.)

⁸⁰ *CSX Transp., Inc. v. Bd. of Pub. Works of State of W.Va.*, 312 F. Supp. 2d 839, 851 (S.D.W. Va. 2004) (relying on statistical Chi-square analysis to determine true commercial and industrial property market value in West Virginia).

⁸¹ *Kikalos v. United States*, 308 F. Supp. 2d 902 (N.D. Ind. 2003) (Allowed government witness opinion, based on percentage mark-up method and government statistics, in liquor store owner's suit to recover overpayment of taxes).

—Market share liability⁸².

—Lost chance litigation⁸³.

The list above is disparate in various ways, but each example involves deciding a necessary element in a case to the extent possible based on statistics⁸⁴.

What some see as difficulties with American legal systems or puzzles that need resolution that they are trying to highlight with weird hypotheticals and solve with formal theories and algorithms do not seem very disturbing viewed through a naturalistic eye. One sees those systems from beginning to end channeling the parties in both civil and criminal cases to offer competing explanations for what happened, how it happened, and once in a while why it happened. The parties provide the evidence that triggers deliberation and suffer the consequences if the evidence is inadequate in the view of the fact finders. The fact finder is left to ponder over those competing explanations by reference to their own knowledge and experience, and perhaps to construct its own explanation of the underlying event. Fact finders employ their natural reasoning processes (including where appropriate probability), and a conclusion is reached as to the verdict⁸⁵. This is all facilitated by liberal rules of admissibility that make virtually any plausibly relevant evidence admissible; the absence of formal rules of sufficiency, which is left to the final judgment of the fact finder almost always; and little concern over what might be labeled «statistical» evidence. There are no algorithms operating here, nor any impossible computational demands made⁸⁶. The system as a whole and fact finders' deliberations are organic

⁸² *Sindell v. Abbott Laboratories*, 26 Cal. 3d 588 (1980).

⁸³ See FISCHER, 2001.

⁸⁴ The objection might be made that in many of these cases there is other evidence in addition to the statistical evidence, and thus that I may misunderstand the argument that focuses on naked statistical evidence. To the contrary, this highlights the unrealism of the discussion. There is no trial involving only «statistical evidence». See n. 57, *infra*. Someone has to present the evidence, authenticate whatever sources are being relied upon, face cross-examination and rebuttal, and so on. If there is a problem here it lies in statistical evidence being the necessary basis of decision. American legal systems have no difficulty with that phenomenon.

⁸⁵ See ALLEN and PARDO, 2019a. Given the actual state of American trials, it could not be otherwise. Evidence comes in serially, fact finders are instructed not to form conclusions or deliberate on the evidence until all the evidence has been heard, and the number of discrete pieces of evidence would overwhelm any computational capacity humans possess. See ALLEN, 1997. These points are true regardless of whether one is talking about objective or subjective Bayesianism. FRANKLIN, 2011. Moreover, there are no virtually no «objective» data points in a typical trial, and no one cares about the truly «subjective» beliefs of fact finders—their job is to determine the facts in a disinterested fashion. These debilitating problems with standard probability theory are some of the reasons behind the paradigm shift that is occurring from the conceptualization of juridical proof as probabilistic to explanatory.

⁸⁶ Including that trials are a search for knowledge, if I may refer to that as an algorithm; rather, there is an effort to construct a reliable decision making process subject to policy constraints. See *supra.*: n. 68. The two primary uses of «knowledge» (other than as elements of offenses and causes of action) are FRE 602's reference to «personal knowledge» and FRE 702's reference to expert knowledge. These references have quite a complicated relationship to «knowledge», as they authorize witnesses to testify in

and fluid rather than shoehorned into «probability», «evidential probability», «probabilistic knowledge», «belief states», «belief functions», «fuzzy logic», or whatever the theory of the day is. If there is shoe horning, it occurs as a result of party choice, not ontological necessity, and even then the fact finder is free to disregard the party's efforts and to conclude that some other explanation is the most plausible on offer.

The empirical description may be apt, but its moral foundation lacking, and intuition mining of weird hypotheticals that ignore the actual operation of legal systems and make impossible epistemological demands may so demonstrate. Perhaps, but there is one more reason for doubt, which ends this article on the same solipsistic note that it began. This literature has an annoyingly imperialistic capacity to refer not just to the various authors' intuitions, but «our» and «ordinary»⁸⁷ intuitions. Many of these do not coincide with mine⁸⁸, and it is interesting where the ones that do not are located. They are usually found in a priori analysis that involves searching for the single key or algorithm that resolves problems rather than engaging in empirical work⁸⁹. Testing intuitions against a naturalistic view of the process informed by an empiricist's slant, there is no conflict at all⁹⁰. American legal systems are not designed to facilitate intuition mining, guarantee that no wrongful decision favoring either party will be made, nor involve a search for «knowledge» in any but the most casual

the same trial to inconsistent propositions, both of which cannot be true, and therefore both of which cannot be knowledge under any conception of knowledge that requires truth.

⁸⁷ See, e.g., Moss, 2018: 217.

⁸⁸ For example, Moss seems to embrace the «intuition» that «one should avoid convicting innocent people at any cost». Moss, 2018: 215. If she does, I demur. Indeed, that «intuition» is a perfect example of the distance between formal reasoning and empirically informed inquiries. And common sense for that matter. There are on average over a million felony convictions each year in the United States, Bureau of Justice Statistics at <https://www.bjs.gov/index.cfm?ty=tp&tid=233>, and probably over 10 million misdemeanors convictions. See NATAPOFF, 2018; NATAPOFF, 2020: 159-162. Having studied the criminal justice process and the field of evidence for over forty years, my conclusion, for what it is worth, is that the only way to stop the convictions of innocents is to stop convictions of everybody, which would lead to social disaster. Such an intuition is an example of the failure to account for the deadly dilemmas of governing. ALLEN and SHAVELL, 2005; LAUDAN and ALLEN, 2008.

⁸⁹ See, e.g., the list of such views in Moss, 2018: 208-210. This also explains why criticisms that some of the explanations of relative plausibility do not meet what might be called philosophical standards of conceptual analysis miss the point. See, e.g. PUNDIK 2008: 36-40. As I have been arguing throughout this article, the methodology of conceptual analysis is not terribly helpful for explaining juridical proof, which requires instead a realistic embracing of complexity and adaptability. This is not to say that conceptual analysis may not serve other purposes in explicating aspects of juridical proof. As I have noted, for example, some of what are offered as general explanations could in fact be enlightening about intellectual tools to be employed along with others in structuring and deciding a case.

⁹⁰ That, of course, may appear to some as circular. I would draw a difference inference—that many of the a priori theorists might do well to spend more time focusing on whether their and what they think are «ordinary» untutored intuitions are justified rather than on how to justify them. Obviously, much of this article is compatible with much of social epistemology, as the original ALLEN and LEITER, 2001, article noted. On social epistemology, see FULLER, 2002, and GOLDMAN, 1999.

sense⁹¹. American legal systems give parties a civilized forum to get a definitive resolution of their dispute that could emerge from any corner of the sprawling human domain by a hopefully disinterested and unbiased fact finder that operates by reference to minimal, not maximal, standards of rationality, precisely because a priori one has no idea what a dispute may entail. It is up to the parties to construct the framework for the disputation, which is done of necessity idiosyncratically. Holding trials to the standards of most formal theories of evidence, proof, or knowledge would make them unworkable, and the costs of that outcome would greatly exceed its benefits, generating the very harms the «normative» analysis is attempting to minimize.

A search for «the» singular key to whatever problem is under investigation probably explains the use of reductive hypotheticals like the ones discussed. The reductionist efforts miss that American legal systems are complex adaptive processes, which is also why individual data points or the intuitions of particular individuals do not count for much. In a closed deductive and static system, inconsistencies are a problem. In complex dynamic ones, they are the order of the day. The fascination with *Smith v. Rapid Transit* is a good example. That is a single, old case, subject to various interpretations. Even if when decided it stood for the proposition that «statistical» evidence is insufficient for a verdict, so what? Other cases plainly say the opposite. In the Amazon rain forests, species come and go, just like in the American legal systems. There are not uniform intuitions that are the uncontested starting points for analysis. There is complexity everywhere. Even cases that may seem alike when presented in the context of weird hypotheticals may, in a real-world context, have quite different implications. The only way to know is to go look⁹².

In any event, the academics' intuitions concerning juridical proof are certainly not borne out by the real world. In fact, the resilience of this «intuition» is especially surprising⁹³ given the plentiful and directly contradicting evidence that statistical ev-

⁹¹ See n. 68, *supra*. Others have suggested this as well. ENOCH et al., 2019: 35.

Here is David Lewis, taking for granted the point we argue for in the text here (and elsewhere):

What if some far-fetched possibility is called to our attention not by a sceptical philosopher, but by counsel for the defence? We of the jury may wish to ignore it, and wish it had not been mentioned. If we ignored it now, we would bend the rules of cooperative conversation; but we may have good reason to do exactly that. (After all, what matters most to us as jurors is not whether we can truly be said to know; what really matters is what we should believe to what degree, and whether or not we should vote to convict). LEWIS, 1996: 560.

⁹² For an early example in the legal literature of discontent with intuition mining as the primary epistemological method, see ALLEN, R.J. and B. LEITER, 2001. See also ALLEN and LEITER (2003): 892: «The naturalistic turn in epistemology came about in no small measure because of the failure of formal theories (elaborated intuitions) to achieve the goal of justifying assertions of knowledge. Justification, in other words, seemed not to be a matter of just consulting one's intuitions, and the effort to justify assertions of knowledge in that manner came to be perceived as the problem rather than the solution. In order to gain knowledge, one needs not just to think but to investigate».

⁹³ With apologies to Mark Twain, that the death of statistical evidence has been greatly exaggerated has been known for over 35 years, yet the myth seems to persist in some circles. See, e.g., ALLEN, 1986: 429, n. 67.

idence, naked or otherwise, faces any serious problems in American jurisprudence⁹⁴. Of course, it is also possible that those who express this intuition are indifferent to what the real juridical world is and the game being played serves the purposes of other disciplines, in which case this article can undergo a timely and well deserved demise—its most basic point is that this form of intuition mining is not explanatory of juridical proof whatever other value it may have, and that what is required is a naturalistic approach to the problem. Or it could be that the description of the American legal systems is itself either ill-informed or, even if accurate, woefully misguided. I would be delighted if this analysis played a role in generating knowledge by stimulating a response justifying either conclusion. The job is not likely to be done with weird hypotheticals and their handmaidens.

How does any of this apply to the European or Spanish world systems? I hesitate to venture far into unknown terrain, but I will observe the discordance between the general European and Spanish scholars' heavy emphasis on pursuing «truth» in civil dispute resolution and the inefficiencies of many of those systems which, in the pursuit of truth, avoid making mistakes by making it difficult to decide cases—indeed in some cases essentially impossible. But not deciding cases is to decide them, as the status quo favors one party or the other; thus, interminable delay in civil cases perpetuates the very same mistakes that the pursuit of truth is supposed to suppress⁹⁵.

- ALLEN, R. J., 1986: «A reconceptualization of civil trials», in *Boston University Law Review*, 66(3): 401-437.
- , 1991: «On the significance of batting averages and strikeout totals: a clarification of the “naked statistical evidence” debate, the meaning of “evidence,” and the requirement of proof beyond a reasonable doubt», in *Tulane Law Review*, 65(5): 1093-1110.
- , 1997: «Rationality, algorithms and juridical proof: A preliminary inquiry», in *The International Journal of Evidence & Proof*, 1(special): 254-275.
- , 2013: «Taming complexity: rationality, the law of evidence and the nature of the legal system», in *Law, Probability and Risk*, 12(2): 99-113.

⁹⁴ Given that the intuition is inconsistent with reality, it is also not surprising that the American legal system is pretty much oblivious to the complicated efforts to justify the intuition. The only case that I know of that cites to the various efforts to domesticate the statistical evidence «problem» is an Israeli case authored by Alex Stein, a former American academic fluent in these debates who was recently appointed to the Israeli Supreme Court. See *The State of Israel v. Aviv Bezalel*, The Supreme Court of Israel sitting as the Final Court of Criminal Appeals, Criminal Appeal 2921/18 and Criminal Appeal 2961/18. For example, I could find no American opinion in Westlaw's case database that relied on or discussed epistemic safety, epistemic sensitivity, or evidential probability. Search terms included, among others: «evidential probability», «epistemic probability», «“epistemic” & “sensitivity”», «epistemological safety», «epistemological /10 safety», «“epistemological” & “safety”», «epistemological sensitivity», «epistemological /10 sensitivity», «“epistemological” & “sensitivity”», and «epistem! & (safety /30 sensitivity)». The entire set of search strings is available from the author.

⁹⁵ See ALLEN, 2015.

- , 2015: «The Perils of Comparative Law Research», published in Spanish as «Los Peligros De Investigación En Derecho Comparado», in Jordi Ferrer Beltran & Carmen Vazquez (eds.), *Debatiendo con Taruffo*. Marcial Pons, 2016: 21-40.
- , 2020: «Debate: Legal Probabilism—A Qualified Rejection: A Response to Hedden and Colyvan», in *Journal of Political Philosophy*, 28(1): 117-128.
- ALLEN, R.J. and B. LEITER, 2001: «Naturalized epistemology and the law of evidence», in *Virginia Law Review*, 87(8): 1491-1550.
- , 2003: «Naturalized epistemology and the law of evidence: Reply to Redmayne», in *The Michigan State Law Review* 2003(4): 885-892.
- ALLEN, R.J. and M.S. PARDO, 2019a: «Relative plausibility and its critics», in *The International Journal of Evidence & Proof*, 23(1-2): 5-58
- , 2019b: «Clarifying relative plausibility: A rejoinder», in *The International Journal of Evidence & Proof*, 23(1-2): 205-217.
- ALLEN, R. J. and A. SHAVELL, 2005: «Further Reflections on the Guillotine», in *The Journal of Criminal Law and Criminology*, 95(2): 625-636.
- BENNETT, W. and M. FELDMAN, 1981: *Reconstructing Reality in the Courtroom*. London: Tavistock Feldman.
- BLOME-TILLMANN, M. 2002: «Statistical Evidence, Normalcy, and the Gatecrasher Paradox», *Mind* 123(514): 563-577.
- , 2017: «More Likely Than Not' Knowledge First and the Role of Bare Statistical Evidence in Courts of Law», in Carter, A., Gordon, E. & Jarvis, B (eds.), *Knowledge First - Approaches in Epistemology and Mind*, Oxford, UK: Oxford University Press: 278-292.
- CALLEN, C. R., 1982: «Notes on a Grand Illusion: Some Limits on the Use of Bayesian Theory in Evidence Law», *Indiana LJ*, 57(1): 1-44.
- , 1991: «Kicking in rocks with Dr. Johnson: a comment on Professor Allen's theory», in *Cardozo Law Review*, 13(2, 3): 423-444.
- COHEN, L. J., 1977: *The Probable and the Provable*, Oxford University Press.
- , 1981a: «Subjective probability and the paradox of the gatecrasher», in *Arizona State Law Journal*, 1981(2): 627-634.
- , 1981b: «Can human irrationality be experimentally demonstrated?», in *Behavioral and Brain Sciences*, 4(3): 317-331.
- DI BELLO, M, 2019: «Proof paradoxes and normic support: socializing or relativizing?», in *Mind*, 128(512): 1045-1084.
- DIAMOND, S., B. MURPHY, M. ROSE, 2012: «The kettleful of law in real jury deliberations: Successes, failures, and next steps», in *Northwestern University Law Review*, 106(4): 1537-1608.
- EDWARDS, W., 1986: «Summing up: The society of Bayesian trial lawyers», in *Boston University Law Review*, 66(3): 937-41.
- ENOCH, D., SPECTRE, L, FISHER, L., 2012: «Statistical evidence, sensitivity, and the legal value of knowledge», in *Philosophy & Public Affairs*, 40(3): 197-224.
- ENOCH, D. and T. FISHER, 2015: «Sense and sensitivity: Epistemic and instrumental approaches to statistical evidence», in *Stanford Law Review*, 67(3): 557-611.
- ENOCH, D. and L. SPECTRE, 2019: «Sensitivity, Safety, and the Law: A Reply to Pardo», in *Legal Theory*, 25(3): 178-199.
- FISCHER, D. A., 2001: «Tort recovery for loss of a chance», in *Wake Forest Law Review*, 36(3): 605-655.
- FRANKLIN, J., 2011: «The objective Bayesian conceptualisation of proof and reference class problems», in *Sydney Law Review*, 33(3): 545-561.
- FULLER, S., 2002: *Social Epistemology*, Indiana University Press.
- GARBER, D. and S. ZABELL, 1979: «On the emergence of probability», in *Archive for History of Exact Sciences*, 21(1): 33-53.
- GARDINER, G, 2020: «Profiling and proof: are statistics safe? », in *Philosophy*, 95(2): 161-183.
- GOLDMAN, A., 1999: *Knowledge in a Social World*, Oxford University Press.
- HACKING, I., 2006: *The Emergence of Probability (2nd ed)*, Cambridge: Cambridge University Press.

- Hedden, B and Colyvan, 2019: M, «Legal probabilism: a qualified defence», in *Journal of Political Philosophy*, 27(4): 448-468.
- JONES, G. T. and R. HAGTVEDT, 2001: «Sample data as evidence: Meeting the requirements of Daubert and the recently amended federal rules of evidence», in *Georgia State University Law Review*, 18(3): 721-748.
- KAPLAN, J., 1967: «Decision theory and the factfinding process», in *Stanford Law Review*, 20(6): 1065-1092.
- KAYE, D., 1979a: «Naked statistical evidence», in *Yale Law Review*, 89(3): 601-611.
- , 1979b: «The law of probability and the law of the land», in *University of Chicago Law Review*, 47(1): 34-56.
- , 1979c: «The paradox of the gatecrasher and other stories. », in *Arizona State Law Journal*, 1979 (1): 101-110.
- , 1982: «The limits of the preponderance of the evidence standard: Justifiably naked statistical evidence and multiple causation», in *Law & Social Inquiry*, 7(2): 487-516.
- LAUDAN, L. and R. J. ALLEN, 2008: «Deadly dilemmas», in *Texas Tech Law Review*, 41(1): 65-92.
- LEITER, B., 2011: «Naturalized jurisprudence and American legal realism revisited», in *Law and Philosophy* 30(4): 499-516.
- , 2020: «What is a realist theory of law? », Available at SSRN: <https://poseidon01.ssrn.com/delivery.php?ID=658002065071097112110111067081073098030040014018086061111066121075071096018068090030030053038123014014001023069117079127108106046002025038074087123121067123085125067034019009003020006094099026107110025007072096093073113087068112113094127121127113084&EXT=pdf>.
- LEMPERT, R., 1986: «The new evidence scholarship: Analyzing the process of proof», in *Boston University Law Review*, 66(3): 439-477.
- LEVANON, L., (2019): «Statistical Evidence, Assertions and Responsibility», in *Modern Law Review* 82(2): 269-292
- LEWIS, D., 1996: «Elusive knowledge», in *Australasian journal of Philosophy*, 74(4): 549-567.
- LITTLEJOHN, C., 2017: «Truth, knowledge, and the standard of proof in criminal law», *Synthese* published online at <https://link.springer.com/content/pdf/10.1007%2Fs11229-017-1608-4.pdf>.
- MOSS, S., 2018: *Probabilistic Knowledge*, Oxford: Oxford University Press.
- , 2021: «Knowledge and Legal Proof», forthcoming in *Oxford Studies in Epistemology* vol. 7, Oxford: Oxford University Press.
- NANCE, D A., 2016: *The Burdens of Proof: Discriminatory Power, Weight of Evidence, and Tenacity of Belief*, Cambridge University Press.
- NATAPOFF, A., 2018: *Punishment without Crime: How our Massive Misdemeanor System Traps the Innocent and Makes America More Unequal*, Basic Books.
- , 2020: «Atwater and the misdemeanor carceral state», *Harvard Law Review Forum* 133(6): 147-178.
- NATIONAL RESEARCH COUNCIL, 2000: «Reference Guide on Statistics», in *Reference Manual on Scientific Evidence* (2nd ed), National Academies Press.
- PARDO, M.S., 2005: «The field of evidence and the field of knowledge», in *Law and Philosophy* 24(4): 322-392.
- , 2013: «The nature and purpose of evidence theory», in *Vanderbilt Law Review* 66(2): 547-613.
- , 2018: «Safety vs. sensitivity: Possible worlds and the law of evidence», in *Legal Theory*, 24(1): 50-75.
- PARDO, M. S. and R. J. ALLEN, 2008» «Juridical proof and the best explanation», in *Law and Philosophy* 27(3): 223-268.
- PARK, R. C. and M. J. SAKS, 2005: «Evidence scholarship reconsidered: results of the interdisciplinary turn», in *Boston College Law Review*, 47(5): 949-1031.
- PILLER, C., 2019: «Beware of safety», in *Analytic Philosophy*, 60(4): 307-335.
- POSNER, R. A., 1999: «An economic approach to the law of evidence», in *Stanford Law Review*, 51(6): 1477.
- PRIEL, D., (2020), «Evidence-based jurisprudence: An essay for Oxford», *Analisi e Diritto* 2020, available at <https://poseidon01.ssrn.com/delivery.php?ID=59411307001311511901112408812003107800>

- 206906708700605009306701503000306910311411909309803101111900102904410807712
411608608903105505702803201300401711211212309306601405802806506411309806911
3021099123101092097111121099113027007007008079100015025081065&EXT=pdf.
- PUNDIK, A., 2008: «What is Wrong with Statistical Evidence? The Attempts to Establish an Epistemic Deficiency», Available at SSRN: <https://ssrn.com/abstract=1134656> or <http://dx.doi.org/10.2139/ssrn.1134656>.
- , 2011: «The epistemology of statistical evidence», in *The International Journal of Evidence & Proof*, 15(2): 117-143.
- RIBEIRO, G., 2019: «Relevance, probative value, and explanatory considerations», in *The International Journal of Evidence & Proof*, 23(1-2): 107-113.
- ROTH, A., 2010: «Safety in numbers? Deciding when DNA alone is enough to convict», in *New York University Law Review*, 85(4): 1130-1185.
- SANCHIRICO, C. W., 2001: «Character Evidence and the Object of Trial» in *Columbia Law Review*, 101(6): 1227-1311.
- SMITH, M., 2010: «What else justification could be? », in *Nous*, 44(1): 10-31.
- , 2018: «When does evidence suffice for conviction? », in *Mind*, 127(508): 1193-1218.
- VIDMAR, N. and V. P. HANS, 2007: *American Juries : The Verdict*. Amherst, N.Y.: Prometheus Books.
- WAGENAAR, W. A., et al., 1993: *Anchored Narratives: The Psychology of Criminal Evidence*, St Martin's Press.
- WEINBERG, J. M., NICHOLS, S., STICH, S., 2001: «Normativity and epistemic intuitions», in *Philosophical Topics* 29(1/2): 429-460.
- WELLS, G. L., 1992: «Naked statistical evidence of liability: Is subjective probability enough?» in *Journal of Personality and Social Psychology*, 62(5): 739-752.

HOW TO THEORIZE ABOUT STATISTICAL EVIDENCE (AND REALLY, ABOUT EVERYTHING ELSE): A COMMENT ON ALLEN

David Enoch

The Philosophy Department and the Faculty of Law

The Hebrew University of Jerusalem.

David.Enoch@mail.huji.ac.il

ABSTRACT: In responding to Prof. Allen's paper, I make several general methodological points: about the use of hypothetical cases, about the point of theorizing, and about the role of idealization. Then I make some more specific points about his claims about (and against) previous work on statistical evidence.

KEYWORDS: legal epistemology; statistical evidence; legal theory.

SUMMARY: 1. SEVERAL ISSUES SET ASIDE.—2. WHAT'S THE POINT?.—3 WEIRD HYPOTHETICAL CASES.— 4. MORE ON STATISTICAL EVIDENCE: 4.1. How *does* the law treat statistical evidence?. 4.2. Sensitivity. 4.3. Incentives.—5. CONCLUSION.—6. BIBLIOGRAPHY.

RECOMMENDED CITATION: ENOCH, DAVID, 2020: «How to Theorize about Statistical Evidence (and Really, about Everything Else): A comment on Allen», in *Quaestio facti*, 2: 285-298. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22455

I want to thank Ronald Allen for giving such close attention to my and my co-authors' work on statistical evidence¹. I also want to start by noting how successful his piece is in at least one of his stated goals—namely, to be provocative.

Revisiting the central theme from ALLEN AND LEITER (2001), Allen insists that internalizing the lessons from naturalized epistemology—presumably, primarily against a priori reasoning and for more empirically minded research of, well, everything—undermines a whole body of literature in the theory of evidence law: discussions of statistical evidence, and what, if anything, makes this kind of evidence more suspicious than other, direct, non-statistical evidence. Allen is strongly critical of a priori reasoning about this and related problems, the kind of reasoning that is not sensitive to the dynamic and complex nature of the American legal system, and that is of no help to the jurist. Central to his arguments is a sustained criticism of relying on «weird hypotheticals»—cases from which nothing of value can, he seems to think, be learned. Though my and my co-authors' work on these matters is unsurprisingly especially dear to my heart, I think it's safe to say that it plays merely the role of a (central) example in Allen's paper, which seeks to establish a more general conclusion, namely, that at least in so far as we're interested in helping the jurist understand the American legal system, this kind of research program is entirely misguided.

Here's my plan for this comment. In the first section, I flag several claims that I will not be taking issue with here. In section 2, I comment on what I take to be the point of theorizing about statistical evidence—the answer I give to this question seems to be very different from the one Allen takes for granted. In section 3, I discuss specifically Allen's rejection of «weird hypotheticals», a discussion that leads to a more general one, of the role of idealization in theorizing. These two sections are, then, both methodological, and entirely general—they do not amount to a specific defense of my own views about statistical evidence (though they're relevant for such a defense, of course). Indeed, they're even more general than that: the methodological picture I sketch applies, as far as I can see, to any *theorizing* effort, and because I think it is almost obvious, I wouldn't have bothered the reader with it if Allen's text hadn't so clearly called for it. In section 4, I quickly comment on some more specific points Allen makes about our views of statistical evidence.

1. SEVERAL ISSUES SET ASIDE

I will not be using the phrase «naturalizing epistemology». This phrase—as Allen notes—has come to mean different things, perhaps (he doesn't say *that*) to the

¹ The relevant papers are all co-authored, with either Talia Fisher, or Levi Spectre, or both. See ENOCH, SPECTRE and FISHER (2012), ENOCH and FISHER (2015), and ENOCH and SPECTRE (2019). I thank Tali and Levi also for reading and commenting on this paper. I also thank a reader for *Quaestio Facti*.

point where it's no longer helpful. When he explains what he has in mind with that phrase he mentions two central commitments: first, a commitment to something like «ought-implies-can», the relation of which to standard philosophical discussions of naturalizing epistemology I do not see (and a commitment which, for the most part, and after suitable qualifications, I share); and second, something about «the difference between a priori conceptual/normative analysis on the one hand and on the other empirical inquiry into the actual state of some phenomena». I am going to set aside general worries—not just about what naturalizing epistemology is, but also about whether we *should* naturalize epistemology (for the record, I think that the only senses in which we should are senses that render this claim trivial), and indeed about the relevant nature of Allen's project (he seems to *recommend* abandoning a priori normative discourse, seemingly equally enthusiastic both for rejecting normative discourse and for engaging it, making recommendations, etc.). As far as I can see, all the interesting topics in Allen's paper can be stated and discussed without the catch-phrase (in some circles) «naturalizing epistemology».

Second, Allen is out to reject a probabilistic conception of legal evidence (though he also seems to think it's a dead horse he's beating, talking of «the demise of the probabilistic explanation»). I am not defending such a conception, and I am certainly not making any claims about the history of evidence law and evidence law theory. I do think—how can one not?—that something about probabilities is often very relevant to any plausible conception of evidence, so that dismissing the relevance of the probability calculus entirely when commenting on legal evidence is unjustified. (Some of what Allen says seems to pull in that direction). But I am certainly not committed to the claim that the probability calculus is the only game in town².

Third, I will not be commenting here on «relative plausibility theory», the alternative view of legal evidence that Allen sketches here (and develops in detail elsewhere). While I have some worries about it³, I don't know it well enough to comment seriously about it, and my focus in this comment lies elsewhere.

2. WHAT'S THE POINT?

What are we after, when we're theorizing about statistical evidence? What is *the point* of such theorizing? There needn't be a single, privileged, canonical point, of course. Different theorists may be after different things. But Allen is explicit about what it is that he is after. He is after «...understanding juridical proof from the inside...». And while this wording may not be sufficiently clear, when he criticizes some of his targets Allen is clearer, saying such things as that the target views fail

² It's not clear to me how exactly Allen is using the term “reductionism”, but what I say in the text here may amount to agreeing with him about rejecting reductionism, or a specific version thereof.

³ Chief among them is the thought that talk of relative plausibility amounts not to a solution to any problem, but rather to renaming it.

«... if the objective is to understand the actual state of the subject of the inquiry—to help us jurists in other words». But while helping people—jurists or otherwise—is often a good idea, I fail to recognize *my* aim in these words.

When I theorize—about statistical evidence, or about anything else—the aim is to reach knowledge and understanding of important relevant truths. The aim, in other words, is entirely epistemic. The questions I ask about some suggested theory of statistical evidence are questions like: Is it true? Is it enlightening? Does it promote insight into the nature of statistical evidence, and indeed evidence more generally? Are the arguments supporting it convincing? Does the theory have problematic implications elsewhere in our system of beliefs? And while it may be nice if an American⁴ lawyer or judge somewhere shows interest in my paper, this is not an adequacy constraint on a theoretical paper, and it certainly isn't *the point* of such theorizing.

The water here may be muddied by Allen's use of the word «jurist». «Jurist» may mean anyone thinking about the law. In this sense, jurists should of course be interested in greater knowledge and understanding, and so in this sense, theorizing (in my sense) does help jurists. In another sense⁵, though, «jurist» just means *the practitioner*, a lawyer or a judge. Practitioners too, however, may be interested in greater understanding—not necessarily for understanding's instrumental payoffs, if there are any, but because practitioners too may be reflective and sophisticated, and if they are, a good theory (one that facilitates knowledge and understanding) *does* help them. The only kind of jurist for whom such understanding is nothing is the narrow-minded, what's-in-it-for-the-way-I-argue-my-next-case kind of practitioner⁶. I fully acknowledge that what I say about statistical evidence may not give the

⁴ Throughout the paper, Allen talks specifically about the *American* legal system (if my app is right, the word «American» appears in his paper 27 times). This seems to me like the academic analogue of American exceptionalism—a less dangerous exceptionalism than many others, of course, but not a nicer or better-supported one. A theory of statistical evidence is—at least if all goes well—an attempt to understand statistical evidence in general, not in some specific jurisdiction. Even if it's about helping jurists, it need not be about helping American jurists in particular. This American exceptionalism is especially troubling seeing that a central target for Allen here is work by three Israelis (though one with dual Israeli-American citizenship), published mostly in international journals; and seeing that Allen's paper (like this comment on it) is coming out in a journal published by a Catalan university in Spain, which states on its website: «What *Quaestio facti* is not: from the outset, it is important to emphasize that it is not a journal centered around the evidence law of one country or another. This is why we only publish original works that touch upon general problems related to legal evidential reasoning». See <https://www.quaestiofacti.com/journal-evidential-legal-reasoning/?lang=en>

Allen's American exceptionalism is rendered significantly better, I think, neither by the short disclaimer about other countries at the end of his paper nor by his willingness to describe his own paper as solipsistic.

⁵ At least one online dictionary defines «jurist» thus: «1. An expert in or writer on law. 2. North American: Lawyer or judge». <https://www.lexico.com/en/definition/jurist>

⁶ At times, it seems very clear that Allen has this kind of practitioner in mind. Consider (from footnote 94): «For example, I could find no American opinion in Westlaw's case database that relied on or discussed epistemic safety, epistemic sensitivity, or evidential probability». This is interesting—many

narrow-minded practitioner anything they are after. I fail to see, though, how this amounts to any kind of objection to my theory, or to the general communal theorizing effort of which it is a part.

Does insisting on the epistemic, theoretical point of theorizing—rather than on assisting the narrow-minded practitioner—show that we theorists are not really interested in the real world? Of course not. It is the real world that we are theorizing *about*. Consider Allen's example of the Amazon River, whose dynamic nature he uses as an example for the dynamic nature of (American) law, purportedly a problem for my kind of theorizing about it. We theorize about the Amazon River all the time, and when we do, it seems to me, we try to reach greater knowledge and understanding of important truths about it (and about related phenomena). The Amazon River is the topic of such theorizing, and obviously, scientists studying it are deeply interested *in it*. That the river doesn't care about such theorizing is really neither here nor there. Analogously, if you are an evidence law theorist, you should of course pay close attention to legal practice, as it is, after all, the object of your inquiry (more about the relation between such attention and idealization below). But if your theory does well in epistemic terms, then that it may not be of use to the narrow-minded practitioner is not more a failure of your theorizing than it is a failure of a scientific study of the Amazon River that the river doesn't care⁷.

What about the dynamic, ever-changing nature of both the Amazon River and (American) law? Of course these objects of inquiry are dynamic, and of course a good theory of them will have to take that into account, to accommodate—and explain, and supply deeper insight into—the relevant changes. But from this obvious (and uncontroversial, I think) observation nothing follows about the nature of the sought-after theory itself. Indeed, very often good theorizing amounts precisely to finding patterns and regularity in the flux. The planets are in motion, but Kepler was after the unchanging laws that explain their motion. Even if at the end of the day he failed, surely he can't be criticized for failing to accommodate the dynamic nature of the planets. The same goes for the Amazon River, and indeed, for dynamic social systems, like the law (American or otherwise). We may, of course, fail: perhaps there is less regularity than we may have thought (and hoped). Or perhaps some posited pattern fails to accommodate the flux. But, first, all of this needs showing—with the planets, the Amazon, and the law alike—and second, such a failure will always be a price. We will seek the most general and stable theory we can find, or in other words, as much regularity in the flux as we can get. That (to an extent) a priori, normative reasoning searches for a regularity in the flux that is the law is not, then, a bug but a feature. It's a part of what theorizing *is*.

things are. But how is this a problem for the target views, unless Allen takes it to be a desideratum to give the narrow-minded practitioner something he or she wants?

⁷ Sometimes, such science will also have pragmatic payoffs, and when it does, this is great. But its value as an instance of theorizing in no way depends on whether this is so.

3. WEIRD HYPOTHETICAL CASES

One of Allen's main complaints about the literature on statistical evidence is its reliance on «weird hypotheticals», like the famous Gatecrashers and Blue Bus cases. The nature of his dissatisfaction with such cases is not entirely clear to me, but it is very clear that it is grounded in the fact that such cases are too remote from actual legal practice. This is presumably why, when discussing such cases, Allen time and time again draws attention to the complexities of real life, and to the unrealistic nature of hypothetical cases ignoring them. Why is it, he asks, that only the statistical evidence is present (as is often stipulated)? How come no one in the Gatecrashers scenario keeps receipts? Who the hell represents the relevant parties, and why are they so negligent? And so on.

But what *exactly* is the problem that Allen sees with relying on hypothetical cases, and why is it that he thinks it's ok, in this case (but perhaps not in others) «to reject» or «to fight» the hypothetical? More than one thing, I think, is going on.

At one point (footnote 14), Allen refers to doubts coming from experimental philosophy about the use of intuitions in general, and in particular to empirical doubts about the ubiquity of Gettier intuitions in epistemology. But I don't think this is Allen's central worry, because he doesn't pursue it in detail. (Also, this worry—if cogent—applies to *any* use of intuitions, about pretty much all kinds of cases. It's not entirely clear to me whether this is a result Allen welcomes.) Myself, I am not enthusiastic about such empirical work (see, for instance, NAGEL 2012). But for the fans, it may be significant to point out that even the high priests of experimental philosophy no longer seem to go for the simple view Allen is citing here (see MACHERY, STICH, ET AL 2015).

At another point, Allen complains: «This literature has an annoyingly imperialistic capacity to refer not just to the various authors' intuitions, but “our” and “ordinary” intuitions. Many of these do not coincide with mine, and it is interesting to see where the ones that do not are located». Many things are interesting, including studying the causal history of intuitions (which is, I think, what Allen refers to at the end of the quoted text). But Allen seems to be confused about how it is that relying on intuitions is supposed to be done in a philosophical discussion. When I rely on something being intuitively thus-and-so, it is of course only my own intuitions that I am relying on. If I use an annoyingly imperialist «we» or «our», this is just an invitation to my reader or interlocutor to think about the matter, and see if they are with me. No sanctions follow if they are not. And the probative value I assign to intuitions is always limited—they are not supposed to be a direct route to infallible truths. Rather, they are good starting points. If two theories T1 and T2 are equally good (as theories) in other respects, and T1 is and T2 is not consistent with an intuition I have, this gives me a reason to go with T1 rather than T2. This reason may be outweighed by competing reasons, and furthermore, if you don't share my intuition, you don't have that reason. Myself, because I don't do empirical work (and, as you

already know, am not a fan of the work of this kind that I know), I will not often commit one way or another regarding «the folk's» intuitions, or «ordinary» ones. (For the record, though – there's apparently some empirical evidence showing that «ordinary» intuitions are on our side, not Allen's⁸.) I don't think there is anything special about this paragraph's relying-on-intuitions-101: I take it this is what all of us do, all the time, across subject matters and disciplines (even if we don't always call what we're doing by that name).

But to get to the heart of Allen's problem with weird hypotheticals, I suggest that we insist on receiving a direct, explicit, clear response to one of these cases. Faced with a weird hypothetical case, then—a case in which many factors are cleaned out by stipulation, two compared cases regarding which we use such phrases as «all other things are equal», and so on—what is his response? Take a version of the Blue Bus case, and compare two cases, such that in one there's only the statistical evidence (say, about market share), and in the other only non-statistical evidence (say, an eyewitness, or a video⁹). And I am now stipulating that all other things are equal, that there isn't any additional information, etc. The case is unrealistic, granted. But can we get a straight answer, please, about *this* case? Here are the main options I can think of:

- (i) The case is incoherent.
- (ii) In that unrealistically clean case, there is no important difference between the statistical evidence and the non-statistical evidence.
- (iii) In that unrealistically clean case, there is still an important difference between the statistical evidence and the (better) non-statistical evidence.

Let's see, then: had the case been strictly speaking incoherent (i), Allen would have been entitled to reject our insistence on a (further) answer. But it's hard to see why we should think it's incoherent – where's the *p*, such that somewhere hidden in the description of the case there's both *p* and (something that implies) *not-p*? At the very least, Allen has not established such incoherence. But if the case is coherent, it seems, we are entitled to an answer as between (ii) and (iii)¹⁰—to repeat, about the case *as described*. How does Allen respond? At times it seems that he is mostly impatient with such cases, that he does not *care* about them. But this can't be the whole story—why should we care, for instance, that he doesn't? We really are owed a straight answer.

⁸ There's some empirical literature on the Wells Effect. See—addressing specifically our 2012 statistical evidence paper—PINILLOS, JARAMILLO AND HORNE (2019).

⁹ At times, Allen writes as if the only non-statistical evidence is witness testimony. («The real difference between statistical evidence and testimony lies in the trial convention of assuming reliability of witnesses until credibility is attacked. This convention obviously does not apply to statistical evidence, ...») This is not so, of course.

¹⁰ Perhaps Allen may insist that the case, like most cases, is under-described, and that different ways of filling in the details will have us landing on option (ii) or (iii) in the text. But this too has to be shown—what details are missing, and is there a coherent way of filling them in in a way that preserves the all-other-things-equal restriction? If so, what's Allen's straight answer about *that* new case?

If Allen goes with (iii), it's pretty much game over. If the intuitive difference between statistical and non-statistical evidence survives the artificial cleaning out of all other factors, we are still in need of an explanation of this difference, and then all the other factors—the ones Allen likes relying on—cannot play a role in the needed explanation.

So I think Allen is best read as going with (ii)¹¹. What, then, of the intuitions—*my* intuitions, that is, and my co-authors', and yours (no?), etc.—that seem to pull so strongly in the opposite direction? I think—but it's hard to tell—that Allen is putting forward *debunking explanations* of such intuitions; that is, he is putting forward explanations of why it is that we have those intuitions, that show them in an unflattering epistemic light, that show why we would have those intuitions even if they're misleading. This is a perfectly legitimate move, of course—and an entirely common one in the a priori reasoning genre Allen likes to say he rejects—but it doesn't come for free. It's not enough to gesture in the debunking-explanation direction. Rather, those pushing this line have to do the work. What are the details of the debunking explanation? Just how debunking is it? How is it a better explanation of the relevant intuitions compared to more vindicating ones? And having considered the suggested debunking explanation, do the purportedly debunked intuitions lose their force? Or do they retain some immediate force, but one that is lost on reflection (perhaps as perceptual impressions do in some optical illusion cases)? It's not impossible that all this work can be done. But Allen has (so far) done very little of it.

Allen insists that the Blue Bus hypothetical «contributes little of interest about the nature of juridical proof». Perhaps this explains his impatience—why even bother giving a straight answer to a question that can teach us so little, if anything at all. Perhaps he can insist, then, that he doesn't need to decide as between (ii) and (iii). He may even be willing to accept (iii), but insist that nothing follows from this artificial case with regard to the real-world legal phenomena. This worry raises important questions that I can only briefly address here¹².

There are, in the real world, no frictionless surfaces. But no one—I think—thinks that Newtonian mechanics with frictionless surfaces is of no value, or even that it doesn't teach us anything about the real world. Some idealizations, some models, are of value, even of value in explaining the real world, despite involving such fictions as frictionless surfaces. On the other hand, if you're doing aerodynamics, then assuming away air-resistance is not an acceptable move—by assuming away air-resistance, you aren't simplifying the questions we were asking ourselves in aerodynamics. Rather, you're assuming those questions away. The difference between the two cases, I suggest, is that in the case of Newtonian mechanics a lot of its core insights and value

¹¹ «In comparing the statistical evidence of bus distributions to testimonial evidence, the bus distribution data (assuming it maps perfectly onto accident rates) is at least as good as the testimonial evidence».

¹² I discuss them in more detail in the appendix on ideal and non-ideal theory in my «Against Utopianism: Noncompliance and Multiple Agents» (2018). Also see the references there.

survive even if all surfaces are assumed to be frictionless. When one understands something in a system with frictionless surfaces, one gains an understanding into (a partial picture of) the real world as well. Of course, one can't then proceed to apply one's conclusions or calculations (or vehicle designs) from the frictionless model to the real world as is. At the very least, adjustments will have to be made¹³. But to argue that no understanding about the real world is at all gained by doing Newtonian mechanics with frictionless surfaces seems ludicrous. And this is what's missing from aerodynamics without air-resistance—not much, perhaps nothing, of what's of interest and value in doing aerodynamics survives this «idealization».

Without pretending that this is an adequate discussion of idealization in philosophy (and elsewhere)¹⁴, the thing to ask now is whether discussing statistical evidence using the idealizations involved in the hypothetical cases Allen dislikes so much is more like doing Newtonian mechanics with frictionless surfaces, or more like doing aerodynamics without air resistance. And it is here that I find his complaints so unconvincing. Think about the suspiciousness of relying on statistical evidence in the real world¹⁵; now assume all these «friction»—like factors away, so that you remain with an unrealistic clean case of statistical evidence. Does it feel like you've assumed away the problem? Of course not. These other factors are much more like friction to Newtonian mechanics, and much less like air-resistance to aerodynamics.

Of course, as elsewhere, one should not proceed too quickly from normative recommendations in the idealized friction-less world to the real world. Adjustments will have to be made. Perhaps the non-reflective practitioner cares more about these adjustments than about the theoretical insights. But if what we are after is also an understanding of «the nature of juridical proof»—its *nature*, that is, presumably, some central truths about it that are at least to an extent constant across a fairly wide range of circumstances and instrumental payoffs (and jurisdictions)—then *of course* models and hypothetical cases are of value. They may very well be indispensable.

4. MORE ON STATISTICAL EVIDENCE

All of this was, as I said at the outset, fully general: Sure, it applies to statistical evidence and to my views thereof, but it's not in particular about either. (For instance,

¹³ As I argue in ENOCH 2018, in some cases the relation between ideal theory and non-ideal theory resists this natural and simple adjustment model.

¹⁴ That such idealization is indispensable in science as well may be dialectically important in our context, because it shows that even if Allen wants to emulate science, perhaps because he's doing «naturalized epistemology», he can't escape idealization. Again, ENOCH 2018.

At least some of Allen's complaints against the use of the probability calculus are also, I think, based on a failure to appreciate the important role of idealized models.

¹⁵ When Allen insists that courts do rely on statistical evidence all the time, he conflates statistical evidence with other phenomena. See the next section.

everything I said so far equally applies to defending Martin Smith's normic-explanation theory from Allen's related criticisms). But I want to conclude with a few more specific points, directly responding to some of Allen's more specific criticism of what we say about statistical evidence. I restrict myself to three points: about characterizing both statistical evidence and the legal treatment thereof, about Sensitivity, and about incentives.

4.1. How *does* the law treat statistical evidence?

Allen seems puzzled over what the fuss is about. It's just false, he says, that the (American) law has any problem with statistical evidence. It takes such evidence into account all the time.

This criticism fails in both directions. First, we've never claimed that courts—American courts included—*never* admit statistical evidence. Indeed, one of our papers (ENOCH AND FISHER 2015) is precisely a discussion—in a law review—of some doctrinal complications and subtleties that it would not have been as natural to discuss in the original, more abstract and philosophical paper (ENOCH, SPECTRE, AND FISHER 2012).

Second, and more importantly, many of Allen's examples are utterly unconvincing. In order to multiply examples, Allen lumps together all cases in which courts are happy to take into account evidence that has something to do with probabilities. But the relevant sense of statistical evidence is much, much narrower than this. We explain this in some detail in responding to Pardo (in ENOCH AND SPECTRE 2019, 183; a paper Allen refers to in a different context) who makes, we think, the same mistake:

When we—following the literature—speak of statistical evidence, we think of examples such as Blue Bus, and the phenomenon it is an example of. This is the phenomenon sometimes called base-rate evidence, sometimes market-share evidence, or sometimes naked statistical evidence... But when Pardo speaks of statistical evidence, he seems to be thinking of any evidence at all that involves a statistic. We never claimed—who would?—that any evidence involving statistics is poor or should be inadmissible, or indeed is insensitive. So Pardo's example of *Tyson Foods*—where statistics were admitted in order to prove something about the circumstances, rather than as direct base-rate data for guilt or liability—is just irrelevant for our discussion. (Footnotes and page references omitted.)

As far as I can see, all of this applies directly also to Allen's examples.

4.2. Sensitivity

Central to our discussion of statistical evidence is the claim that in paradigmatic cases of statistical evidence such evidence is—even if not misleading—not sensitive to the truth. That is, roughly, even had the accused not crashed the gates, the court that relies on statistical evidence would have still found him guilty. Not so in the case

of non-statistical evidence¹⁶. This, we think, is an important difference between statistical and non-statistical evidence, though it falls short of vindicating a differential treatment of these different kinds of evidence (because the law, in a slogan, shouldn't care about Sensitivity or other purely epistemic notions¹⁷). Here, I want to set the record straight regarding some of what Allen says about Sensitivity.

At one point, Allen insists that the difference between statistical and non-statistical evidence has nothing to do with Sensitivity. Here is what he says: «The witness' testimony is not more "sensitive" [than statistical evidence]; rather, one simply does not know what causes her mistakes and thus one does not know how they will be distributed. One form of evidence is being preferred for no other reason than ignorance of its distributional effects». But first, to repeat, we do *not* think that Sensitivity is something the law should care about, so this talk of *preferring* one kind of evidence over another misses its mark¹⁸. Second, in order to evaluate Allen's claim about the nature of the difference between testimony and statistical evidence we would need an account of counterfactuals, of causation, and more. I humbly suggest not engaging in all of those discussions here—though I would advise caution with way-too-quick assumptions that take for granted grand metaphysical claims that are hotly debated.

Allen says: «When it comes to the actual trial setting, a complex body of evidence consisting of several pieces of relevant and interrelated statistical information that converges in a way that reveals a 70% likelihood that a blue bus caused the harm is at least as good, if not better, evidence as a body of evidence that includes only the testimony of a single witness who is known to be 70% reliable in identifying the color of the bus that caused the harm». This is problematic, though, and not just because here too Allen seems to take any probabilistic evidence at all to be statistical in the intended sense. The main problem here is that whether the complex body of evidence he describes is indeed as good as (or better than) the testimony may depend on further factors about this body of evidence—perhaps including, as far as anything Allen has said, on whether the pieces of evidence in this body are sensitive. Here too, then, there is no viable objection to anything we say or think about Sensitivity.

¹⁶ Notice that the relevant question from the point of view of the court cannot be whether the relevant judgments are sensitive, because sensitivity entails truth, and the court is at that point uncertain whether the relevant judgment is true. The important point is that judgments that are based on non-statistical evidence are sensitive-if-true, whereas judgments that are based on statistical evidence are typically insensitive even if true.

¹⁷ We elaborate on and argue for this claim—in more detail than we have done in the past—in “Does Legal Epistemology Rest on a Mistake” (ENOCH, FISHER AND SPECTRE, ms).

¹⁸ Allen writes: «If what they mean is that there really is no interesting difference between the two forms of evidence, it would have been helpful, and saved a lot of time, for them to have just said that». (footnote 50). Well, we think that there *are* interesting differences between the two kinds of evidence, in both epistemic and incentive-related terms. But we don't think that there's a necessary difference in their reliability (the context of this footnote of Allen's), and we think that Sensitivity doesn't and shouldn't matter to the law. I think we've been pretty clear on all of these things, so that if time was wasted, this is not on us.

4.3. Incentives

Allen rightly notes that on our picture what—if anything—justifies the law's differential treatment of statistical and non-statistical evidence is not directly the epistemic Sensitivity story, but rather an incentive story. We note that accepting statistical evidence will, in some cases, erode the incentive effect that the substantive law seems to be interested in securing. The Gatecrashers case can again serve as a good example. If you are standing there, deliberating whether to buy a ticket or crash the gates, then the substantive law that makes gatecrashing a crime is supposed to give you an incentive to buy a ticket. But if you know that you may be convicted on the basis of statistical evidence alone, you know that your own action has little effect on your legal fate, or anyway, less effect than it has in a legal regime that is more suspicious of statistical evidence: After all, if so many people crash the gates, then the statistical evidence will be available against you even if you happen to buy a ticket. And if so, whatever incentive we may have hoped the criminal law supplied to buy the ticket rather than crash the gates is seriously eroded by the willingness to rely on naked statistical evidence. Notice how modest our claim here is: we're not saying that accepting statistical evidence always, necessarily, eliminates the relevant incentives. What we are saying is that the evidence law regime—and in particular, the attitude towards statistical evidence—may *affect* the relevant incentives, and at least often, in a counter-productive way.

Allen's main complaint about this story—on top of the complaints about «weird hypotheticals» already discussed above—is that they are unrealistic in what they require of the relevant agents. Addressing a similar incentive story about the Blue Bus case, he claims that such a story presupposes that bus drivers take the evidence regime into account in their deliberation. But they clearly don't. So the incentive story is obviously false.

This, however, is just not how incentive-explanations work, partly because this is not how incentives work. Sure, *one* way in which agents—people, organizations, etc.—respond to incentives is by explicitly deliberating about the incentives' significance for them and their interests. But it is an extremely impoverished picture of action and its responsiveness to incentives (and other reasons) to think of this as the *only* way in which incentives may work. Incentives serve to structure choice situations, and they infiltrate actions in other ways as well. Sometimes we can point to the precise mechanisms that explain responsiveness to incentives—explicit deliberation about the incentives, imitating others who explicitly deliberate about the incentives, falsely thinking that one acts for one reason while in truth one is responding perhaps subconsciously to incentives, an organization's decision-making procedures being sensitive to incentives without any single person in them being similarly sensitive, and so on. Sometimes we can't even pinpoint the specific mechanism that renders an agent sensitive to the relevant incentives. And really, this doesn't much matter—what matters for the plausibility of incentive-explanations is not *how* the relevant agent is

sensitive to the relevant incentives, but *that* this is so. Nothing about explicit deliberation need be implied (nor, it should go without saying, is any universal generalization about everyone being perfectly sensitive to incentives all the time).

Notice also how, if you (and Allen) reject the point of the previous paragraph, your conclusion is of much wider scope (and partly for that reason utterly implausible). Does Allen really want to suggest—on the basis of a priori reasoning, no less!—that nothing in evidence law supplies any incentive at all for primary behavior? Indeed, if his objection to our incentive story is to be taken seriously, he must reject anything remotely resembling law-and-economics explanations of pretty much anything, and vast parts of the social sciences. He may want to do this, but he's going to have to do a more convincing job arguing for such surprising claims.

5. CONCLUSION

I agree with Allen's general tone, according to which we should be cautious in applying lessons from epistemology – or from other philosophical sub-disciplines – to legal theory. Legal doctrines of causation should probably not be sensitive to the finer points in the metaphysics of causation; if need be, we can disambiguate a legal and a metaphysical notion of causation, and take things from there. How the criminal law should treat attempts should probably not be sensitive to the finer points in the philosophy of action; if need be, we can disambiguate a legal and a philosophical notion of attempt (or of trying), and take things from there. And the same holds, of course, for epistemology and the law of evidence. This is a point I've now made in some detail, both in more general contexts¹⁹ and specifically in the context of a discussion of statistical evidence²⁰. But nothing resembling Allen's conclusions seems to me to follow: Nothing against a priori reasoning, nothing about the illegitimacy of considering “weird hypotheticals”, and at the end of the day, nothing by way of his more specific claims about statistical evidence either. Nor does it follow that philosophy has no lessons of value to a legal theorist: In fact, if Allen is sincere about his attempts to implement in legal theory the lessons of naturalized epistemology, he cannot consistently²¹ reject all attempts to incorporate philosophical insight into studying the law.

¹⁹ See the last section (How Not to Do Law and Philosophy) in Enoch (2012).

²⁰ See section IV (On the Relations between Law and Epistemology) in Enoch and Spectre 2019.

²¹ Though perhaps – it's really hard to tell – consistency is not something he takes so seriously: “In a closed deductive and static system, inconsistencies are a problem. In complex dynamic ones, they are the order of the day.”

6. BIBLIOGRAPHY

- ENOCH, D., 2012: "Comment on Yaffe's *Attempts*", *Jerusalem Review of Legal Studies* 6: 20 – 35.
 — 2018: "Against Utopianism: Noncompliance and Multiple Agents", *Philosophers' Imprint* 18.
- ENOCH, D., SPECTRE, L., and FISHER, T. , 2012: "Statistical Evidence, Sensitivity, and the Legal Value of Knowledge", *Philosophy and Public Affairs* 40, 197-224.
- ENOCH, D. and FISHER, T., 2015: "Sense and Sensitivity: Epistemic and Instrumental Approaches to Statistical Evidence", *Stanford Law Review* 67, 557 – 611.
- ENOCH, D. and SPECTRE, L., 2019: "Sensitivity, Safety, and the Law: A Reply to Pardo", *Legal Theory* 25, 178-199.
- ENOCH, D., FISHER, T. and SPECTRE, L. (ms): "Does Legal Epistemology Rest on a Mistake? On Fetishism, System Design, and Conscientious Fact-Finding".
- MACHERY, E., STICH, S., ROSE, D., CHATTERJEE, A., KARASAWA, K., STRUCHINER, N., SIRKER, S., USUI, N. & HASHIMOTO, T., 2015: "Gettier across Cultures", *Nous* 51, 645-664.
- NAGEL, J., 2012: "Intuitions and Experiments: A Defense of the Case Method in Epistemology", *Philosophy and Phenomenological Research* 85, 495-527.
- PINILLOS, N. A., JARAMILLO, S. & HORNE, Z., 2019: "Asymmetric belief sensitivity and justification explains the Wells Effect", *Proceedings of the 41st Annual Cognitive Science Society*.

NATURALIZED EPISTEMOLOGY AND THE LAW OF EVIDENCE: METHODOLOGICAL REFLECTIONS

Michael S. Pardo

Professor of Law,

Georgetown University Law Center

ABSTRACT: This paper discusses Ronald Allen's article, «Naturalized Epistemology and the Law of Evidence Revisited», and reflects on how epistemology can contribute to our understanding of the evidentiary proof process. I first situate Allen's critique of recent philosophical scholarship, distinguishing between general theoretical accounts of proof (including the theory that Allen and I have defended), on one hand, and the applications of specific epistemological concepts or issues to law, on the other. I then present a methodological picture that diverges in some respects from the one that emerges from Allen's critique. In discussing this alternative methodological picture, I explain how epistemology can contribute to legal evidence and proof while avoiding the problems that Allen identifies.

KEYWORDS: Epistemology; legal proof; relative plausibility; statistical evidence; safety; sensitivity.

SUMMARY: 1. INTRODUCTION.—2. NATURALIZED EPISTEMOLOGY AND THE LAW OF EVIDENCE: THEN AND NOW.—3. AN ALTERNATIVE METHODOLOGICAL PICTURE.—4. CONCLUSION.—5. BIBLIOGRAPHY.

RECOMMENDED CITATION: PARDO, MICHAEL S., 2020: «Naturalized Epistemology and the Law of Evidence: Methodological Reflections», in *Quaestio facti*, 2: 299-311. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22484

1. INTRODUCTION

Theoretical scholarship on legal evidence and proof has been on the rise throughout the world¹. Roughly, and for purposes that will be clear in a moment, we can distinguish two different strands in the fabric of theoretical evidence scholarship. The first strand involves the application of theoretical concepts and insights from other academic disciplines. For example, within this strand, one might draw on developments in epistemology in order to apply some philosophical concept or insight to an issue in the legal domain². A second strand of theoretical scholarship, also drawing on interdisciplinary resources, aims to provide a general theoretical account of the nature and structure of a legal system's evidentiary proof process as a whole and its component parts³. A prominent example is the debate between probabilistic and explanatory accounts of proof and the ensuing shift from the former to the latter⁴. Although we can draw a rough distinction between these two stands, they are related. On one hand, specific epistemological considerations may provide support, or create problems, for general theoretical accounts⁵. On the other hand, operating within a general theoretical framework (and its attendant assumptions) may affect the plausibility of a particular application of epistemology to law and whether it successfully illuminates a legal issue⁶.

My purpose in drawing attention to these related but distinct strands of scholarship is to situate Ron Allen's article, «Naturalized Epistemology and the Law of Evidence Revisited», and my perspective in discussing it. I now find myself in the somewhat awkward position of being caught in an apparent tension between these two strands of scholarship. On one hand, Allen and I are the proponents of a general explanatory account of legal proof, known as relative plausibility⁷. On the other hand, my scholarship has also drawn on concepts and insights in epistemology in analyzing aspects of the evidentiary proof process, employing in the process some of the methodological aspects (e.g., «weird hypotheticals») that Allen powerfully critiques in his article. So, what gives?

In the discussion to follow, I will explain why I do not see a tension between these two strands of scholarship. I will also explain why I am more sanguine than Allen in the potential for epistemology to illuminate evidentiary issues in law. In focusing on these specific goals, my hope is that the discussion will help to elucidate more general

¹ The existence of this journal, devoted to evidential legal reasoning, is further evidence of this rise.

² See, e.g., PARDO, 2018; PARDO, 2010; PARDO, 2007.

³ See, e.g., PARDO & ALLEN, 2008; ALLEN & PARDO, 2019a; ALLEN & PARDO, 2019b.

⁴ See the symposium on «Relative Plausibility and Its Critics», *International Journal of Evidence & Proof*, 2019 (23): 1-217.

⁵ See, e.g., GOLDMAN, 2002; LAUDAN, 2008; HAACK, 2014, each raising epistemological problems for subjective probabilistic accounts of legal evidence.

⁶ See PARDO, 2019, arguing that recent theoretical discussions of the «proof paradoxes» mistakenly assume a probabilistic account of standards of proof.

⁷ PARDO & ALLEN, 2008; ALLEN & PARDO, 2019a; ALLEN & PARDO, 2019b.

methodological issues and also to clarify possible ways in which epistemology can contribute to our understanding of evidence law, evidence theory, and legal proof. The next section provides some background and summarizes Allen's critique. Section 3 then provides and defends an alternative methodological picture, explaining how scholarship consistent with that picture can avoid the problems that Allen identifies.

2. NATURALIZED EPISTEMOLOGY AND THE LAW OF EVIDENCE: THEN AND NOW

Scholars exploring the relevance of «naturalized» epistemology for law have emphasized an instrumental connection between epistemic considerations and truth (or factual accuracy)⁸. The philosopher Alvin Goldman uses the term «veritistic» to refer to epistemic evaluation along this dimension⁹. According to Goldman, epistemic evaluation along this dimension examines the tendency of rules or practices to produce true beliefs or judgments (as opposed to errors or ignorance), and, accordingly, he argues that «it is natural to evaluate existing [legal] procedures along the veritistic dimension»¹⁰.

In their 2001 article, «Naturalized Epistemology and the Law of Evidence», Ron Allen and Brian Leiter adopt Goldman's veritistic approach and argue that naturalized epistemology «provides the most appropriate theoretical framework for the study of evidence»¹¹. The veritistic approach on which they rely is «normative» and «regulative»—«it is concerned with the production of knowledge, meaning in part *true* belief»¹². In assessing evidence law along this dimension, they emphasize two important constraints that epistemic evaluation «must honor»¹³. First, normative conclusions must respect the «ought implies can» constraint¹⁴. Second, epistemic

⁸ GOLDMAN, 1999; ALLEN & LEITER, 2001. On naturalized epistemology, see QUINE, 1969; KORNBLITH, 1994. On naturalism and legal theory, see LEITER, 2017. «Naturalism» may refer to a variety of methodological or substantive claims; Allen's «naturalized» approach is primarily a methodological one. See ALLEN, 2020: 1-2 [citations are to manuscript pages].

⁹ GOLDMAN, 1999: 69. Goldman is a pioneer in the branch of epistemology known as «social epistemology». See ALLEN & LEITER, 2001: 1497 («Social epistemology is simply that branch of naturalized epistemology concerned not with individual knowers but with the social processes and practices that inculcate belief»).

¹⁰ GOLDMAN, 1999: 272. See also *ibid.*: 278 («The whole point of this chapter is to consider which systems or procedures of fact-finding are epistemically superior»). See also HAACK, 2014; LAUDAN, 2008.

¹¹ ALLEN & LEITER, 2001: 1493. In particular, their naturalized approach is one of applied social epistemology. *Ibid.*: 1497 («[N]aturalized social epistemology must consider the range of empirical sciences that examine the social mechanisms of belief-inculcation. In what follows, we shall often speak of “naturalized epistemology” and “social epistemology” interchangeably.»)

¹² *Ibid.*: 1498 (original emphasis).

¹³ *Ibid.*: 1499.

¹⁴ *Ibid.*

evaluations ought to be determined by instrumental considerations; they should be assessed according to the «the actual consequences» of competing epistemic norms, evaluating «which are the most effective means of producing knowledge»¹⁵. With these methodological considerations in hand, Allen and Leiter examine theoretical approaches to legal proof as well as specific types of evidence and evidentiary rules¹⁶. In addition to their specific conclusions, the article is also useful methodologically, providing insightful examples of how to apply epistemology to the law of evidence.

In revisiting the relationship between naturalized epistemology and the law of evidence, Allen now applies some of the methodological considerations from that previous article to recent philosophical discussions of legal proof¹⁷. By «naturalized epistemology», Allen clarifies that he is primarily referring to «inquiries—analytical or empirical—into how the human mind engages with its environment and the implications of that form of engagement for Western legal systems»¹⁸. Examples of such inquiries include: «how such legal systems structure dispute resolution, the relationship between that structure and the field of evidence with its resultant rules, and how all of this is influenced by cognitive capacities»¹⁹. Allen's approach to investigating these issues follows the methodological constraints described above, including: the ought-implies-can principle; a focus on «the actual state» of the legal system and its actors (informed by the best available empirical evidence); and a concern for the instrumental effects that any normative recommendations are likely to have on the legal system²⁰.

In light of these methodological considerations, Allen first points to the success of «relative plausibility» as a general theoretical account²¹. This theory explains the evidentiary proof process, and its components, in terms of explanatory criteria and considerations²². After describing the success of relative plausibility, Allen critiques recent efforts by scholars relying on epistemology to provide «alternative», «competing explanations» of legal evidence and proof²³. Unlike relative plausibility, these accounts «are at odds with an empirically oriented naturalized approach» and suffer from various «shortcomings»²⁴. These «shortcoming» include: (1) the problem of the «weird

¹⁵ *Ibid.*

¹⁶ They discuss probabilistic and economic theories of evidence as well as specific rules concerning demeanor and character evidence. *Ibid.*: 1503-49.

¹⁷ ALLEN, 2020.

¹⁸ *Ibid.*: 2.

¹⁹ *Ibid.*

²⁰ Allen contrasts his methodology with «a priori, conceptual», «a priori, top-down» and «conceptual/normative» analyses. *Ibid.*: 1, 4. As noted, however, his methodological approach includes within its scope both «empirical» and «analytical» inquiries. *Ibid.*

²¹ The success of relative plausibility is contrasted with the failures of probabilistic theories. *Ibid.*: 4-10.

²² PARDO & ALLEN, 2008; ALLEN & PARDO, 2019a; ALLEN & PARDO, 2019b.

²³ *Ibid.* 2-3. He discusses the Blue Bus and Gatecrasher hypotheticals and focuses the majority of his critique on recent philosophical discussions by ENOCH ET AL., 2012, and SMITH, 2018.

²⁴ *Ibid.*: 1.

hypothetical»—«[h]ypotheticals are posited that are supposed to reflect some aspect of the legal system or the field of evidence, but do not»²⁵; (2) «some theorizing neglects critical aspects of the actual state of affairs»²⁶; and (3) «much of the theorizing [...] makes essentially impossible epistemological demands»²⁷. Because of these problems, the alternative explanations are «implausible» and have little utility for law²⁸.

I, of course, share Allen's views about relative plausibility. I also agree with many aspects of his general methodological discussion as well as many of his specific criticisms. Nevertheless, I remain more sanguine about the genre of scholarship that Allen critiques—in my view, epistemology may provide a greater source of understanding for legal evidence and proof than Allen's article may suggest²⁹. In the next section, I will sketch an alternative picture.

3. AN ALTERNATIVE METHODOLOGICAL PICTURE

The methodological picture I will describe shares a number of details with the one that emerges from Allen's critique. In particular, there are five shared aspects that I take as a starting point:

First, and perhaps least controversially, theorizing about legal evidence and proof should accurately describe the object of inquiry. This is easier said than done when attempting to describe a complex entity such as the evidentiary proof process. But trying to avoid either misdescribing the object of inquiry or neglecting critical aspects of it are details of Allen's naturalized approach with which I assume no one would take issue.

Second, the ought-implies-can constraint is also one that, at least in principle, should be uncontroversial and, thus, I agree that theorizing should avoid making impossible demands, epistemological or otherwise.

²⁵ *Ibid.*: 3.

²⁶ *Ibid.*

²⁷ *Ibid.*

²⁸ *Ibid.*: 1, 3, 6. It is important to note that although Allen is critical of some attempts to apply epistemology to law (i.e., what he calls the «cottage industry» focusing on statistical evidence, *ibid.*: 3), his argument is *not* a wholesale rejection of the relevance and significance of epistemology for evidence law. Indeed, both ALLEN & LEITER, 2001, and Allen's current paper (as well as some of our work on relative plausibility) involve applications of epistemological insights to law. There is no contradiction here. Allen's methodological approach, in my view, is thus best seen as charting a middle path between some types of epistemological applications, on one hand, and more extreme skeptical positions, on the other. For more extreme expressions of skepticism about the relevance of epistemology for evidence law, see PARK, 2001: 2067; ENOCH ET AL., 2012: 211 («[T]he law should not care about knowledge, or indeed about epistemology in general»).

²⁹ For reasons discussed below, some of the possible differences between our methodological pictures are likely to be a matter of focus and framing, as opposed to genuine substantive disagreement. See *ibid.*: 14-15 n. 48.

Third, the instrumental constraint, as outlined by Allen and Leiter, is an important one for social institutions such as law³⁰.

Fourth, I share Allen's anti-reductionist impulse. No theory is likely to provide a «key» or an algorithm for understanding all aspects of legal proof based on a small subset of variables. This is true for probabilistic approaches to legal proof, and it is also true for theories that draw on additional epistemological concepts (e.g., knowledge, justification, warrant, safety, sensitivity, normic support, and so on).

Fifth, as mentioned, I share with Allen the view that our explanatory theory, relative plausibility, provides the best available general account of the evidentiary proof process. Thus, I envision work that draws on additional aspects of epistemology as being complementary to, rather than in competition with, relative plausibility³¹.

Thus, the alternative picture that I envision is empirically informed, anti-reductive, and consistent with the best general theoretical account of legal proof (relative plausibility). According to this picture, epistemic concepts and issues help to inform, clarify, and illuminate features of the evidentiary proof process. One theoretical upside of this methodological approach is that it may help to make *explicit* features or issues that are *implicit* in the evidentiary proof process. Elucidating these features and issues may, on the one hand, contribute to a better understanding legal proof, and, on the other hand, unearth features and issues so that they may then be subjected to more informed normative scrutiny. In discussing this methodological picture, I will defend some possible uses of «weird hypotheticals»³², and I will illustrate how this methodological approach can avoid the problem of «impossible epistemological demands»³³.

Allen's primary targets in his critique of «weird hypotheticals» are the well-known Blue Bus and Gatecrasher examples and their variants³⁴. These examples are typically invoked in both the evidence-law and philosophical literatures in the service of illustrating something problematic with statistical evidence. As Allen discusses, a «cottage industry» has arisen using the examples to distinguish statistical evidence from non-statistical evidence, attempting to explain exactly *why* the former is prob-

³⁰ ALLEN & LEITER, 2001: 1499. See also GOLDMAN, 1999; LAUDAN, 2008.

³¹ Relative plausibility draws on insights from the philosophical literature on «inference to the best explanation». See LIPTON, 2004; HARMAN, 1986. Because of the differing context and the complexities of legal proof, however, there are ways in which the inferential process differs from IBE as it is articulated in the philosophy of science and in epistemology. We discuss some of these differences in ALLEN & PARDO, 2007. See also PARDO & ALLEN, 2008: 243-47 (discussing why some of common objections to IBE in philosophy do not apply in the legal context).

³² ALLEN, 2020: 3.

³³ *Ibid.*

³⁴ In Blue Bus, the primary evidence linking the defendant company to a particular accident is evidence that the defendant owns a majority of the buses in the jurisdiction. The hypothetical is loosely based on dicta in *Smith v. Rapid Transit*, 58 N.E.2d 754, 755 (Mass. 1945). In Gatecrasher, the primary evidence against a particular defendant, alleged to have entered a ticketed event without paying, is evidence that a majority of attendees entered without paying. See COHEN, 1977: 75.

lematic in a way that the latter is not³⁵. He complains that the examples contribute «little of interest regarding the nature of juridical proof», do not attend to the «actual conditions of legal systems,» and «obscure rather than enlighten»³⁶.

I agree that these examples are more misleading than helpful in understanding legal proof and that «the attention focused on these (and similar) hypotheticals may have simply impeded progress»³⁷. My diagnosis of the reasons why, however, differs in some respects from Allen's. In my view, there are two central problems with these examples³⁸. The first problem, as Allen discusses, is that the *question* they are being used to answer is not a genuine problem for law in need of analysis or explanation³⁹. In other words, I share with Allen the view that there is not a *general* problem in law involving «statistical evidence»⁴⁰. Therefore, analyses or criteria that seek to explain a general, and meaningful, distinction between statistical and other evidence in law are doomed to fail (whether they rely on weird hypotheticals or not). Because the hypotheticals are being put in the service of answering an ill-formed question, it is not surprising that after decades of trying no one has successfully answered the question or «solved» the problem⁴¹.

The second problem is that, even if the question being addressed were genuine, the examples do not provide enough information to answer it. The hypotheticals are framed so as to imply that the evidence appears to be sufficient on its face to satisfy the applicable standard of proof⁴². Implicit in this framing, however, is the assumption that the standard of proof is a probabilistic threshold (say, 0.5 for the preponderance standard), which the statistic in the example surpasses. But this assumption is itself highly questionable⁴³. If, instead, standards of proof are better understood as explanatory thresholds, as Allen and I argue, then the evidence is no longer necessarily sufficient on its face to meet the standard of proof⁴⁴. Without more information,

³⁵ ALLEN, 2020: 3.

³⁶ *Ibid.*: 8, 10-11.

³⁷ *Ibid.*: 5.

³⁸ These examples and these issues are discussed in more detail in PARDO, 2019.

³⁹ See ALLEN 2020: 27-29 (discusses the uses of statistical evidence throughout the law).

⁴⁰ PARDO, 2019: 265 («It is therefore a mistake to infer that the examples capture something essential about legal doctrine»). To be clear, the point is that there does not appear to be a general issue for statistical evidence *as distinct from* non-statistical or «individualized» evidence. Sometimes the former will be probative and sometimes not, and the same holds true for the latter. The quality, epistemic features, and probative value of legal evidence cut across the distinction between statistical and other evidence.

⁴¹ For further discussion, see PARDO, 2019: 253-66.

⁴² This apparent sufficiency then creates a puzzle for those who conclude that the evidence is insufficient. *Ibid.*

⁴³ *Ibid.* Also problematic is the assumption that the statistic expresses the probative value of the evidence. *Ibid.*: 257-61.

⁴⁴ ALLEN & PARDO, 2019a: 6-15.

we simply do not know what to infer about the evidence and its sufficiency⁴⁵. Thus, the problems with these examples (and with analyses involving the examples) arise not because the hypotheticals are «weird» (although they are, for the reasons Allen discusses)—the problems arise because the hypotheticals are underdetermined and do not provide a clear answer to the question for which they are being invoked.

By contrast, I will discuss examples of «weird» hypotheticals that I think help to illustrate issues involving legal evidence and proof and that are consistent with the general methodological picture outlined above.

The first example is one that Allen critiques for making «impossible or outlandish epistemological demands»⁴⁶. The example, presented as a critique of the importance of epistemic sensitivity for law, was designed to illustrate that even insensitive evidence may nevertheless be reliable and highly probative:

Drug Weight: A criminal defendant is tried for possessing a large amount of an illegal substance. According to the criminal statute at issue, the prosecution must prove that the defendant possessed 500 grams or more of the substance. A chemist who sampled and tested the substance will testify as an expert that the amount seized was over a kilogram. The expert is well qualified, employed an acceptable and reliable methodology, and otherwise satisfies the criteria for admitting expert testimony. The chemist's process, however, tends to overestimate weight by a very small amount (say, by one gram or less)⁴⁷.

What exactly does this weird hypothetical help to illustrate? The evidence is highly probative in establishing the disputed fact of consequence—namely, that the amount seized was over 500 grams. Nevertheless, when we examine this evidence in terms of its apparent sensitivity (i.e., looking at the closest possibilities in which the disputed fact *is false*⁴⁸), the evidence is insensitive because the expert's testimony may still report that the evidence meets the threshold even when it is slightly below the threshold. The point of the example, however, is that even if this were true, it does not and should not matter in cases such as the example because, given the large amount involved, the actual case is far away from that possibility. The key point is that the closest possibility in which a belief, verdict, etc. is false *may not be a close*

⁴⁵ PARDO, 2019: 282 (arguing that the examples «remain underdetermined because of the nature of the standards [of proof] and the limited amount of information in the examples»).

⁴⁶ ALLEN, 2020: 14 n. 48.

⁴⁷ PARDO, 2018: 66.

⁴⁸ Determining which alternative possibilities to consider can be a tricky issue in philosophical theorizing. One important difference between philosophy and law is that, consistent with relative plausibility, it is typically left to the parties to specify the alternative possible explanations that should be considered. Another important difference is that safety and sensitivity are «factive»—they assume that the propositions at issue are true and compare those circumstances to circumstances in which the propositions are false. In law, the disputed propositions are typically not known to be true or false at the time of decision. Therefore, the relevant comparisons are: *assuming the proposition is true*, would the same decision follow if it were false (sensitivity) or how easily could the same decision follow if the fact were false (safety).

possibility and thus *may not have much of an effect of the probative value of evidence*⁴⁹. This conclusion follows from a general epistemological detail about sensitivity that, I argue, carries over to the context of legal evidence and proof. Because of this detail, whether evidence is sensitive or not does not tell us much about its probative value. By contrast, I argue, epistemic safety (or something like it) better tracks these considerations⁵⁰. In the example, the testimony appears to be safe⁵¹—even if there is a small range in which the expert's process is unreliable, the result here is so far outside of that range that the expert could not easily be mistaken about whether the amount exceeds 500 grams. Thus, epistemic safety may help to illustrate this feature of legal evidence⁵².

What, then, is Allen's complaint about the hypothetical? In arguing that it makes impossible epistemological demands, he writes that:

If one «knows» that the process «overestimate[s] by a» gram, then one «knows» what the true amount is. If the chemist's mistake is within a range, then again you know the range of the true weight. This hypothetical makes sense only if you already know what you need to know, and if you do, then one simply presents it to the factfinder. If one does not «know» the error, then the argument about safety does not run through⁵³.

But the hypothetical can work to illustrate its intended point without making impossible (or outlandish or peculiar) epistemological demands. We do not need to *know* the exact amount by which the process errs in a particular case—I agree that if we knew this, then we would know the true amount and could present this to the factfinder. Suppose that there is a small range in which the process is unreliable; sometimes the result is slightly over the true amount and sometimes it is slightly under, but (given our best available knowledge) the former tends to happen more often than the latter. Thus, if called upon to make a precise judgment in a particular case in which the true weigh was near the statutory threshold, then there is a real danger of a false positive (and false conviction)⁵⁴. The idea of safety helps to illustrate this feature.

Moreover, Allen's own critique seems to draw on something like this epistemic consideration. He writes:

⁴⁹ See WILLIAMSON, 2002: 159-60; PARDO, 2018: 66.

⁵⁰ *Ibid.*: 69-74.

⁵¹ See also note 48.

⁵² This is not to suggest that safety captures all aspects of probative value; rather, it draws attention to an epistemic feature of the latter than may otherwise go unnoticed. For further discussion, see PARDO, 2018.

⁵³ ALLEN, 2020: 15 n. 48.

⁵⁴ As well as a false negative. Allen may object that, in the counterfactual, the factfinder should simply be presented with the fact that the particular testimony is within the range (and thus it would not be relied upon). Possibly—but such an argument seems to rely on the sort of a priori assumption that Allen rejects. By contrast, a naturalized approach that pays close attention to the actual conditions of the legal system suggests that reliance on and deference to prosecution experts, even when their testimony exceeds the limits of known reliability, is a problem in criminal cases. NATIONAL RESEARCH COUNCIL, 2009.

The fact that the Blue Bus Company owns 70% of the buses does not entail that there is a 70% chance that the Blue Bus Company caused the harm in question. The Blue Bus Company may own 70% of the buses but have extremely effective safety measures in place; they may have well-seasoned drivers with impeccable driving records; they may not even have buses that frequently visit the location of the harm at the time in question. Actual factfinders could easily regard this statistic as being insufficient to ground liability because these relevant possibilities remain unaccounted for by the statistic about ownership⁵⁵.

Indeed, this is exactly right. These «unaccounted for» «relevant possibilities» could easily be the case (i.e., the situation is unsafe) and this, in turn, diminishes the probative value of the evidence⁵⁶. When the alternative possibilities could *not* easily be the case, then the evidence will have higher probative value in proving the disputed fact. Nothing in my previous analysis, or in this discussion, is meant to suggest that safety is the «key» to probative value or the only relevant concept for understanding legal evidence. Rather, it provides one more analytical tool for understanding and making explicit features of legal evidence, or the process of legal proof, that are implicit and may otherwise go unnoticed⁵⁷. And it is possible to use this tool without making impossible epistemological demands.

My second, and final, example concerns Gettier-type cases and their potential relevance to legal proof. Within epistemology, this particular sub-genre is probably the best example of «weird hypotheticals» that seem far removed from the practical realities of legal proof⁵⁸. As I have argued, however, even Gettier cases contain potential insights for law and help to unearth unnoticed features of legal evidence and proof⁵⁹. Within epistemology, the examples are typically used to illustrate that even true and epistemically justified beliefs may fall short of qualifying as «knowledge» because of an accidental or coincidental relationship between the belief's truth and the justifying evidence. A similar problematic gap between evidence and truth may arise in the context of legal verdicts. Here is a «weird hypotheticals» that illustrates such a gap:

Framed Defendant: The police arrest a motorist and plant drugs in his car. He is convicted at trial of illegal possession based solely on testimony from the arresting officers and the planted drugs. As it turns out, the defendant did have illegal drugs in his car at the time that never were discovered. The verdict that the defendant possessed drugs is therefore both true and justified (that is, the evidence at the time of the trial is sufficient to establish a conviction beyond a reasonable doubt), but the truth and the justifying evidence are disconnected. The truth of the verdict is purely coincidental or accidental⁶⁰.

⁵⁵ ALLEN, 2020: 17-18.

⁵⁶ See note 48.

⁵⁷ Allen likely agrees with this methodological point. See *ibid.*: 20 (distinguishing the «many tools to use in thinking about evidence» from «efforts to find the solution or key» to legal proof). This is why I noted above, see note 29, that the differences between our methodological pictures are likely more a difference in focus and framing rather than substantive disagreement.

⁵⁸ GETTIER, 1963. See also LYCAN, 2006 («The Gettier problem became a leading focus, if not the focus, of disenchantment with the definition-and-counterexample method of analytic philosophy»).

⁵⁹ PARDO, 2010.

⁶⁰ *Ibid.*: 57.

My view is that such a gap is a problem that renders the verdict an error in need of correction⁶¹. Moreover, the hypothetical helps to illustrate a genuine epistemological issue that may arise for courts. One context in which it may arise is when criminal defendants seek a new trial based on newly discovered evidence. The prosecution may respond to such claims by pointing to other evidence of guilt that was not presented at the trial and that purports to show the defendant's guilt⁶². This then raises a question for courts reviewing the issue: must the evidence that purports to establish guilt have been «internal» to the trial and presented to the factfinder? In *Reese*, for example, the court answered this question in the affirmative—concluding that the prosecution must present any additional incriminating evidence in a new trial⁶³.

To be clear, nothing in the «weird hypotheticals» necessarily tells us how the law ought to respond in such situations⁶⁴. At the very least, however, the hypotheticals help to reveal epistemic issues that may arise in the context of legal proof. Once the issues are made explicit⁶⁵, there is then space for a more informed debate about how the law does, or ought, to proceed (as opposed to leaving the issue implicit and possibly to the whims of individual decision-makers).

4. CONCLUSION

The methodological picture I have described is one way in which epistemology may contribute to the law of evidence. It is intended to be an optimistic one and an invitation to epistemologists. For anyone taking up this invitation, I think there are three important points that Allen's critique reveals. First, it is important (and sometimes difficult) to accurately describe the underlying legal details. Second, there is a complex relationship between epistemology and law. It is unlikely that useful connections will involve a simple application of some epistemic concept or issue

⁶¹ See *ibid.* This problem arises even when the gap does not arise because of government misconduct. See *ibid.*: at 51-52 (presenting additional examples).

⁶² See, e.g., *Pennsylvania v. Reese*, 663 A.2d 206 (1995).

⁶³ *Ibid.*: 210 («Because the jury did not hear evidence of other explanations [...], it would have been improper for the [reviewing] court to have considered it»).

⁶⁴ In addition, nothing in the methodological picture that I have presented relies on privileging «intuitions» about the hypotheticals. See ALLEN, 2020: 5 (critiquing reliance on intuitions about the cases). The role of intuitions in epistemology is a controversial topic. See GOLDMAN, 2012; CAPPELEN, 2012. In my view, this issue need not be resolved for epistemology to contribute to legal proof.

⁶⁵ Because, under my methodological picture, epistemology can contribute by unearthing epistemic issues implicit in law, it is not a criticism of this picture that the law does not explicitly refer to the epistemic concepts. See ALLEN, 2020: 31 n. 94. Moreover, given the complex relationship between epistemology and law, the latter may be useful to law even when a concept or issue differs between the two contexts. For example, aspects of «knowledge» may be useful for understanding law, even when legal verdicts fail to qualify as «knowledge», according to philosophical analyses. See PARDO, 2010: 54-55. See also PARDO, 2018 (arguing that safety illuminates aspects of legal evidence regardless of whether it is a necessary condition for knowledge); MOSS, 2021.

to law—a particular concept or issue may have different features, or play different roles, in the legal context than it does in other epistemological contexts. Third, it is probably time to stop trying to «solve» issues related to the Blue Bus and Gatecrasher hypotheticals and, for those interested in contributing to our understanding of the law of evidence, to focus elsewhere⁶⁶.

5. BIBLIOGRAPHY

- ALLEN, R.J. & B. LEITER, 2001: «Naturalized Epistemology and the Law of Evidence», in *Virginia Law Review*, 87(8): 1491-1550.
- ALLEN, R.J. & M.S. PARDO, 2007: «Probability, Explanation, and Inference: A Reply», in *The International Journal of Evidence & Proof*, 11: 307.
- 2019a: «Relative Plausibility and Its Critics», in *The International Journal of Evidence & Proof*, 23(1-2): 5-58
- 2019b: «Clarifying Relative Plausibility: A Rejoinder», in *The International Journal of Evidence & Proof*, 23(1-2): 2-5-17.
- ALLEN, R.J., 2020: «Naturalized Epistemology and the Law of Evidence Revisited», in this journal.
- CAPPELEN, H., 2012: *Philosophy without Intuitions*, Oxford: Oxford University Press
- COHEN, L.J., 1977: *The Probable and the Provable*, Oxford: Oxford University Press.
- ENOCH, D., SPECTRE, L. AND FISHER, T., 2012: «Statistical Evidence, Sensitivity, and the Legal Value of Knowledge», in *Philosophy & Public Affairs*, 40(3): 197-224.
- GETTIER, E., 1963: «Is Justified True Belief Knowledge?», in *Analysis* 23: 121.
- GOLDMAN, A.I., 1999: *Knowledge in a Social World*, Oxford: Oxford University Press.
- 2002: «Quasi-Objective Bayesianism and Legal Evidence», in *Jurimetrics*, 42: 237.
- 2012: «Philosophical Naturalism and Intuitional Methodology», in GOLDMAN, A.I. (ed.), *Reliabilism and Contemporary Epistemology*. Oxford: Oxford University Press.
- HAACK, S., 2014: *Evidence Matters: Science, Proof, and Truth in the Law*, Cambridge: Cambridge University Press.
- HARMAN, G., 1986: *Change in View: Principles of Reasoning*, Cambridge, MA: MIT Press.
- KORNBLITH, H., 1994: «What is Naturalistic Epistemology?», in KORNBLITH, H. (ed.), *Naturalizing Epistemology*. Cambridge, MA: MIT Press.
- LAUDAN, L., 2008: *Truth, Error, and Criminal Law: An Essay in Legal Epistemology*, Cambridge: Cambridge University Press.
- LEITER, B., 2017: «Naturalism in Legal Philosophy», in ZALTA, E. N. (ed.), *Stanford Encyclopedia of Philosophy*, available at: <https://plato.stanford.edu/entries/lawphil-naturalism/>.
- LIPTON, P., 2004: *Inference to the Best Explanation*. London: Routledge.
- LYCAN W.G., 2006: «On the Gettier Problem Problem», in *Epistemology Futures*, Oxford: Oxford University Press
- MOSS, S., 2021: «Knowledge and Legal Proof» in *Oxford Studies in Epistemology* Vol. 7, Oxford: Oxford University Press.
- NATIONAL RESEARCH COUNCIL, 2009: *Strengthening Forensic Science in the United States: A Path Forward*, Washington D.C: The National Academies Press.
- QUINE, W.V., 1969: «Epistemology Naturalized», in *Ontological Relativity and Other Essays*, New York: Columbia University Press.

⁶⁶ Possible epistemic issues include those involving testimony, epistemic injustice, experts, deference, disagreement, group agency and collective judgments, and specific issues involving the inferences and reasoning in particular types of legal cases, among many others.

- PARDO, M.S., 2007: «Testimony», in *Tulane Law Review*, 82: 119.
- 2010: «The Gettier Problem and Legal Proof», in *Legal Theory*, 16: 37.
- 2018: «Safety vs. Sensitivity: Possible Worlds and the Law of Evidence», in *Legal Theory*, 24(1): 50-75.
- 2019: «The Paradoxes of Legal Proof: A Critical Guide», in *Boston University Law Review*, 99: 233.
- PARDO, M.S. and ALLEN, R.J., 2008: «Juridical Proof and the Best Explanation», in *Law & Philosophy*, 27(3): 223-68.
- PARK, R.C., 2001: «Grand Perspectives on Evidence Law», in *Virginia Law Review*, 87: 2067.
- SMITH, M., 2018: «When Does Evidence Suffice for Conviction?», in *Mind*, 127(508): 1193-1218.
- WILLIAMSON, T., 2002: *Knowledge and its Limits*, Oxford: Oxford University Press.

STRONG, WEAK, OR APPARENT NATURALIZATION? RELATIVE PLAUSIBILITY THEORY AND CONCEPTUAL ANALYSIS

Nicola Muffato

Università degli Studi di Trieste

Associate Professor, Department of Legal Science,

Languages, Interpretation, and Translation

ABSTRACT: The paper questions the naturalized epistemological framework for Allen's relative plausibility theory (RPT). It is divided into two main sections: §1 is devoted to show that the weak/modest naturalism of Allen's central arguments make them hardly distinguishable from a traditional variety of conceptual analysis, which is neither a priori nor dependent on intuitions; §2 deals with the "empirical" character of RPT, presenting some objections to its methodological assumptions and the ambitious project of providing a unifying account of the profoundly different facets of the process of proof in USA legal systems without relying on a sufficiently wide array of sociological and empirical data and without specifying the bridge principles governing the relations between the various domains of inquiry.

KEYWORDS: naturalized epistemology; conceptual analysis; relative plausibility theory; jurisprudence; methodology.

SUMMARY: 1. NATURALIZATION AND CONCEPTUAL ANALYSIS.— 2. THE EMPIRICAL CHARACTER OF THE RELATIVE PLAUSIBILITY THEORY.— 3. REFERENCES

RECOMMENDED CITATION: MUFFATO, NICOLA, 2020: «Strong, Weak, Or Apparent Naturalization? Relative Plausibility Theory And Conceptual Analysis», in *Quaestio facti*, 2: 313-324. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22465

In «Naturalized Epistemology and the Law of Evidence Revisited»¹, Professor Ronald J. Allen presents a summary of several interesting and provocative arguments against the main competitors of the conception of legal evidence and proof he defends, relative plausibility theory (henceforth: RPT). Since I am quite sympathetic with almost all the objections he offers and with the general purpose of his criticism—to wit: one would better study how the rules on evidence of real legal systems are applied and followed, before trying to abstract from their conceptual articulation in order to rationalize them—I will not try to refute his refutations. I shall instead question the naturalized epistemological framework for RPT, taking cues from the quick brush strokes Allen gives in the target paper and the references to his related works, where this model is explored more in detail. In my view, not only does not the author carry until the end the naturalization project, but RPT does not even need such a support to overcome its a priori rivals. The paper is divided into two sections: §1 is devoted to show the weak naturalism of RPT's epistemological framework; §2 deals with the empirical evidence for RPT.

1. NATURALIZATION AND CONCEPTUAL ANALYSIS

1.1. In a path-breaking article written in 2001 with Brian Leiter, «Naturalized Epistemology and the Law of Evidence», Allen maintains that naturalized social epistemology provides a fruitful way of understanding evidence law. The authors endorse Alvin Goldman's attribution of a *normative* task to this kind of philosophical thought: the regulation of the social mechanisms and practices of belief-inculcation so that they could causally lead in a reliable way to knowledge². From this point of view, epistemology is seen as a *normative methodology* of belief-acquisition, as a method of theory appraisal, and, eventually, as a proposal of partial reconceptualization of a practice: given certain conditions and contexts, it prescribes compliance with certain models and methods of belief-acquisition (e.g., of scientific inquiry) considered better than others because more reliable in yielding to true beliefs. In the case of evidence law, the task consists in *evaluating* a specific system of legal rules on evidence—thus seen *also* as a mechanism or practice of belief-inculcation—in the light of the accepted methodological prescriptions³ and a new conceptual apparatus⁴.

A *naturalized* (normative) epistemology, however, is also characterized by its continuity with and dependency upon empirical science(s)⁵. Ascending to the level

¹ ALLEN, 2020.

² Cf. ALLEN & LEITER, 2001: 1498: «Social epistemology, in the sense we adopt, is *normative* or *regulative* in its ambitions. We want to ask, as Goldman puts it in his recent important book on the subject, “Which [social] practices have a comparatively favorable impact on knowledge as contrasted with error and ignorance?”».

³ Cf. ALLEN & LEITER, 2001: 1501-1503.

⁴ Cf. ALLEN, 1986.

⁵ Cf. ALLEN & LEITER, 2001: 1497.

of meta-epistemology, we can say that the methods for epistemological theorizing ought to be empirical/naturalistic, and scientific. As shown by Leiter⁶, this thesis is ambiguous. According to a first reading («Methods Continuity» thesis, henceforth MCT), it means that epistemological rules should reflect or imitate some relevant aspects/contents of scientific methods and choices. If so reconstructed, epistemology turns out to be an abstraction or a generalization from scientific methods⁷: its norms and principles are gathered through a bottom-up process that moves from the rules scientists implicitly or explicitly follow when engaged in the thoughtful practice of theorizing⁸.

According to a second reading, epistemology has a distinctive method—e.g., conceptual explanation or conceptual analysis—which does not abstract from scientific methods, but must be constrained by the *results* of the theories developed following those methods («Results Continuity» thesis, henceforth RCT): philosophical theses and conceptual claims must be supported or justified by the results of empirical theories. However, Leiter does not clarify what it means for empirical evidence (expressed by empirical propositions) to «support» or «justify» conceptual claims. Since he seems to rely on Goldman's view, I guess he assumes that epistemological theorizing does not start within science⁹, but must in any case have a factual basis of *reliable* cognitive processes on which epistemic concepts and properties supervene. For example, epistemic justification—a normative concept—supervenes on a social, reliable, causal mechanism of belief-inculcation, which can be studied empirically.

It must be noted, however, that in both readings epistemological rules only make sense if they can be accomplished with or applied by human beings—the rationalist «“should” implies “can” principle»—and are seen as instrumental as to the achievement of epistemic goals. A naturalized approach to the evaluation of a system—*restitutio*: a dynamic order—of legal rules of evidence consists in confronting the latter with a feasible methodology of belief-acquisition. Formal models (e.g., expected utility theory, probability theory, economic analysis) and other reductionist conceptions (e.g., sensitivity theory) of legal decision-making and fact-seeking fail this test because they commit the unforgivable sin of de-contextualization: their hypotheticals are too artificial and when applied in concrete contexts they generate pragmatic oddities or conceptual impossibilities. In brief, such models conflict with the facts about how we happen to reason and the limitations posed by real legal systems to the ascertainment of truth.

Allen and Leiter say they are doing normative epistemology. But then they add: «Whether epistemology in general should be naturalized is, in any case, irrelevant

⁶ Cf. LEITER, 1997: 289-290; 2017: §1.

⁷ A problem I will not examine consists in the selection of the appropriate methods a naturalized jurisprudence should adopt.

⁸ Cf. NIXON, 2004.

⁹ This would explain Leiter's apparently incoherent claim that «the “concept” of law is not illuminated or fixed by empirical inquiry in the natural and social sciences» (LEITER, 1997: 302).

to the question of whether naturalized epistemology provides a fruitful way of understanding evidence law. The latter is our contention»¹⁰. In the quoted passage, the authors seem to concede that even if the naturalization project in epistemology were wrong or unconvincing or useless from a meta-epistemological point of view, evidence law could in any case be usefully (correctly?) described (interpreted?) through the lens of naturalized epistemology. I'm not sure I've understood their point: given MCT or RCT, if naturalized epistemology cannot be defended, neither can a naturalized theory of evidence law. Maybe what they want to say is that the best way to explain evidence law consists in reconstructing it as at least partially informed by the empirical methods recommended by a naturalized normative epistemology. This explanation would not depend on the acceptance of the latter, even if Allen and Leiter do in fact accept it. In other words, the authors would be maintaining that evidence law is a mechanism of belief-inculcation, «a practice that has as one of its elements the production of knowledge»¹¹ and whose rules are, at the same time, legal, epistemic, and «naturalistic». This would be a strong contention, if its naturalism were demanding. Unfortunately, as we will see, this is not the case.

1.2. In a passage of the target paper, Allen retains the feasibility and instrumental requirements, but offers a significantly different definition of «naturalized epistemology»: «As I am using the phrase, it refers to inquiries—analytical or empirical—into how the human mind engages with its environment»¹². The first thing to note here is that Allen has shifted from normative methodology to descriptive theorizing. I hesitate to qualify it as a piece of «descriptive methodology», because the phrase «inquiries into how the human mind engages with its environment» lends itself to two alternative interpretations: on one side, it could amount to a description or an explanation of the methods we follow in order to know and deal with our environment; on the other, it could be a description or an explanation of (i) some relevant features of our biology, or (ii) the capacities and abilities we display when immersed in our social and natural environment, or (iii) the functioning of a social epistemic practice. According to this second view—which, I suspect, is more in tune with Quine's replacement naturalism than with Goldman's version of RCT—epistemology does not describe «methods». Be as it may, such inquiries do not produce by themselves the regulation of social practices.

Second, the engraved disjunction—«analytical or empirical»—is a bit confounding, given that naturalization and rejection of the analytic/synthetic distinction are generally seen as coextensive. What do the two disjuncts mean, then? Let us start from «analytic». On a first interpretive hypothesis, Allen uses here the adjective as roughly synonymous with «conceptually regimented» but divorced from the allegedly a priori necessities accessible through the ordinary intuitions to which armchair

¹⁰ Cf. ALLEN & LEITER, 2001: 1496-1497.

¹¹ Cf. ALLEN & LEITER, 2001: 1500.

¹² ALLEN, 2020: 2.

philosophy appeals. In fact, he ties his own conceptualizations «to either realistic hypotheticals or the actual state of affairs»¹³, trying to make them sensible to empirics, that is: to evidence gained through experience. In this case, the disjunction probably reveals the adoption of RCT.

On a second interpretation, quite typical in the literature concerned with the methodology of empirical research in social sciences, while «empirical», narrowly intended, is referred to methods and techniques of data collection, «analytic» is predicated of methods and techniques of description, modeling, and explanation of the data. It follows from this construction that legal analysis is part and parcel of empirical social sciences—such as, e.g., sociological jurisprudence or forensic psychology—and that Allen's epistemology is grounded on MCT rather than RCT.

Which interpretation is to be preferred? The explicit reference to Goldman's works suggests the first. However, as I anticipated, it is not clear how much Allen's method of epistemological theorizing—especially his RPT—is naturalistic. On one hand, some of the arguments he presents against other theories of evidence applied to law—e.g., the conjunction problem, the reference class problem, the objection against a probabilistic appraisal of evidence based on the computational intractability of updating prior probabilities, the mental experiment designed to show how base rates and assignments of probabilities must be related in order for a standard of proof to minimize expected loss¹⁴—or supporting his own view—e.g., the argument for the reciprocity of the burdens of persuasion¹⁵, the distinction between narratives and explanations¹⁶, the metaphor of legal systems as evolutionary optimization problems and neural networks¹⁷—are not only genuinely conceptual, but also unlikely to be brought back to supervenience relations on (or reduction relations to) a unique set of precisely identifiable natural properties. On the other hand, without the support of precise laws of supervenience (or of concept- and theory-reduction), sheer reference to empirical or scientific evidence in the context of the explanation of a social practice, even if useful, is not sufficient to qualify a theory as naturalistic or empirical, in the sense of naturalized epistemology (and jurisprudence). This comment, however, needs a clarification.

If we strip away from RCT the methodological requirement of equipping a naturalized epistemology with an explanation of the principles governing the relation of evidential support or justification between a set of empirical propositions and a *conceptual* claim, RCT itself reduces to a generic prohibition to disinterest in scientific or empirical facts about human behavior/psychology/biology, and nature in general. The latter, however, is a simple expression of good common sense. Its rejection would

¹³ ALLEN, 2020: 3.

¹⁴ Cf. ALLEN, 1997: 346.

¹⁵ Cf. ALLEN, 2019: 12-13.

¹⁶ Cf. ALLEN & PARDO, 2019a: 7, fn 7, 17, fn 86; 2019b: 211, fn 20.

¹⁷ Cf. ALLEN, 2013: 112-113; 2019: 20-23.

amount to a very peculiar philosophical move, a move that surely cannot be sensibly made when trying to elucidate the concepts that articulate our understanding of a social practice (e.g., the legal practice of civil trials in a U.S.A. jurisdiction)¹⁸. Otherwise, how could they be analyzed as articulating a *certain* practice in the first place? If concepts depend on social linguistic and inferential practices, then the facts about these practices—about what we do when we talk and engage in reasoning—cannot be irrelevant to non-formalist conceptual analysis. What I'm saying is that weakening RCT results in an epistemological technique which can be hardly distinguished from a quite common and traditional variety of conceptual analysis: conceptual clarification¹⁹.

Still, according to this conception of philosophy, empirical evidence does not affect concept-use and conceptual relations in a simple, straightforward way. More specifically, conceptual claims are neither justified nor supported by empirical evidence alone. A defender of conceptual clarification might contend that we learn our concepts through conditioning, training, and enculturation, not by sheer induction or abduction from our past and present experiences, and that, analogously, even if we often modify our concepts under the impact of empirical discoveries, this does not imply that the new concepts are induced or abducted from the latter: rather, the practice—e.g., the way we identify things—effects that a certain empirical evidence E is treated as a new *criterion* for the application of the concept C, thereby establishing a new conceptual relation between E and C.

In other words, some empirical propositions are *treated as rules* of concept-use and thus become relevant for the *meaning* of certain interactions, not as evidence for testable generalizations or hypotheses about individual or social human biology. This is the reason why Ludwig Wittgenstein portrayed conceptual analysis as a therapeutic clarification based on perspicuous *descriptions* of language- and concept-use rather than as a theoretical *explanation* of causal relations between sensorial inputs and linguistic outputs. A consequence of this approach is the rejection of the simplistic conclusion that «philosophical claims are always vulnerable to the success of a *posteriori* inquiry»²⁰: empirical discoveries cannot refute philosophical claims, because the latter are not propositions, but representation-rules whose function is to show conceptual relations. A philosophical claim can only be «refuted» showing that it is incompatible with other already «accepted» philosophical claims.

Secondly, the fact that naturalized normative epistemology forbids any ultimate appeal to intuitions is not a differentiating trait of this conception, since concep-

¹⁸ This means that the theories criticized by ALLEN are not only *empirically*, but also *conceptually* inadequate.

¹⁹ However, it is also strikingly similar to other kinds of conceptual analysis: compare, for example, Allen's RPT and Jules Coleman's pragmatic conceptual analysis of tort law (as exposed in COLEMAN, 2001).

²⁰ Cf. ALLEN & LEITER, 2001: 1495.

tual clarification can easily comply with it. As Constantine Sandis aptly noticed, those who believe the contrary (e.g., experimental philosophers, Leiter, but also defenders of conceptual analysis, as Kenneth E. Himma²¹) mistakenly identify «claims about how “one” (or “we”) would ordinarily apply a word [or use a concept, for the case] with beliefs or intuitions about correct usage»²². Conceptual clarification need and does not rely on intuitions: it focuses on the *rules* implicit in the linguistic and non-linguistic behaviors of the participants to a social practice in order to describe its conceptual articulation, using as analytic tools logic(s), philosophical distinctions, appropriately contextualized mental experiments. Besides, those who defend this way of philosophizing have their arguments against the equation of *competence* (and correct use) with majoritarian use or natural dispositions²³.

Third, «the two constraints of naturalized [normative] epistemology: “ought implies can” and the instrumental character of normative advice»²⁴ are simply not characteristic of naturalized epistemology alone. Many defenders of traditional (normative) epistemology would profess adherence to such principles—even if in some limited cases their application may be controversial.

In sum, Allen’s meta-epistemology is surely incompatible with *some kind* of a priori, «immodest», «formal» conceptual analysis, but this is not sufficient to qualify as «naturalized» his epistemology, since it is not clearly distinguishable from a traditional way of practicing conceptual clarification. In fact, he often defends his theses resorting to conceptual arguments, and there are reasons to doubt that RPT mainly rests on empirical grounds, as I’ll try to show in the following section.

2. THE EMPIRICAL CHARACTER OF THE RELATIVE PLAUSIBILITY THEORY

2.1. Let us now turn to the meaning of «empirical». In a paper coauthored with Michael Pardo, Allen claims that the task of their research project «is empirical: what is the best explanation of the data, where “the data” are observations of how the American legal system structures proof at trial?»²⁵. However, in what sense is it *empirical* an *explanation* of data, consisting in *observations* of the ways a specific legal system structures proof at trial? Italics signal the need for a clarification of these notions, because they are controversial and equivocal.

²¹ Cf. LEITER, 2003: 43-44; HIMMA, 2015: 74-75, 77. See also ALLEN & LEITER, 2003: 892.

²² Cf. SANDIS, 2010: 183.

²³ Cf. MEDINA, 2002: ch. 4 and 5; 2006: 35.

²⁴ Cf. ALLEN & LEITER, 2001: 1503.

²⁵ ALLEN & PARDO, 2019a: 7.

According to the authors, RPT seeks to offer the best empirical account of juridical proof and the entire litigation process in American legal systems, seen as a complex, organic, dynamic, and adaptive system²⁶. As they precise,

[RPT] is about the entire process of proof, including (1) the form, securing, and presentation of evidence, (2) the forms of argumentation employed at trial, (3) the manner in which humans process and deliberate on evidence, (4) the trial structure created by the rules of evidence and procedure, (5) the structure of litigation before and after trial, (6) the manner in which judges and juries, on the one hand, and trial and appellate judges, on the other hand, interact, and (7) to some extent about the meaning and nature of rationality²⁷.

It seems then that Allen and Pardo aim at explaining empirically *each one* of the seven objectives enumerated and their mutual connections. This ambition, and the efforts devoted to realize it, are praiseworthy. However, the declared objectives are very heterogeneous and ask for different methods of inquiry, some of which can be doubted as to their empirical (or scientific) character. I'm referring, especially, to the explanation of (4): its object is a normative legal order, a social practice in which standards, rules, and principles play a central role. In this case, the doubts do not stem from the idea that norms are not empirical objects—an idea Allen is committed to reject, given his naturalism. The problems are rather methodological.

Let us assume that empirical social sciences can describe—among other things—the beliefs, value-judgments, rules/methods, and choices of the members of a community engaged in a social practice. Since all these aspects are expressed through their discourses, the discourses of the social sciences are (at least in a loose sense) meta-linguistic. Now, a first methodological problem for the social scientist (or the historian) is that the recognition and description, in the meta-language of the social science, of a complex linguistic normative structure (e.g., an ideology, a philosophical current, a legal order), and even its reduction to specific representative sentences, are only possible through the difficult selection and application of a set of conceptual criteria—that is, norms—of identification. A second methodological problem depends on the fact that all social practices distinguish a correct way of participation and engagement from more or less serious errors: an adequate description of the practice must then take into account this normative component (even if the theorist ends up re-describing it as a collective illusion, ideology, or cognitive dissonance). A third methodological problem concerns the empirical adequacy of the theories which ascribe meanings to social behaviors.

The choice of the criteria of identification is guided by the purpose of giving the best explanation of the practice, reproducing the beliefs, values, methods, and choices of the subjects whose practices we wish to describe. However, this preliminary methodological operation is extremely delicate: since the application of the selected criteria implies a simplification, an idealization, and the imposition of a descriptive

²⁶ Cf. *ibidem*; ALLEN, 2020: 3.

²⁷ ALLEN & PARDO, 2019b: 207-208.

model to the practice, the risk of projecting the theorist's views onto the practice to be described lurks constantly in the background. It is thus essential to distinguish the normative criteria functional to the description of social values, beliefs, methods, and choices of the practitioners from the values, beliefs, methods, and choices the theorist considers rational and justifiable. In the case of legal analysis, it is important to specify if the attitudes and methods expressed in the discourses of the legal operators are identified by looking at the actual discourses the majority of each subgroup of practitioners produces, or at the discourses they say they are producing, or wanting to produce, or at the «best examples» of such discourses according to the observer.

Social theorists and historians are accustomed to such problems and try to avoid them making explicit their criteria of identification, varying such criteria in order to confront the resulting explanations, including in the description the attitudes and (linguistic or non-linguistic) actions they consider irrational/unjustified, but endowed with a relevant social diffusion, and accounting for the distance between the explicit (self-)definitions or (self-)representations of the practitioners and their real behavior. The officials of a legal order, on the other hand, do not care about all these descriptive facets (and do not praise explanations that preserve the «pluralism» of the practice) when they try to set what the law is or «says» on a specific matter. Neither do formal legal doctrine and descriptive jurisprudence/legal theory, even if they work on a different level of systematic and generality. This happens because, while operating with a (not too precise) common sense concept of law/legality, they are concerned with deriving from the law itself—not from the *concept* of law—the (more precise) technical criteria for its description—*rectius*: for the recognition of its sources and the determination of its normative contents—in order to find out the legal solution to a practical problem. Officials and formal legal doctrine assume law's criteria of self-delimitation as given: an assumption an empirical social scientist cannot accept, but must substitute with a testable theory about the *effectiveness* of the preferred rule of recognition.

What about RPT? As shown by the array of examples in the footnotes of the target paper, Allen is a profound connoisseur of statutory and common law of evidence in the U.S.A. Nonetheless, his theory does not follow the methods of empirical social sciences for the identification of their explanatory units. Nor does it proceed—at least to my knowledge—from a systematic and detailed recollection of sufficiently wide and statistically representative samples of procedural documents, proffered evidence, reported judicial decisions, verdicts of the juries, interviews to professional and lay legal operators, etc.²⁸ In this sense, Allen is neither developing nor relying on a sociological theory of evidence law to support his claims about American legal systems. On the other hand, if the «observations» of the ways a specific legal order structures proof at trial consist in the «best»—more comprehensive, coherent, sim-

²⁸ Allen shows an acute perception of many concrete aspects of the litigation practice, but does not provide an explicit systematic empirical support to the explanations of (1), (2), (5), (6) either.

ple, efficient, etc...—literal, systematic and teleological interpretations of statutes, precedents, jury instructions *et similia*, the scrutiny of legal doctrines and the analysis of the concepts they incorporate, then it seems that RPT's explanation of (4) assumes the legal practice's self-identification criteria: it is elaborated from the internal point of view, not from the external point of view of empirical social sciences²⁹. The practice, however, could well be more or less incoherent, or insensible to some features of legal rules, formal legal doctrines and best jurisprudential reconstructions (such as RPT): a possibility Allen does not explore empirically.

2.2. Where does then lie the empirical base of RPT? Almost entirely in the explanation of (3), which recruits the «story model» resulting from the researches in social and cognitive psychology conducted by Nancy Pennington and Reid Hastie. Still, even leaving aside concerns about the external and ecological validity of simulated trials, such a base does not suffice, in the light of the ambitious objectives of RPT, as an empirical theory. Here is a brief summary of some objections.

Daniel Simon argued that the story model «is of limited use in cases where there is not much of a story to tell, such as when the evidence is fragmented [...]. The same would apply to cases that hinge on technical judgments of patents, trademarks and antitrust, discerning causes of physical events, setting damage awards and so forth»³⁰; again, «A narrative [...] is of little relevance to negligence cases where the contested issue is a failure in appreciating a risk, to identification cases, or to cases in which the material facts concern the physical conduct of the defendant, the quality of a product, the extent of the damages, and the like»³¹. Allen might reply that explanations in a trial context are not stories, and that holistic reasoning supplies the lacking part in RPT: but in that case he should go into details and make explicit testable hypotheses about the holism he endorses. Besides, RPT neglects the experimental evidence and scholarship on biases and external factors interfering with jury's decisions that exclude or impair plausibility assessments³².

However, let us grant, *arguendo*, that Allen has sufficient empirical support for the claim that fact-finders process—or even cannot but process—evidence at trial in a holistic way, comparing alternative explanations and assessing their relative plausibility. Does this claim in turn support the conclusion that American legal systems—statutory rules, precedents, jury instructions, doctrines—are best explained by RPT? No, since the two are quite independent: much empirical (and conceptual) work is needed to link the explanation of (3) to an empirically well corroborated explanation of (4). And even if such a link were to be provided, there are good reasons to suppose that RPT would still need a conceptual framework whose elucidation cannot

²⁹ Moreover, one may object that some aspects of the formal legal doctrine seem inconsistent with RPT (cf. SCHAUER, 2019: 182).

³⁰ SIMON, 2019: 84.

³¹ SIMON, 2004: 564

³² Cf. SPOTTSMOOD, 2014.

be reduced to theoretical posits or empirical evidence for (3), (4), etc., as I tried to highlight in §1.2.

Finally, let us consider a different critical issue. In a rejoinder to some critics³³, who lamented a lack of precision in the definition of the object of RPT and its analytic tools, and concern about empirical literature, Allen and Pardo defended their view arguing that a narrower focus on empirical legal scholarship would have «limited utility in furthering the understanding of an entity such as juridical proof» at the level at which they want to explain it. In this regard, they invoke the cognitive complexity of the reasoning of the different legal actors in particular contexts and the factual complexity of the huge amount of cases the legal systems deal with to express the «doubt that there will be a well-organized science (or philosophical thought) of how people decide one explanation is better than another, even though that is precisely what people do in a virtually infinite number of ways»³⁴.

I think this reply misses the point of the criticism, which does not ask RPT to find a generalization or covering law in order to account or predict the decisions of individual cases. The objection, instead, is that an *empirical* explanation of the nature of legal proof should pay more attention to the results of empirical social sciences, as RCT prescribes. Consider human behavioral biology: its central object, human behavior, is extraordinarily complex, and can be approached scientifically from many different specific points of view (socio-biology, ethology, genetics, endocrinology, neurology, etc.). It is possible to elaborate a «big picture» of human behavior only transcending the boundaries of each discipline and reconstructing the complex connections between the phenomena studied by each one of them at different levels. Just like Allen tries to do with legal proof. But for RPT to become an *empirical* big picture, each explanation composing it should be empirically informed, as are the theories that illuminate different aspects of human behavior.

3. REFERENCES

- ALLEN, R.J., 1986: «A Reconceptualization of Civil Trials», in *Boston University Law Review*, 66(3): 401-437.
- , 1997: «Reasoning and its foundation: Some responses», in *International Journal of Evidence and Proof*, 1 (special issue): 343-347.
- , 2013: «Taming complexity: Rationality, the law of evidence and the nature of the legal system», in *Law, Probability and Risk*, 12: 99-113.
- , 2019: «Standards of Proof and the Limits of Legal Analysis», in *Diritto & Questioni Pubbliche*, 19(2): 7-25.
- ALLEN, R.J., LEITER, B., 2001: «Naturalized Epistemology and the Law of Evidence», in *Virginia Law Review*, 87(8): 1491-1550.

³³ Cf. SCHAUER, 2019; SIMON, 2019; VÁZQUEZ, 2019.

³⁴ ALLEN & PARDO, 2019b: 210.

- , 2003: «Naturalized Epistemology and the Law of Evidence: Reply to Redmayne», in *The Michigan State Law Review*, 2003(4): 885-892.
- ALLEN, R.J., PARDO, M.S., 2019a: «Relative plausibility and its critics», in *The International Journal of Evidence and Proof*, 23(1-2): 5-59.
- , 2019b: «Clarifying relative plausibility: A rejoinder», in *The International Journal of Evidence and Proof*, 23(1-2): 205-217.
- COLEMAN, J.L., 2001: *The Practice of Principle*, Oxford: Oxford University Press.
- HIMMA, K.E., 2015: «Conceptual Jurisprudence. An Introduction to Conceptual Analysis and Methodology in Legal Theory», in *Revus*, 26: 65-92.
- LEITER, B., 1997: «Rethinking Legal Realism: Toward a Naturalized Jurisprudence», in *Texas Law Review*, 76(2): 267-315.
- , 2003: «Beyond the Hart/Dworkin Debate: The Methodology Problem in Jurisprudence», in *The American Journal of Jurisprudence*, 48(1): 17-51.
- MEDINA, J., 2002: *The Unity of Wittgenstein's Philosophy: Necessity, Intelligibility, and Normativity*, Albany: State University of New York Press.
- , 2006: *Speaking from Elsewhere*, Albany: State University of New York Press.
- NIXON, J., 2004: «What is theory?», in *Educar*, 34: 27-37.
- SANDIS, C., 2010: «The Experimental Turn and Ordinary Language», in *Essays in Philosophy*, 11(2): 181-196.
- SCHAUER, F., 2019: «One small step towards a metatheory of evidence and proof», in *The International Journal of Evidence and Proof*, 23(1-2): 176-183.
- SIMON, D., 2004: «A Third View of the Black Box: Cognitive Coherence in Legal Decision Making», in *The University of Chicago Law Review*, 71(2): 511-587.
- , 2019: «Thin empirics», in *The International Journal of Evidence and Proof*, 23(1-2): 82-89.
- SPOTTSWOOD, M., 2014: «Bridging the gap between Bayesian and story comparison models of juridical inference», in *Law, Probability, and Risk*, 13(1): 47-64.
- VÁZQUEZ, C., 2019: «Less probabilism and more about explanationism», in *The International Journal of Evidence and Proof*, 23(1-2): 68-74.

IN DEFENSE OF WEIRD HYPOTHETICALS

Barbara A. Spellman

Professor of Law; Professor of Psychology
University of Virginia

«Our statements about the external world face the tribunal of sense
experience not individually but only as a corporate body».
Willard van Orman Quine¹

ABSTRACT: Professor Allen (this issue) critiques the value of using “weird” hypotheticals to mine intuitions about legal systems. I respond by supporting the value of “thin” hypotheticals for providing information about how people reason generally, rather than for revealing peoples’ specific answers. I note that because legal systems are the products of many minds thinking about how other minds operate, the object of inquiry is metacognition—that is, understanding how reasoning works.

KEYWORDS: legal epistemology; reasoning; metacognition; psychology and law

SUMMARY: 1. INTRODUCTION— 2. A NATURALISTIC TURN: THE RISE AND FALL AND RISE OF HEURISTICS AND BIASES— 3. WHAT DO «YOU» CARE WHAT OTHER PEOPLE THINK? THE VALUE OF THIN HYPOTHETICALS. 3.1. We are all guilty of overusing «simple» tools. 3.2. Through Thin and Thick. 3.2.1. From Abstract to Concrete to Just

¹ QUINE, 1951.

Right. 3.2.2. «Real» Legal Hypotheticals— 4. THE BLUE BUS BEEPS— 5. HYPOTHESIZING ABOUT HOW OTHER PEOPLE WILL THINK— 6. BIBLIOGRAPHY

RECOMMENDED CITATION: SPELLMAN B. A., 2020: «In Defense of Weird Hypotheticals», in *Quaestio facti*, 2: 325-337. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22477

1. INTRODUCTION

It is difficult to be a refutator of a refutator: in his essay², Ron Allen exhibits some deft refutating. It is also difficult for me to want to agree with Ron Allen about anything—even when I think he’s right³. And I do think he’s right that relative plausibility is a better explanation (than probabilities) of how people reason about evidence, and, in fact, about many things. More specifically, as a cognitive scientist, I agree with his conclusions: that the attempt to explain evidentiary reasoning with probabilities is misguided and that what I will call «thin hypotheticals» (rather than «weird hypotheticals»)⁴ are misused in trying to gather information useful for understanding our legal system. However, in my view, the fact that they can be misused does not mean that they cannot be extremely useful.

In Section 2, I describe why I believe that naturalized epistemology is a good way to think about human reasoning generally. There I lay the groundwork for the main argument in Section 3—that thin hypotheticals are, in fact, quite useful to gain understanding about how people reason about law and evidence. Part of my argument is that I often don’t care much about «what» people say in response to those hypotheticals but I care about «why» they say it. Empiricists who believe that people are giving odd or incorrect answers can try to chase down whether it is because of experimenter demands, the way the information is presented, the particular contents of the hypothetical, incompetence with statistics, or some general characteristic of human reasoning. We should care *why* because when we understand the cause, we can get ideas about what other situations would result in similar «bad» answers and how to fix them.

In Section 4, I’ll take us for a ride on the blue bus and consider relative plausibility. Finally, Section 5 steps back to consider how the legal system is a complex adaptive system and describes how thinking about how other people do think, and

² ALLEN, 2020 (this issue).

³ I once cut off an interminable audience member «question» at the end of a long conference day. Professor Allen took a vote among the attendees regarding my, admittedly, controversial behavior.

⁴ Allen uses the term «weird hypotheticals» and argues that they ignore important aspects of the legal system including its values. I believe that the modifier «thin» as introduced by Gilbert Ryle and made popular by the anthropologist Clifford Geertz captures those deficiencies better than «weird». A «thin» description is a surface level one, often of just a behavior; a «thick» description puts a behavior within a context or culture.

can think, (i.e., metacognition) is a more essential foundation of the system than is relative plausibility.

2. A NATURALISTIC TURN: THE RISE AND FALL AND RISE OF HEURISTICS AND BIASES

Let's start with a very brief and somewhat idiosyncratic view of the recent history of the psychology of judgment and decision making⁵. In the 1950's and 1960's⁶, there was some psychology research that endorsed the view of the «Rational Man», with studies that were interpreted as showing that humans are good at statistics and other types of formal reasoning.

But things turned in the late 1960's and 1970's; not all was found to be well with human's ability to think logically or statistically. An early important paper by Tversky and Kahneman (1974) called «Judgment under Uncertainty: Heuristics and Biases» demonstrated how experiment participants made various kinds of reasoning errors when presented with simple reasoning tasks that relied on basic concepts from logic or statistics. The authors concluded: «What is perhaps surprising is the failure of people to infer from lifelong experienc [...] fundamental statistical rules...». ⁷ For the next 20 years, during the «Reign of Heuristics and Biases»⁸, discovering different ways in which human reasoning was flawed became a cottage industry⁹ for experimental psychologists who argued, in effect, that people are stupid. But that was not Tversky and Kahneman's point: they acknowledged that people were making mistakes, but the interesting thing was that the mistakes were systematic rather than random. Thus, these errors could be useful to researchers when trying to uncover what people¹⁰ are really doing when they are, supposedly, or should be, using logic or statistics¹¹.

The 1990's were filled with explanations for the errors. Some lines of work argued that study participants were getting the «wrong» answer because they were filling in

⁵ See SPELLMAN and SCHNALL, 2009: 126-130, for a less brief but equally idiosyncratic explication.

⁶ PETERSON and BEACH, 1967.

⁷ TVERSKY and KAHNEMAN, 1974: 1130.

⁸ SPELLMAN and SCHNALL, 2009: 127

⁹ This cottage industry was probably located next door to the cottage industry that Allen mentions (ms p. 3), where unnamed disciplines create weird hypotheticals.

¹⁰ It is likely that all, and certain that most, of the experiments that I describe in this article had participants who were «WEIRD» people. WEIRD is an acronym for Western, Educated, Industrial, Rich, and Democratic—qualities of the people and their countries who comprised the overwhelming majority of psychology research participants up to 2010 (and continue to do so). HENRICH, HEINE & NORENZAYAN, 2010.

¹¹ I am happy to state that I believe that people are «bad» at the conscious use of probabilities and statistics. But I also believe that people often make judgments «as if» they are using probabilities and statistics without even knowing that they do so.

missing information in the sketchy experimental problems or vignettes with other (typically sensible) information¹². Another line of work argued that evolution could explain the errors, that humans did not evolve using probability theory, and that people were much better at reasoning when numbers are described in terms of frequencies rather than probabilities¹³. And a line of work that grew out of Kahneman's lab suggested that many of the errors arose from «attribute substitution»—that when a problem is missing information, or too difficult, or has to be done under time pressure, reasoners will substitute an easy-to-think about way of answering rather than grappling with the difficult demands¹⁴.

One example, relevant to the blue bus problem, comes from the base-rate fallacy. Experiment participants read brief personality descriptions of people who were said to be either engineers or lawyers. The descriptions were said to be drawn from a group of 100 descriptions, of which 70 were engineers and 30 were lawyers, or vice versa. Participants were remarkably insensitive to the proportions, making about the same profession-probability judgments for each personality description regardless of the base-rate condition.

Another famous example is the «Linda problem», for which many people fall prey to the conjunction fallacy. Participants read about an intelligent 31-year old woman who, as a college student, was concerned with progressive issues and had participated in demonstrations. When asked whether it was more likely that Linda was a bank teller or that Linda was a bank teller who was active in the feminist movement, participants chose the latter, despite that answer being logically impossible¹⁵. In each of these examples, rather than engaging in complicated thinking about likelihoods, participants were engaged in the simpler task of thinking about what the subjects of the vignettes were like.

Dual process theory developed out of the recognition that there were similar patterns of evidence in a wide range of research domains¹⁶. It suggests that people use two different paths in reasoning: the «heuristic system» (known as System 1), in which reason is fast and effortless and the «logical system» (known as System 2), in which reason is slow and conscious¹⁷. Note that although this model provides a convenient way to think about how people reason, and has been useful in getting people to make better decisions,¹⁸ it should be thought of as a metaphor rather than an accurate description of a mechanism.

¹² E.g., CHENG and NOVICK, 1990, for causal attribution

¹³ GIGERENZER and HOFFRAGE, 1995.

¹⁴ KAHNEMAN and FREDERICK, 2002.

¹⁵ KAHNEMAN and TVERSKY, 1982.

¹⁶ SMITH and DECOSTER, 2000.

¹⁷ There are other qualities assigned to the two systems and debates over which are essential qualities. KAHNEMAN, 2011; EVANS & STANOVICH, 2013.

¹⁸ See KAHNEMAN, 2011; THALER & SUNSTEIN, 2009. Of course, it can also be used to find ways to get people to make worse decisions.

3. WHAT DO «YOU» CARE WHAT OTHER PEOPLE THINK?¹⁹ THE VALUE OF THIN HYPOTHETICALS

I don't care much about what people think of as answers to weird hypotheticals, but I do care why they think it. To figure that out takes a lot of work—more than using the same thin hypothetical over and over to try to tap into «intuitions». To actually learn something, we need at least some comparison conditions, preferably thoughtful ones, and hopefully, other research that reveals the same underlying explanation (or theory) of the result.

The short problems or vignettes from the heuristics and biases era can be viewed as thin hypotheticals and we have learned many things from them. For example, for some problems, most people jump to incorrect intuitive conclusions²⁰ but others do not; whereas for other problems, the proportions are reversed. What affects how such judgments are made? We know that judgments are affected by characteristics of: the problem—e.g., whether it is phrased using probabilities or frequencies²¹; the person—e.g., people who are high in Need for Cognition are more likely to use the more logical reasoning system²²; and the environment in which the reasoning takes place—e.g., how much time is available, whether the person is multitasking.

About the blue bus and gatecrasher hypotheticals, Allen states that «for these hypotheticals to be helpful (to jurists), they must be crafted without losing sight of the actual object of inquiry, namely, the legal system»²³. Well, yes, and no. When I first read Allen's statement, I thought: «But where is the legal system?»—riffing on the philosopher Gilbert Ryle's story of a foreign visitor to Oxford University, who, having been shown colleges, libraries, sports fields, etc., eventually asks: «But where is the university?»²⁴. In fact, Allen does provide a very inclusive description of «the American legal system» and juridical proof quoted from Allen and Pardo (2019)²⁵. The description includes a lot about how trials are structured and how litigation proceeds using evidence, including «(3) the manner in which humans process and deliberate on evidence».

To my reading, however, the most important part of Allen's description needs us to take a step back to something assumed but not said that encompasses it all: the structure and processes and forms of the systems of proof were designed by human minds to regulate the actions of other humans as they attempt to influence other minds. So, what, really, is the object of inquiry? The legal system or the human

¹⁹ Yes, the title of the 2001 book about Richard Feynman.

²⁰ Note that it is not always the case that a fast intuitive answer is wrong whereas the slow conscious one is correct; however, most research problems are set up that way.

²¹ GIGERENZER and HOFFRAGE, 1995.

²² KAHNEMAN and FREDERICK, 2002.

²³ ALLEN, 2020 (this issue): ms p. 272.

²⁴ RYLE, 1949: 16.

²⁵ ALLEN, 2020 (this issue): ms pp. 276-277, quoting extensively from ALLEN & PARDO, 2019: 207-208.

mind (as it interacts with and within that system)? Allen's claim that weird hypotheticals ignore important aspects of actual legal systems—because they do not capture all the information in trials, and they tend not to represent all the values that go into a trial decision—is absolutely true. But it is mostly irrelevant.

3.1. We are all guilty of overusing «simple» tools.

I agree with Allen that we (lawyers, philosophers, psychologists, and others I'm sure) are all guilty of using the same «simple» hypotheticals over and over. Certainly, when a large part of a community knows about the blue bus or about vehicles that manage to sneak into a park,²⁶ it creates a shorthand way of speaking about ideas to indoctrinate others. But is it useful beyond that?

Thin hypotheticals remind me of thin experiments—or what would be called experiments with low external validity (i.e., unlikely to generalize outside the laboratory). In experimental psychology, when a simple experimental manipulation creates some interesting and easy-to-obtain data, the field goes wild using it, and often extrapolates far beyond its likely relevance.

For example, in the DRM paradigm, experiment participants listen to a list of words (usually nouns) that are all related to one word, which is not presented. In one version, they hear 12 words including: «bed», «rest», «tired», «dream», but they never hear the word «sleep». Immediately after hearing the final word, they are asked to recall the entire list. Typically, about 40% of the participants will recall the never-presented word «sleep». The first paper using this technique²⁷ has been cited over 4500 times since 1995. The fact that a simple experiment can create a «false memory» is a lovely existence proof that very simple short-term false memories, under conditions of high demand, can be easily created. But how relevant should researchers believe this result is for false memories of witnesses in the courtroom?

Recently the so-called experimental philosophers have been stuck on replay, using the Trolley Problem²⁸ to study morality so often that it has become a joke and generated dozens of memes. The Trolley Problem seems to me to be an exemplary «thin hypothetical» (and also a «weird hypothetical»)—especially when people are asked «what would you do?» as opposed to «what should a person do»? Researchers can add more people to the tracks, give more time, put the participant in the role of the engineer—but what have we learned about real morality? Currently, there is some push-back to contextualize morality research and get away from such thin hypotheticals.²⁹

²⁶ SCHAUER, 2008.

²⁷ ROEDIGER and McDERMOTT, 1995.

²⁸ FOOT, 1978; THOMSON, 1985. The memes can be found by using a Google Image search for «trolley problem».

²⁹ SCHEIN, 2020.

Using the same thin hypotheticals over and over just to check people's intuitions is certainly not useful. As Allen remarks³⁰, he often doesn't share the same intuitions as others. Nor should he. Intuitions don't come from nowhere. The heuristics and biases (H&B) literature took advantage of the fact that with thin hypotheticals many or most people jumped to the same wrong intuitive conclusions—but because of personality, knowledge, or training, not everyone jumps. There is a fascinating «adversarial collaboration» paper by Kahneman, whose H&B work often showed intuitions getting people to the wrong answer, and Klein, whose work on naturalistic decision-making more often showed intuitions getting people to the correct answer. They came to a surprising amount of agreement: that there were important differences between the problems they used (Kahneman's hypotheticals vs. Klein's real-world situations); and that intuitions develop when people get reliable feedback about their answers³¹. Of course, not all hypotheticals have a correct answer. Some people pretend that the blue bus hypothetical has a correct answer (liable!) and that anyone who doesn't come to that answer is a fool. But they would be wrong.

3.2. Through Thin and Thick

3.2.1. *From Abstract to Concrete to Just Right*

Creating variations of hypotheticals that are thicker does not necessarily mean that people will reason better about them. The Wason Selection Task is a deductive reasoning task that was widely used by cognitive psychologists in the 70's and 80's. In the original version, participants are shown four index cards, each with a letter (A or B) or a number (4 or 7) written on it. They are told that each card has a number on one side and a letter on the other. There is a rule that: if there is a vowel on one side, the card must have an odd number on the other side. Participants are then asked to turn over all, and only, the cards they need to turn over to make sure that the rule is enforced. Typically, in this abstract version of the task, most participants turn over the A (to check that there is a 7 on the other side). But few people turn over the other necessary card³². Conclusion? People are not good at modus tollens reasoning or falsification? People are not good at abstract reasoning but they would be better if there were context?

In another variation, the four cards had the ages of people in a restaurant on one side—15 and 50—and what they were drinking on the other side—beer or tea. The rule was: if a person is drinking an alcoholic beverage, they must be over 21. In this

³⁰ ALLEN, 2020 (this issue): ms 293-294, near fn 87.

³¹ KAHNEMAN and KLEIN, 2009.

³² That would be the 4—to make sure that there is *not* an A on the other side. WASON, 1968. Wason was a big fan of Popper and claimed that this experiment showed that people were bad at trying to falsify hypotheses.

example, most people correctly turned over the beer card and the 15 card. Why do they get it right? Because it is a culturally familiar situation? Many variations later, we learned that just because the people and activities in the vignettes are familiar, doesn't mean participants would get it right. For example, participants are not so good with the rule: if a person is eating haddock, they must be drinking gin³³. This is an arbitrary pairing although, of course, people have experience with other food pairing rules.

So, are people only good at deductive reasoning with rules for which they have specific experience? In another set of studies, participants are told about an isolated tribe of people who sometimes wear a bit of volcanic rock. The tribespeople have a rule: if one is going out at night, then one must tie a small piece of volcanic rock around one's ankle. For this version, only 28% of participants knew to check both those going out at night and those not wearing a rock. However, of participants who were told that the tribe believed that vicious spirits roamed the night but a volcanic rock around their ankle would keep them safe, 86% knew who needed to be checked³⁴. Thus, people can be good at reasoning about things way outside their experience, and which might seem arbitrary, if they are told that the rule follows a type of culturally common reasoning rule (here, one about precautions). That is, not too abstract, not simply concrete, but in that human reasoning sweet-spot of knowledge-based generalization.

3.2.2. «Real» Legal Hypotheticals

After this romp through experimental psychology, let's not forget that hypotheticals are also the bailiwick of lawyers and law professors. We use them in classes, at the end of a case excerpts in casebooks, and in legal arguments. Crucially, the hypotheticals we choose tend to be neither thin nor random. They are not thin because they assume a case that is «just like» the current one, about which we are well informed, but with one thing changed. And they are typically not random because the creator of the hypothetical knows what it means to change something that might matter; that is, might make the hypothetical case worthy of a different result. It is in these comparisons that students learn the important features, and the limits, of doctrine; that lawyers argue regarding precedent; and that courts explore possible future implications of their decisions.

Imagine for a second: what would the study of law be without the use of hypotheticals?

³³ MANKTELOW and EVANS, 1979. Beer, tea, haddock, and gin: Yes, many of these researchers were British.

³⁴ CHENG & HOLYOAK, 1989: 296-299.

4. THE BLUE BUS BEEPS

So back to the blue bus. What are some things that this hypothetical can tell us relevant to the legal system despite it being thin?

Varying the blue bus hypothetical gives us the opportunity to ask questions like: do people answer «not liable» because they don't understand simple probabilities? Are they not sensitive to the fact that 80% is greater than 50%? Do they object to using a 51% threshold for liability judgments? Or are they just answering the question quickly, without really thinking about it, but if forced to engage more deeply they would answer «liable»?

Various experiments based on the blue bus hypothetical have answered those questions—sometimes with «regular people» and sometimes with real judges as experimental participants. For example, Wells³⁵ asked participants to read variations of the blue bus story and then to estimate «the probability that a bus from the Blue Bus Company killed Mrs. Prob's dog» and to answer the question «If you were a juror in this case, would you rule against the Blue Bus Company and force them to pay damages to Mrs. Prob?» In Experiment 1, psychology student participants read either standard (naked statistics) volume-of-traffic testimony (i.e., that 80% of the traffic was from the Blue Bus Company, which owned 80% of the buses) or «weigh-attendant» (human) testimony (i.e., in which a country transportation official had noted that a blue bus had logged into a nearby weigh station about 10 minutes before the accident and the log book entries were right about bus color 80% of the time). If you've read Allen's essay, you won't be surprised to learn that participants were much more likely to award damages in the weigh-attendant (human) version (67%) than in the standard (naked statistics) version (8%). But for both versions, participants rated the probability of the damage having been done by a blue bus as about 80%.

In Experiment 2, changing the standard statistical version to a rate-of-accidents version (i.e., that 80% of accidents were caused by the Blue Bus Company), didn't change the results of approximately equal probability ratings and much more imposition of liability in the weigh-attendant variation. Experiment 3 added a new condition and also added new types of participants: Master of Business Administration (MBA) students and judges. The patterns looked the same: probability judgments always about 80% and many more plaintiff verdicts in the weigh-attendant than the naked statistics variations.

Another set of blue bus studies³⁶ found that having participants make a liability decision within 1-2 minutes versus about 6 minutes versus after a group deliberation did not affect probability estimates and did not affect liability decisions. However, in

³⁵ WELLS, 1982. Note that in the experiments that follow, in some studies participants were asked both to estimate the probability and whether they would impose liability; in other studies, those questions were asked of different participants.

³⁶ WRIGHT, MACEACHERN, STOFFER and MACDONALD, 1996.

a version in which the participants read that the Blue Bus Company owned 99.9% of the buses in the area, liability went up from 30% (in the standard condition) to only 63% (in the 99.9% condition).

I learned a lot reading these studies: within these thin hypotheticals, people are sensitive to base rate probabilities, people are more willing to impose liability with greater probabilities but, still, even very very high probabilities alone are not sufficient for them to want to impose liability – an intriguing finding.

My favorite set of blue bus experiments varies how easy it is for participants to construct counterfactuals in which the Grey Bus Company was at fault. In the basic version of the story, participants learned that the bus's tire tracks matched 8 out of 10 of the blue buses and only 2 out of 10 of the grey buses. In the alternative version, other participants learned that the tire tracks partially matched one of the blue buses and one of the grey buses, but the investigator said there was an 80% chance that it was the blue bus. In addition to being asked the probability that a blue bus killed the dog, and whether they would find against the Blue Bus Company, participants were also asked how easy it was to think that a grey bus had run over the dog. As with Wells (1982; which the initial study here was based on), participants did not differ in believing that there was about an 80% chance that it was a blue bus. However, participants varied between conditions in other judgments: those who had been told that the tracks matched 8 blue and 2 grey buses were less likely to find against the Blue Bus Company and said it was easier to think of a grey bus hitting the dog than participants in the 80%-accuracy condition. In fact, participants' ratings of ease of thinking predicted their verdicts.³⁷

The finding that the ease of creating counterfactuals affects judgments is not unique to bus accidents. During the Reign of Heuristics and Biases, a busload of studies involved vignettes in which people engaged in «normal» behavior (e.g., taking a typical route) or «unusual-for-them» behavior (which would provoke counterfactual reasoning)³⁸ and found that people were judged to be more causal of whatever bad outcome occurred to them when they had chosen the abnormal behavior rather than their usual one³⁹. Later, various theories of causal attribution looked more precisely at how people used counterfactuals in making causal judgments and found that, among other things, people make such judgments as if they consider the probability of the counterfactual (or counterfactuals) occurring and the conditional probability of its occurrence leading to the outcome in question⁴⁰.

³⁷ NIEDERMEYER, KERR, and MESSÉ, 1999.

³⁸ ROESE, 1997.

³⁹ SPELLMAN and MANDEL, 1999. The effect of abnormal vs normal circumstances is relevant to Allen's question about the usefulness of "normic" reasoning, ALLEN, 2020, this issue -- somewhere around fn 56.

⁴⁰ GILBERT, TENNEY, HOLLAND and SPELLMAN, 2015; PETROCELLI, PERCY, SHERMAN and TORMALA, 2011; SPELLMAN, KINCANNON and STOSE, 2005.

Considering more counterfactuals can lead to lower causal attributions to the questioned activity; a client's defense that creates more counterfactuals to a prosecution's (or plaintiff's) theory of the case can result in smaller chance of finding causation⁴¹. These findings are related to work by Pennington and Hastie⁴² on the story model, showing that mock jurors are more likely to find a suspect not guilty when the defense has constructed an alternative story to the prosecutor's story rather than just trying to poke holes in that story. Lawyers have plenty of anecdotal stories to this effect. Ironically, all of these findings are consistent with... relative plausibility theory.

5. HYPOTHESIZING ABOUT HOW OTHER PEOPLE WILL THINK

Hypotheticals also permeate the legal system in the form of metacognition: thinking about what other people might think. This is a theme of Saks and Spellman's book⁴³: *The Psychological Foundations of Evidence Law*. We argue that evidence rulemakers—legislators and judges—are trying to do several different things when formulating or interpreting rules. For one thing, they are, as Allen notes, juggling values⁴⁴. Trials are not just about getting to the correct answer; they also are supposed to embody other values like fairness, finality, and efficiency. But putting those other values aside for the moment, in crafting rules of evidence law (e.g., regarding character, or hearsay, or balancing probativeness against prejudice), rulemakers are acting like applied psychologists. They may consider what effect a type of evidence «should» have on jurors, consider what effect they believe it would «actually» have on jurors, and try to figure out how the evidence should (or should not) be presented so that jurors will «get it right». But what makes us think that the rulemakers themselves will get all those steps right?⁴⁵

Yet this is how it works. Hypotheticals and counterfactuals are important tools for understanding legal systems, mostly because legal systems are a product of human minds for engaging with human minds. It is indeed a complex adaptive system— involving people with different roles and values, incomplete and uncertain evidence that leads to known mistakes, and multiple institutions with different powers to revise the processes. At every step, we want to guess what others will think, and we may base our guesses on our intuitions, on their responses to thin hypotheticals, or on a deeper understanding of people's reasoning processes across people and contexts.

⁴¹ SPELLMAN and KINCANNON, 2001; TENNEY, CLEARY and SPELLMAN, 2009.

⁴² See HASTIE, 1999: 232, for a review of this research.

⁴³ SAKS and SPELLMAN, 2016.

⁴⁴ ALLEN, 2020 (this issue)—currently at top of manuscript p. 272—he doesn't use the word juggle but notes that «...trials pursue many objectives in addition to truth»—cite however useful

⁴⁵ Sometimes rulemakers do get it right. Read the book.

One hypothetical, weird or normal, thin or thick, tells us close to nothing about human reasoning—the true object of inquiry; but when put together wisely, they go far.

6. BIBLIOGRAPHY

- ALLEN, R.J. (this issue): «Naturalized epistemology and the law of evidence revisited»
- ALLEN, R.J. and LEITER, B., 2001: «Naturalized epistemology and the law of evidence», in *Virginia Law Review*, 87(8): 1491-1550.
- ALLEN, R.J. and PARDO, M.S., 2019: «Relative plausibility and its critics», in *International Journal of Evidence & Proof*, 23(102): 5-58.
- CHENG, P.W. & HOLYOAK, K.J., 1989: «On the natural selection of reasoning theories», in *Cognition*, 33(3): 285-313.
- CHENG, P.W. and NOVICK, L.R., 1990: «A probabilistic contrast model of causal induction», in *Journal of Personality and Social Psychology*, 58(4): 545-567.
- EVANS, J.S. and STANOVICH, K.E., 2013: «Dual-process theories of higher cognition: Advancing the debate», in *Perspectives on Psychological Science*, 8(3): 223-241.
- FOOT, P., 1978: «The problem of abortion and the doctrine of the double effect», in *Virtues and Vices and Other Essays in Moral Philosophy*. Oxford: Oxford University Press: 19-32.
- GIGERENZER, G. and HOFFRAGE, U., 1995: «How to improve Bayesian reasoning without instruction: Frequency formats», in *Psychological Review*: 102(4), 684-704.
- GILBERT, E, TENNEY, E.R., HOLLAND, C. and SPELLMAN, B.A., 2015: «Counterfactuals, control, and causation: Why knowledgeable people get blamed more», in *Personality and Social Psychology Bulletin*, 41(5): 643-658.
- HASTIE, R., 1999: «The role of stories in civil jury judgments», in *University of Michigan Journal of Law Reform*, 32(2): 227-240.
- HENRICH, J., HEINE, S.J. and NORENZAYAN, A., 2010: «The weirdest people in the world?», in *Behavioral and Brain Sciences*, 33(2-3): 61-83.
- KAHNEMAN, D., 2011: *Thinking, fast and slow*. New York: Macmillan.
- KAHNEMAN, D. and FREDERICK, S., 2002: «Representativeness revisited: Attribute substitution in intuitive judgment», in GILOVICH, T., GRIFFIN, D. & KAHNEMAN D. (eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge: Cambridge University Press: 49-81.
- KAHNEMAN, D. and KLEIN, G., 2009: «Conditions for intuitive expertise: a failure to disagree», in *American Psychologist*, 64(6): 515-526.
- KAHNEMAN, D. and TVERSKY, A., 1982: «On the study of statistical intuitions», in *Cognition*, 11(2): 123-141.
- MANKTELOW, K.I. and EVANS, J. St. B., 1979: «Facilitation of reasoning by realism: Effect or non-effect?», in *British Journal of Psychology*, 70 (4): 477-488.
- NIEDERMAYER, K.E., KERR, N.L. and MESSÉ, L.A., 1999: «Jurors' use of naked statistical evidence: Exploring bases and implications of the Wells effect», in *Journal of Personality and Social Psychology*, 76 (4): 533-542.
- PETERSON, C.R. and BEACH, L.R., 1967: «Man as an intuitive statistician», in *Psychological Bulletin*, 68 (1): 29-46.
- PETROCELLI, J.V., PERCY, E.J., SHERMAN, S.J. and TORMALA, Z.L., 2011: «Counterfactual potency», in *Journal of Personality and Social Psychology*, 100 (1): 30-46.
- QUINE, W.V.O., 1951: «Two dogmas of empiricism», in *Philosophical Review*, 60 (1): 20-43.
- ROEDIGER, H.L. and McDERMOTT, K.B., 1995: «Creating false memories: Remembering words not presented in lists», in *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21 (4): 803-814.
- ROESE, N.J., 1997: «Counterfactual thinking», in *Psychological Bulletin*, 121 (1): 133-148.
- RYLE, G., 1949: *The Concept of Mind*. New York: Harper & Row.

- SAKS, M.J. and SPELLMAN, B.A., 2016: *The Psychological Foundations of Evidence Law*. New York: NYU Press.
- SCHAUER, F., 2008: «A critical guide to vehicles in the park», in *New York University Law Review*, 83 (4): 1109-1134.
- SCHNEIDER, C., 2020: «The importance of context in moral judgments», in *Perspectives on Psychological Science*, 15 (2): 207-215.
- SMITH, E.R. and DECOSTER, J., 2000: «Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems», in *Personality and Social Psychology Review*, 4 (2): 108-131.
- SPELLMAN, B.A. and KINCANNON, A., 2000: «The relation between counterfactual (“but for”) and causal reasoning: Experimental findings and implications for jurors’ decisions», in *Law and Contemporary Problems: Causation in Law and Science*, 64 (4): 241-264.
- SPELLMAN, B.A., KINCANNON, A. and STOSE, S., 2005: «The relation between counterfactual and causal reasoning», in MANDEL, D.R., HILTON, D.J., & CATELLANI, P. (eds.), *The Psychology of Counterfactual Thinking*, London: Routledge Research: 28-43.
- SPELLMAN, B.A. and MANDEL, D.R., 1999: «When possibility informs reality: Counterfactual thinking as a cue to causality», in *Current Directions in Psychological Science*, 8 (4): 120-123.
- SPELLMAN, B.A. and SCHNALL, S., 2009: «Embodied rationality», in *Queen’s Law Journal*, 35 (1): 117-164.
- TENNEY, E.R., CLEARY, H.M.D. and SPELLMAN, B.A., 2009: «Unpacking the doubt in “beyond a reasonable doubt”: Plausible alternative stories increase not guilty verdicts», in *Basic and Applied Social Psychology*, 31 (1): 1-8, [Revised for legal application and published with response and replies as: «“This other dude did it!” A test of the alternative explanation defense», in *The Jury Expert*, 21 (4): 37-42.]
- THALER, R.H. and SUNSTEIN, C.R., 2009: *Nudge: Improving decisions about health, wealth, and happiness*. New York: Penguin.
- THOMSON, J.J., 1985: «The trolley problem», in *Yale Law Journal*, 94(6): 1395-1415.
- TVERSKY, A. and KAHNEMAN, D., 1974: «Judgment under uncertainty: Heuristics and biases», in *Science*, 185 (4157): 1124-1131.
- WASON, P.C., 1968: «Reasoning about a rule», in *Quarterly Journal of Experimental Psychology*, 20 (3): 273-281.
- WELLS, G.L., 1992: «Naked statistical evidence of liability: Is subjective probability enough?», in *Journal of Personality and Social Psychology*, 62 (5): 739-752.
- WRIGHT, E.F., MACEACHERN, L., STOFFER, E. and MACDONALD, N., 1996: «Factors affecting the use of naked statistical evidence of liability», in *Journal of Social Psychology*, 136 (6): 677-688.

EL TESTIMONIO ÚNICO DE LA VÍCTIMA EN EL PROCESO PENAL DESDE LA PERSPECTIVA DE GÉNERO (2)

RESPUESTA A LOS COMENTARIOS SOBRE «EL TESTIMONIO ÚNICO DE LA VÍCTIMA EN EL PROCESO PENAL DESDE LA PERSPECTIVA DE GÉNERO», PUBLICADOS EN *QUAESTIO FACTI* 1/2020

José Luis Ramírez Ortiz

Magistrado, Audiencia Provincial de Barcelona

joseluisramirezortiz@gmail.com

jose.ramirez@poderjudicial.es

RESUMEN: En el presente trabajo el autor da respuesta a las objeciones planteadas por Raymundo Gama, Roberta Casiraghi, Federico José Arena y Olga Fuentes Soriano a una de las tesis formuladas en el artículo «El testimonio único de la víctima en el proceso penal desde la perspectiva de género (1) ». En concreto, la tesis de que el testimonio único no corroborado no es suficiente para desvirtuar la presunción de inocencia. El autor concluye que el abordaje del derecho y sus instituciones con las herramientas que suministra el feminismo puede suponer un potente motor de cambio para la consecución de un derecho legal y judicial más igualitario. Al mismo tiempo, destaca que una acrítica recepción de los aportes de la perspectiva de género puede dar lugar a que acabe convirtiéndose en un instrumento más al servicio de la degradación del modelo constitucional procesal penal, enmascarando las verdaderas raíces del problema de la desigualdad de género.

PALABRAS CLAVE: Suficiencia de la prueba, estándar de prueba, testimonio único, presunción de inocencia, perspectiva de género, estereotipos.

SINGLE VICTIM'S TESTIMONY AND GENDER PERSPECTIVE IN THE CRIMINAL PROCEDURE (2) RESPONSE TO REFUTATIONS

ABSTRACT: In this paper, the author responds to the objections raised by Raymundo Gama, Roberta Casiraghi, Federico José Arena and Olga Fuentes Soriano to one of the theses formulated in the article «Single victim's testimony and gender perspective in the criminal procedure (1)». Specifically, the thesis that the sole uncorroborated testimony is not sufficient for proof of guilt in a criminal proceeding as it cannot distort the presumption of innocence. The author concludes that approaching the law and its institutions with the tools provided by feminism can be a powerful engine of change for the achievement of a more equal legal system. At the same time, the author highlights that an uncritical reception of the contributions of gender perspective can contribute to the degradation of the constitutional criminal procedure model, masking the true roots of the problem of gender inequality.

KEYWORDS: sufficiency of the evidence; standard of evidence; single testimony; presumption of innocence; gender perspective; stereotypes.

SUMARIO: 1. AGRADECIMIENTOS.— 2. RÉPLICA A RAYMUNDO GAMA.— 3. RÉPLICA A ROBERTA CASIRAGHI.— 4. RÉPLICA A FEDERICO JOSÉ ARENA.— 5. RÉPLICA A OLGA FUENTES SORIANO.— 6. ÚLTIMAS OBSERVACIONES.— 7. BIBLIOGRAFÍA

RECOMMENDED CITATION: RAMÍREZ ORTIZ, JOSÉ LUIS, 2020: «El testimonio único de la víctima en el proceso penal desde la perspectiva de género (2). Respuesta a los comentarios sobre “El testimonio único de la víctima en el proceso penal desde la perspectiva de género”», publicados en *Quaestio Facti* 1/2020», in *Quaestio facti*, 2: 339-359. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22535

1. AGRADECIMIENTOS

No puedo comenzar sin expresar mi agradecimiento a Raymundo Gama, Roberta Casiraghi, Federico José Arena y Olga Fuentes Soriano por sus pormenorizados e inteligentes comentarios que han servido para clarificar un debate complejo de gran actualidad y de indudable utilidad práctica. Creo que el abordaje del derecho y sus instituciones con las herramientas que nos aporta el movimiento feminista constituye una excelente oportunidad para que reflexionemos colectivamente sobre cómo se legisla y cómo se aplican las normas con la finalidad de promover un derecho legal y judicial más igualitario. Al mismo tiempo, considero que una acrítica recepción de los aportes de la perspectiva de género puede dar lugar a que, en vez de suponer un potente motor de cambio, aquella pueda convertirse en una fórmula instrumental para profundizar en la gradual pero perceptible, desde hace décadas, transformación del modelo constitucional procesal penal en otra cosa, para extender el poder penal en lugar de para limitarlo, enmascarando las verdaderas raíces del problema de la desigualdad de género.

El tema del testimonio único constituye un buen banco de pruebas para evaluar si es posible conjugar ambas perspectivas. Y, desde luego, para ello, los comentarios críticos de Raymundo, Roberta, Federico José y Olga, tienen un enorme valor. Reitero mis agradecimientos.

2. RÉPLICA A RAYMUNDO GAMA

En la primera parte de su estimulante comentario crítico Raymundo Gama reivindica la inseparabilidad de la perspectiva de género de las teorías feministas que se encuentran en su origen, destacando cómo si se desvinculan aquella pierde buena parte de su potencial transformador. A continuación, señala que el rendimiento probatorio de la perspectiva de género no se reduce a las funciones heurística y epistémica a las que aludo en mi trabajo, sino que comprende un espectro más amplio (la prueba en general, sus normas, prácticas, instituciones y presupuestos). Por último, cuestiona mi tesis central (la exigencia de que el testimonio único de la afirmada víctima se encuentre corroborado), señalando que desconoce las exigencias de la perspectiva de género, al operar en detrimento de las víctimas reforzando el escepticismo estructural hacia su credibilidad y la impunidad de cierta clase de delitos.

Coincido plenamente con la primera observación, que aprovecho para complementar con algunas ideas adicionales. Como es sabido, la larga trayectoria y la vinculación del movimiento feminista con objetivos de cambio social explica la existencia de diversas corrientes, lo que, en realidad, impediría afirmar que hay un solo feminismo. De hecho, en la base de las diferencias cabe advertir la mayor o menor proximidad con otros movimientos teóricos y políticos, desde el pensamiento socialista hasta el neoliberal. No son, en consecuencia, del todo coincidentes ni las premisas ideológicas ni las estrategias proyectadas para combatir el enemigo común de la desigualdad¹.

Así las cosas, si la perspectiva de género sirve para identificar aquellas instituciones, reglas y prácticas del derecho que crean, legitiman y perpetúan la discriminación, así como para señalar el camino para derogarlas, transformarlas y/o sustituirlas por otras, tanto la labor de identificación de las instituciones discriminatorias como la de la elección de las estrategias para su erradicación o transformación estará condicionada, en buena medida, por su conexión con otras corrientes de pensamiento político. No pretendo negar la autonomía de algunos aportes del feminismo, pues del mismo modo que desde el feminismo socialista se es consciente de que la discriminación de la mujer no depende en exclusiva de la adscripción a una determinada clase social, de modo que, producida la supresión del sistema de clases, la dominación masculina quedaría erradicada, pues existen otras fuentes de discriminación no dependientes de la lucha de clases, desde el feminismo liberal se es consciente del hecho de que, si se dejan intactas las estructuras socioeconómicas, difícilmente podrá combatirse la subyugación machista. Mi intención es advertir del dato de que los análisis no se realizan en una tabla rasa, sino desde concepciones políticas y culturales más amplias, que pueden divergir hasta ser opuestas, así como que tal aproximación nos permite identificar discursos que, bajo el ropaje terminológico de la perspectiva de género, pueden no acabar vertiendo hacia la emancipación sino hacia el control.

¹ Véanse WATKINS, 2018 y VIVAS LARRUY, 2016.

Del mismo modo, comparto la segunda observación de Gama, que también me permito ampliar. Si en mi trabajo ya destaqué la utilidad de la perspectiva de género para analizar y comprender la organización social, parece claro que a la luz de las herramientas analíticas que la integran puede evaluarse todo el derecho penal sustantivo y procesal y sus instituciones. Bajo este ángulo, la perspectiva de género se proyecta, en primer lugar, sobre el momento legislativo, identificando aquellas soluciones que, compatibles con los rasgos del modelo garantista y de la presunción de inocencia, sean más adecuadas para suprimir las diferencias entre hombres y mujeres. En segundo lugar, vierte en el momento investigativo-jurisdiccional (determinación de los hechos, fijación de la significación típica y determinación de las consecuencias jurídicas). Por último, resulta de particular interés en el momento de la ejecución, ya sea desde la perspectiva estrictamente penal como desde la penitenciaria. Además, en todos estos casos, la posición de la mujer puede tomarse en consideración tanto cuando aparece como víctima como cuando interviene como persona investigada, acusada o penada. El objeto de mi trabajo se circunscribía a un único aspecto: qué puede aportar la perspectiva de género, y en qué grado, al problema del testimonio único en el proceso penal. Dejé así de lado deliberadamente otros posibles rendimientos, tanto probatorios como extraprobatórios, que he tenido ocasión de desarrollar en una reciente monografía², pero eso no significa que los rechace.

Creo así que el verdadero motivo de divergencia radica en el aspecto central de mi tesis: la declaración testifical única no corroborada de la afirmada víctima no puede prestar el suficiente respaldo probatorio a la hipótesis acusatoria para darla por acreditada. A juicio de Raymundo Gama, tal aserto carece de perspectiva de género pues provoca la impunidad de aquellos delitos que refuerzan la dominación estructural de los hombres sobre las mujeres, como los sexuales, en los que no suele haber otros medios de prueba. Además, frente a lo que sostengo, a su entender ni el principio de valoración racional de la prueba ni el derecho fundamental a la presunción de inocencia resultan afectados cuando la condena se produce sobre la base de un testimonio único no corroborado, pues cabría acudir, en línea con la posición que mantiene Jordi Nieva, a una prueba pericial psicológica sobre la credibilidad del testimonio que serviría para obtener «algunos elementos en apoyo de la credibilidad de la víctima»

El juicio sobre los rendimientos que pueden obtenerse de la perspectiva de género en el derecho penal sustantivo y procesal está íntimamente ligado al juicio que hagamos sobre cuáles han de ser las finalidades y metas de esta rama del ordenamiento y cuál ha de ser el papel de los jueces dentro del proceso. En mi artículo partí, en línea con los planteamientos de Ferrajoli y de Vives Antón sobre la configuración del modelo de derecho penal y procesal del Estado constitucional, de la premisa de que la función primordial de aquel es mantener a raya los excesos punitivos del Estado. Lógicamente, son el vehículo del castigo de la persona responsable de la comisión del

² RAMÍREZ ORTIZ, 2019.

hecho delictivo (y, desde este punto de vista, serían instrumentos de aplicación el *ius puniendi*), pero, teóricamente, podría optarse por un modelo de terror punitivo en que el Estado pudiera sancionar a un sujeto sin norma penal y sin proceso previos, de ahí que lo que caracterice al modelo constitucional penal y procesal sea su función de garantía. Pues bien, si esto es así, no constituiría una rama del derecho tendencialmente orientada al cambio social, lo que, en apariencia, podría resultar contradictorio con los postulados del feminismo y hacernos concluir que las herramientas conceptuales que nos ofrece este movimiento serían inaplicables en el derecho penal sustantivo y procesal³.

Tal inaplicabilidad, ciertamente, se produciría de entenderse que lo que se exige es la instauración de un modelo militante o de combate de derecho penal excepcional ante una situación de alarma o emergencia social, pero también de estimarse que tal perspectiva siempre y en todo caso vulnera las garantías constitucionales. Sin embargo, defendiendo una tercera opción: pueden obtenerse ciertos rendimientos de la perspectiva de género en el análisis e interpretación de la ley penal y procesal y de sus instituciones. Ahora bien, nunca debemos perder de vista la función última de esta rama del derecho estatal de límite del poder público. Ello exige tener siempre presente las líneas que delimitan el sistema de garantías, pero también someter a juicio crítico posicionamientos rígidos que pueden enmascarar aplicaciones discriminatorias del derecho penal falsamente garantistas. Como señala con acierto Arduino⁴, se trata de romper el cerco de impunidad que el sistema penal construyó tras siglos de invisibilización de las cuestiones de género sin caer en el punitivismo demagógico, pues en algunas reivindicaciones realizadas desde el discurso de género, aun bienintencionadas, pueden anidar sesgos y manipulaciones autoritarios.

Sobre la base de las precedentes consideraciones creo que puedo seguir sosteniendo que toda condena penal basada en exclusiva en una prueba testifical no corroborada compromete el derecho a la presunción de inocencia. Tal y como defendiendo en mi trabajo, la constitucionalización de dicho derecho y su entronque con el deber de motivar las decisiones judiciales y el principio de prohibición de arbitrariedad de los poderes públicos, exige un modelo objetivo o intersubjetivo del fenómeno probatorio, en el que el foco de atención se proyecta sobre los criterios normativos de justificación. La pregunta relevante, por tanto, es si el testimonio único de la víctima es apto para probar la acusación, no si se puede o no se puede creer. Y lo que vengo

³ En este sentido, Erice MARTÍNEZ, 2018, destaca que el derecho penal no ofrece soluciones totales frente a la desigualdad y la discriminación, sino que, como sistema estructurado de control social, se limita a reaccionar frente la comisión de hechos considerados previamente como peligrosos para los intereses o bienes que se desea proteger. «Por ello, las soluciones a los problemas de desigualdad y discriminación de la mujer en una sociedad discriminadora y violenta hacia ellas nunca podrán ser obtenidas mediante el derecho penal, ni tan siquiera, si se llegaran a construir sus respuestas desde políticas criminales que orientaran la labor legislativa desde una perspectiva que tuviera en cuenta esa situación de desigualdad y discriminación real».

⁴ ARDUINO, 2017.

a sostener, en síntesis, es que, toda vez que los aspectos personales relativos al testigo (gestualidad, apariencia física, etc) son equívocos, y que la coherencia interna del relato no sirve por sí sola para testar la fiabilidad del medio de prueba, acabamos desembocando en aquello que queríamos evitar: un juicio de credibilidad sobre la persona del testigo, en el que de la mera calificación del medio como atendible habrá de inferirse la prosperabilidad de la tesis acusatoria. Y, como sigo defendiendo, lo probatoriamente relevante no es la credibilidad subjetiva sino la fiabilidad de la información, pues un individuo puede ser personalmente muy fiable, (v.gr. por gozar de gran reputación y haber mantenido una trayectoria vital éticamente intachable), pero proporcionar una información no fiable por múltiples factores que la psicología del testimonio tiene identificados que pueden concurrir en el momento de la codificación, en el de la retención o en el de la recuperación de la información.

Como sostiene de Paula Ramos⁵ la valoración exige siempre un recomienzo epistémico. En este punto, sin perjuicio de sus valiosos rendimientos en otros ámbitos, cabe advertir lo erróneo del planteamiento que pretende incorporar al fenómeno probatorio en el proceso penal el lema «Yo sí te creo», que se ha generalizado en ciertos sectores del movimiento feminista. Dicho lema, por lo que nos ocupa, parece tener como punto de partida la tesis de que lo que se transmite en el testimonio es una creencia, una descripción subjetiva de un estado mental. Sin embargo, en el testimonio no hay una transferencia de conocimiento o creencias de una persona a otra. Cada testimonio de una persona a otra implica una comunicación y, a partir de ella, un «recomienzo epistémico», pues quien recibe la información debe, a su vez, ser «epistémicamente responsable» para confirmar o refutar la información recibida. Esto es, quien recibe el testimonio ha de valorar su grado de confirmación. Y, si bien en ciertos contextos en los que la verdad importa menos (v.gr. informaciones poco trascendentes transmitidas en el marco de relaciones de amistad), es razonable que el receptor del testimonio «baje la guardia» ante la persona que confía, en contextos en los que la verdad adquiere un valor preeminente por sus efectos sobre la vida de las personas la labor de confirmación o refutación es inexcusable, lo que remite necesariamente a la corroboración externa. Y aquí, no hemos de pensar solo en contextos institucionales, como el del proceso, sino en otros, como el médico. Así, probablemente no aceptaríamos un diagnóstico de una enfermedad grave proveniente de un médico que nos dijera que alcanzó una alta convicción subjetiva a tal efecto sin realizar previamente pruebas conforme a los protocolos médicos exigibles. No se trata de «creer», sino de «probar»⁶.

De este modo, llegamos al núcleo argumental: razones político-institucionales impiden otorgar al testimonio único no corroborado el valor que se pretende. Camarena Grau⁷ formula la pregunta en los siguientes términos: ¿puedo ser condenado a

⁵ DE PAULA RAMOS, 2019.

⁶ SAN MIGUEL BERGARECHE, 2018.

⁷ CAMARENA GRAU, 2017a.

una pena muy grave por la mera declaración de otro, que está enfrentado a mí, sin comprobar aquellos aspectos de su declaración acreditables, sin aparente dificultad, por otros medios de prueba? Y da una respuesta que combina aspectos epistémicos y políticos. Desde la primera perspectiva, afirma que no es razonable aceptar una creencia en condiciones de evidencia insuficiente. Y, desde la segunda, agudamente refiere que la legitimidad de la respuesta a la pregunta exige que esta sea racionalmente aceptada por todos aquellos ciudadanos a los que ha de vincular, de modo que la posición desde la que deberíamos contestar habría de ser una posición de imparcialidad (esto es, haciendo abstracción de las contingencias y sin saber si vamos a intervenir en un proceso penal o, en caso de participar en él, sin saber qué posición vamos a ocupar ni en qué momento histórico va a tener lugar dicha participación). Y, a tal fin, recuerda el mandato que expresa el artículo 24.2 de la Constitución Española, que consagra el derecho a la presunción de inocencia, añadiendo que, en caso de no exigirse corroboraciones, el poder del Estado no estaría sujeto a verdaderas reglas.

Veámoslo con un ejemplo. Una sentencia de condena podría razonar en un supuesto en el que se acusó por violación al varón pareja de la mujer víctima que esta no presenta déficits de credibilidad subjetiva, realizó un relato coherente y fue persistente en la incriminación, por lo que le otorgó pleno valor de cargo. Pero también podría señalar que aunque no hubo persistencia en la incriminación, la perspectiva de género enseña que las mujeres que padecen violencia a manos de sus parejas suelen presentar pensamientos y emociones contrarias hacia aquellas, sentimientos contradictorios de amor y odio, autonomía y dependencia, reconociéndose como víctimas pero también como corresponsables de la situación, lo que explicaría las retractaciones en las denuncias, en las decisiones sobre la ruptura de la relación o el mantenimiento de la convivencia, y condenar igualmente. Del mismo modo, la sentencia de condena podría argumentar que, aunque el relato de la víctima no persistente presentó trazos de incoherencia, al tratarse de víctima de género especialmente vulnerable (v.gr. por enfermedad mental), la ausencia de coherencia en algunos aspectos podía encontrar explicación en su situación de vulnerabilidad, y condenar igualmente. De idéntica manera, en los tres supuestos las sentencias podrían ser absolutorias, y razonar, en el primero, que el relato del varón acusado provenía de una persona subjetivamente creíble, fue coherente y persistente a lo largo del procedimiento, por lo que no había motivo para otorgar mayor peso acreditativo a ninguna de las declaraciones. En el segundo, que la retractación, en ausencia de otros elementos informativos de corroboración impedía que la prueba pudiera desvirtuar la presunción de inocencia, pues la existencia de distintas versiones obrantes en el procedimiento no permitía confiar en ninguna de ellas. Y en el tercero, que la incoherencia del relato, unida a la existencia de distintas versiones proporcionada por la afirmada víctima obrantes en el procedimiento, impedía otorgar el valor de prueba de cargo al testimonio.

Si nos colocamos por un momento en la posición del acusado, claramente advertimos que algo falla en un sistema en el que, frente a una acusación que lleva asociada penas muy graves, todas las opciones son posibles. Si nos situamos en la posición del

órgano de apelación, descubrimos igualmente que el control institucional carece de reglas para la resolución de la controversia que se plantea en segunda instancia. Si la presunción de inocencia, al parecer de Vives Antón⁸, no es solo un derecho, sino un verdadero modelo, el elemento central y esencial del sistema, la expresión abreviada del conjunto de derechos que definen el estatuto de todo ciudadano en tanto que susceptible de ser sometido a una investigación o proceso, o, en palabras de López Ortega⁹ una garantía constitucional en cuya virtud el legislador está obligado a diseñar un método de enjuiciamiento criminal que atienda a la protección de la inocencia y la libertad de la persona imputada, creo que podemos convenir que la suficiencia del testimonio único no corroborado para la condena no es respetuosa con dicho modelo o método.

No comparto, por otra parte, la objeción de que siempre puede acudir a una prueba pericial psicológica sobre la credibilidad del testimonio para suplir los problemas del testimonio no corroborado. En primer lugar, porque tal prueba no deja de ser una metaprueba y, como tal, no solventaría los problemas de fondo que suscita el testimonio único no corroborado. Pero es que, además, en el ámbito de la psicología del testimonio no solo se ha llamado la atención sobre los riesgos de sobrevalorar esta prueba cuando su objeto lo constituyen las declaraciones de menores de edad, sino que está generalizada la posición de quienes estiman que la misma carece de fiabilidad científica en los supuestos de personas adultas¹⁰.

Por último, creo que aducir que tal tesis carece de perspectiva de género confunde planos, pues si bien es compartible la afirmación de que ha de ponderarse la efectividad de la igualdad al interpretar las normas jurídicas para potenciar los fines de la tutela antidiscriminatoria, no lo es menos que no es fácil ver de qué manera se protege mejor dicho principio defendiendo un sistema probatorio más abierto a la arbitrariedad y el error. En definitiva, cuando el juzgador decide sobre la hipótesis acusatoria a la vista de un solo testigo no corroborado, nunca sabremos si condenó deliberadamente de forma injusta. Pero tampoco, si se equivocó, dadas las particularidades de la prueba testimonial.

3. RÉPLICA A ROBERTA CASIRAGHI

Mi coincidencia con el comentario de Roberta Casiraghi es plena en algunos aspectos que creo muy relevantes, tales como los riesgos de la confianza ciega hacia la prueba testifical, los de la utilización acrítica de la perspectiva de género, la inadmisibilidad de la introducción de tipologías de testigos en función de la naturaleza o clase del delito del que afirmen haber sido víctimas o la defensa del carácter absoluto del

⁸ VIVES ANTÓN, 2011

⁹ LÓPEZ ORTEGA, 2000.

¹⁰ MANZANERO, 2010.

derecho a la presunción de inocencia. También, en la necesidad de preservar el principio de contradicción, aun en aquellos casos en los que se acuda a los mecanismos de la preconstitución del medio de prueba o a los de evitación de la confrontación visual o interrogatorio directo para evitar una eventual victimización secundaria.

La discrepancia radica en el alcance de mi tesis central. A juicio de la autora, es cuestionable la vinculación entre presunción de inocencia e insuficiencia del testimonio único no corroborado, ya que tal asociación confunde los espacios de las reglas de valoración y de decisión. En su opinión, la presunción de inocencia es solo una regla de decisión que opera después de aplicar las correspondientes reglas de valoración para resolver las situaciones de incertidumbre. No puede, por ello, incidir en la valoración de la prueba, que es previa. En segundo lugar, cuestiona la preterminación legal de los valores probatorios, sean positivos o negativos, afirmando que las garantías centrales para evitar los riesgos de abuso o arbitrariedad radican en la contradicción en la práctica del medio de prueba y en la obligación judicial de motivar las decisiones.

Desde finales de los años 90¹¹ un prestigioso sector de la doctrina, proveniente del campo de la filosofía del derecho, ha desarrollado en España un importante trabajo de instauración y análisis de los criterios normativos de justificación en materia probatoria, al tiempo que de modelos racionales de valoración, con la finalidad de dotar de una mínima solvencia a la labor judicial de fijación de hechos¹². Para ello, se parte de la idea de que en el ámbito de la prueba la presunción de inocencia se desdobra en una regla y en un estándar probatorio. La regla probatoria exige para la condena la existencia de una actividad probatoria de cargo, esto es, suficiente, lícita, y producida en el juicio oral con todas las garantías. Por su parte, el estándar probatorio sirve como criterio para determinar cuándo se entiende acreditada la hipótesis acusatoria.

A la luz de tal distinción creo que cabe dar una primera respuesta a la crítica, pues lo que vengo a sostener es la insuficiencia del testimonio único no corroborado para fundar la condena, con lo que aludo a la presunción de inocencia en su vertiente de regla probatoria, resaltando que este medio de prueba, como otros (v.gr. la identificación visual mediante reconocimiento en rueda no corroborada o la declaración testifical de referencia), debido a su debilidad intrínseca, nunca puede, por sí solo, sustentar una sentencia de condena. No obstante, la crítica es sólida, y creo que merece la pena continuar el diálogo.

Efectivamente, como Casiraghi señala, suelen distinguirse dos momentos que integrarían lo que convencionalmente denominamos valoración de la prueba: la valoración, en sentido estricto, y la decisión¹³. La valoración, en sentido propio, consiste

¹¹ Liderando tal movimiento autores de reconocido prestigio tales como Juan Igartua Salaverría, Marina Gascón, Jordi Ferrer Beltrán y Carmen Vázquez entre otros.

¹² En el campo judicial, magistrados como Perfecto Andrés Ibáñez, bajo la influencia de la doctrina italiana, ya venían trabajando en este ámbito.

¹³ Véanse FERRER BELTRÁN, 2007 y GASCÓN, 2010.

en evaluar el grado de apoyo que las pruebas prestan a la hipótesis acusatoria¹⁴. A tal efecto, quien juzga ha de valorar individualmente todos los medios de prueba practicados, tanto los de cargo como los de descargo, e identificar las informaciones provenientes de cada medio de prueba que considere provisionalmente relevantes y fiables y las razones para ello. A continuación, valorará conjuntamente dichas informaciones probatorias y establecerá qué relaciones existen entre ellas y con los hechos objeto de juicio, y establecerá cuáles estima definitivamente relevantes y fiables. El segundo momento analíticamente trascendente dentro del fenómeno global de la valoración probatoria es el de la determinación de si la probabilidad de ocurrencia alcanzada por esa hipótesis es suficiente para afirmar su existencia. En forma gráfica, Gascón señala que «la valoración, en sentido estricto, consiste en medir la probabilidad; la decisión, en aplicar al resultado de esa valoración el estándar de prueba establecido».

El estándar de prueba vendría integrado por los criterios que indicarían cuándo se habría conseguido la prueba de un hecho. Aquí, se trata de determinar si el grado de corroboración que las pruebas aportan a la hipótesis acusatoria es suficiente para tenerla por probada. Y aquí, también, está convencionalmente aceptada, en los últimos años, que tal estándar, que se encuentra en el artículo 24.2 CE, encuentra traducción lingüística en la fórmula «más allá de toda duda razonable».

La metodología examinada, propia del modelo de valoración racional, tiene una evidente vocación pedagógica y, en este sentido, una sana vertiente prescriptiva, por lo que debe aplaudirse y tomarse como verdadera hoja de ruta. Con todo, creo que puede ser matizada desde un punto de vista descriptivo en algún aspecto concreto. Así, cabe señalar que en los procesos valorativos suele producirse una mayor interacción recíproca que la descrita, lo que dificulta la separación nítida y tajante en los momentos que analíticamente se diferencian.

En esta línea, el juicio de fiabilidad individual de cada prueba está íntimamente conectado con el que se realiza en el momento de la valoración conjunta. Aunque sean operaciones que, por razones didácticas o metodológicas, conviene distinguir, en la práctica no son exactamente secuenciales, sino que interactúan en relación de circularidad pues en la valoración individualizada del medio a testar suele contrastarse la información que proporciona el medio de que se trate con las informaciones que provienen de otras fuentes de prueba de modo que, en caso de disparidad absoluta, puede llegar a calificarse de no atendible el medio cuestionado en atención al resultado de dicho contraste. En suma, el juicio de fiabilidad del medio individual de prueba puede depender en ciertos casos de los datos probatorios provenientes de otros medios.

¹⁴ El juicio se instrumenta para determinar si el acusado ha cometido el hecho delictivo, no para realizar reconstrucciones históricas con prescindencia de tal meta. Demostrar la inocencia no es el objetivo del proceso penal, del cual no puede pretenderse que ofrezca un relato de los hechos tal y como sucedieron, sino solo asegurarse de que el acusado es la persona que ha cometido el hecho del que se le acusa.

Por otro lado, la presunción de inocencia puede producir efectos no solo en el momento de la decisión, sino también en el momento de la valoración en sentido estricto, pues la existencia de una duda razonable sobre, v.gr. la presencia de un móvil espurio en la declaración del testigo, se resuelve ya en el mismo momento de testar su fiabilidad y, por tanto, de otorgarle un peso probatorio inicial o *prima facie*.

Con ello considero que puede seguir defendiéndose la vinculación entre la presunción de inocencia y la insuficiencia probatoria del testimonio único no corroborado. No creo, por el contrario, que tal insuficiencia pueda suplirse ni mediante el sometimiento del medio de prueba a la contradicción plenaria ni mediante la motivación de la decisión judicial.

En cuanto a lo segundo, me remito a la respuesta dada al comentario crítico de Gama. Y, respecto de lo primero, estimo que, ante medios de prueba que presentan problemas generales de fiabilidad, fundamentalmente cuando no pueden ser contrastados, si se trata de minimizar los errores o abusos de poder y de salvaguardar la protección de los ciudadanos, presumidos inocentes, ante los riesgos de condenas injustas, siempre es preferible optar por una regla (en nuestro caso, de prueba legal negativa) que ante un estándar (el juez decide siempre que se salvaguarde el principio de contradicción), pues la discrecionalidad acaba traduciéndose en imprevisibilidad y, en no pocos casos, en arbitrariedad. A ello cabe añadir el dato de que si el testigo, v.gr. por error, está íntimamente convencido de la realidad de lo que afirma, poco podrá hacer el contradictorio para dilucidar el valor de verdad de un testimonio sincero pero falso, cuando no se dispone de elementos externos para someter a contraste sus afirmaciones.

4. RÉPLICA A FEDERICO JOSÉ ARENA

Con Federico José Arena hay también muchos puntos de coincidencia, como reconoce en su trabajo. Entre ellos, la conciencia de las limitaciones del derecho penal como medio de solucionar problemas sociales estructurales y de la pena como forma de solventar la violencia en casos concretos.

La primera de sus sugerentes notas críticas me permite traer a colación la idea, de tintes foucaultianos o bourdieuanos, de que el campo jurídico está integrado por un conjunto de sujetos, cada uno de ellos poseedor de saberes teóricos y de prácticas reconocidos, que pugna por privilegiar determinados saberes y prácticas y de erradicar otros. No es posible, por ello, disociar la práctica discursiva de cada sujeto de la posición que ocupa en dicho campo y de sus pretensiones. Al escribir sobre la insuficiencia probatoria del testimonio único de la víctima no corroborado lo hacía desde mi experiencia como juez español de una corte de apelación.

Siendo ese el lugar desde el que hablo matizaría la afirmación de Arena de que, en la práctica, el alcance de la tesis que defiende es menor del que parece, ya que prácticamente siempre cabe identificar elementos de corroboración, si esta se toma

en sentido débil, con lo que los supuestos verdaderamente problemáticos serían excepcionales.

Efectivamente, los supuestos de testimonios únicos no corroborados no son los estadísticamente más frecuentes, pues suelen concurrir otros medios de prueba. Con todo, las condenas basadas en testimonios únicos no corroborados no son residuales. Sobre la base del argumento de que, en otro caso, se produciría un efecto de impunidad al tratarse de delitos cometidos en la clandestinidad, esto es, transcurren entre el autor y la víctima, sin testigos, se dictan muchas sentencias en las que se imponen penas muy graves sin otra base probatoria, soslayando con ello la petición de principio en que se incurre (al dar algo por probado contra reo), al tiempo que la capciosidad en el razonamiento, como ha puesto de relieve Andrés Ibáñez¹⁵, pues las circunstancias del entorno en que suelen acontecer los delitos de este tipo que, según se afirma no han de perjudicar a la víctima, sí deben darse en perjuicio de la persona acusada.

Por otra parte, es conveniente recordar que en muchas ocasiones suele atribuirse en la praxis judicial valor de corroboración a elementos informativos que carecen de tal calidad (v.gr. las pruebas periciales de credibilidad de los testimonios). Además, a raíz de la introducción por la Sala II del Tribunal Supremo de la doctrina del carácter privilegiado del testimonio de la afirmada víctima de la violencia de género, testimonio que sería suficiente para la condena, aun no corroborado, se ha agudizado en los órganos judiciales la desproblematización de la valoración probatoria. Esta viene así a construirse argumentalmente a través de una fórmula estereotipada, aparentemente de nuevo cuño, que no es sino recuperación de fórmulas clásicas, ya que basta con una doble afirmación para dar por acreditado el relato acusatorio: en primer lugar, el testimonio de la víctima es privilegiado; y, en segundo lugar, el juzgador, en virtud del principio de inmediación, ha creído la versión de la víctima. Por lo que respecta a la apelación, ello ha dado lugar a una renuncia al control de los hechos en segunda instancia. Ilustrativa de este estado de cosas es la Sentencia 2/2019 de la sección tercera de la Audiencia Provincial de Jaén (ROJ SAP J 128/2019) que, enfrentada al supuesto del testimonio único, tras reafirmar que la víctima presta un testimonio privilegiado o cualificado «al haber presenciado el hecho» como dice la Sala II, concluye: «En el presente caso, el juzgador de instancia percibió que la víctima, pareja del acusado, decía la verdad, narrando el maltrato y las vivencias que ello le había producido». Esto es, comoquiera que el testimonio de la víctima tiene un peso probatorio específico y superior al del acusado, y el juez «se creyó» a aquella, la condena está justificada. Como puede verse, nos encontramos ante una peculiar combinación del sistema de prueba legal y de la doctrina de la íntima convicción, que impide todo control institucional.

Arena realiza dos aproximaciones a la perspectiva de género (como exigencia contraepistémica, que impone la necesidad de privilegiar otros fines distintos de la búsqueda de la verdad, y como exigencia epistémica, que exige tener por verdadera la

¹⁵ ANDRÉS IBÁÑEZ, 2019.

creencia apoyada en el testimonio de la víctima) a fin de evaluar cómo la distinción repercute sobre la presunción de inocencia. Señala que, concebida como exigencia contraepistémica, podría ser defendida con el argumento de que con el fin de cambiar la historia de discriminación en perjuicio de las mujeres sería necesario, en algunos casos, sacrificar el valor de verdad, considerando más valioso estimar como verdadera la creencia fundada en el testimonio único de la víctima aun cuando ello pudiera aumentar la posibilidad de condenar a un inocente. De ello sigue que cuando afirmo que la presunción de inocencia es un derecho absoluto estoy precluyendo la discusión, que debería mantener en el mismo plano, utilizando otra clase de argumentos a favor de la presunción de inocencia.

Cuando aludí en mi trabajo a la naturaleza absoluta del derecho lo hacía a propósito de algunas posiciones doctrinales que sostienen que, como todo derecho fundamental, no es absoluto y, en consecuencia, es susceptible de sujeción a ponderación cuando entre en conflicto con otros intereses relevantes, entre los que se encuentran la integridad corporal, la libertad individual y la libertad sexual de las mujeres frente a las agresiones de los hombres. A continuación, exponía las razones por las que considero que es política y jurídicamente imposible en el Estado constitucional restringir o atenuar un derecho en el que no cabe distinguir un centro y una periferia, pues todo es núcleo duro ¹⁶. En suma, como ha sostenido Camarena Grau ¹⁷ si en el Estado constitucional se plasma la concepción kantiana del derecho como conjunto de condiciones bajo las cuales el arbitrio de una persona puede conciliarse con el arbitrio de otra según una ley universal de la libertad, de poco serviría tal enunciación programática si, al mismo tiempo, el Estado no reconociera el derecho a la presunción de inocencia cuando pretende privar a una persona de alguno de sus derechos básicos. El carácter estructural de la presunción de inocencia como guía procedimental cuando se trata de privar de derechos, al que me he referido con anterioridad, creo que permite afirmar, frente a lo que sostiene Arena, que hay diferencias categóricas entre dicho derecho y otros intereses legítimos o que juegan en planos distintos. Por otra parte, si la privación estatal de derechos fundamentales fuera posible sin una certeza objetiva sobre los hechos más allá de toda duda razonable cuando se tratara de tutelar la vigencia del principio de igualdad entre hombres y mujeres, dejando a un lado la cuestión acerca de si es posible tal tutela en un proceso penal, no solo no podríamos estar seguros sobre si efectivamente estamos otorgando tutela a dicho derecho. Además, estaríamos afirmando que cuando la acusación se formula por cierta clase de delitos, la persona acusada no tiene todo el derecho a toda la presunción de inocencia en todos los casos, lo que, paradójicamente, lesionaría la aplicación igualitaria del artículo 24.2 de la Constitución.

Por otra parte, Arena afirma que, como exigencia epistémica, sostener que el testimonio único no corroborado permite desvirtuar la presunción de inocencia no

¹⁶ ANDRÉZ IBÁÑEZ, 2019.

¹⁷ CAMARENA GRAU, 2017b.

implica cambiar la regla de la duda razonable, sino aseverar que la duda razonable queda descartada cuando se cuenta con el testimonio de la víctima. El argumento es circular. Creo, por el contrario, que si afirmamos que la duda razonable se descarta mediante una prueba que presenta todos los déficits que describo en mi trabajo, rebajamos el umbral y desnaturalizamos el derecho.

Arena señala, no obstante, que hay varios casos en los que basta con el testimonio de una única persona para considerar como verdadera una creencia (v.gr. la palabra de un único oficial de policía o la de un secretario judicial para tener por producidos ciertos actos), y que en tales casos la razón para ello radica en que los rasgos de la persona son determinantes para considerar creíble y fiable su testimonio (v.gr. el cargo de funcionario público). A su entender, lo mismo cabría predicar de la persona que afirma ser víctima de violencia de género, sobre la base de generalizaciones basadas en la experiencia. Discrepo de las premisas del argumento. Del mismo modo que, en el proceso penal, es incompatible con la presunción de inocencia la presunción de victimización, es incompatible con dicho derecho la presunción de veracidad de las afirmaciones de los funcionarios policiales. Esta doctrina de la presunción de verdad policial, de raigambre preconstitucional, vinculada a una concepción autoritaria de orden público, no tiene amparo en ninguna norma jurídica. En cuanto a los secretarios judiciales, la cualidad de fedatarios públicos se reduce a aquello que afecta a su labor de documentación de actuaciones judiciales, con lo que no hay identidad ni similitud de razón en los casos. En cuanto a las generalizaciones, no queda claro cómo podría determinarse estadísticamente si las personas que afirman ser víctimas dicen la verdad. Si el cálculo se efectuase tomando en consideración la ratio de condenas, el resultado distaría de ser concluyente, dadas las tasas de absoluciones. Y si se tomaran en consideración otros parámetros, la reconstrucción histórica habría de realizarse prescindiendo de las garantías establecidas para las personas acusadas, por lo que constituiría un abuso pretender obtener rendimientos probatorios de tales reconstrucciones. Por último, aun cuando pudiera afirmarse que estadísticamente las víctimas de cierta clase de delitos suelen ser sinceras, ello no significa todavía que sean veraces. Y aunque lo fueran, tal dato nada indicaría sobre si es cierto lo que narra la víctima en el caso concreto objeto de enjuiciamiento.

A continuación, Arena profundiza en la incidencia de los sesgos y estereotipos en la valoración de las declaraciones. Al hilo de su exposición, creo que convendría aclarar un malentendido que no se desprende de su texto pero que a mi juicio se encuentra extendido: en los casos de violencia de género, las sentencias de condena se dictan con perspectiva de género y las absolutorias sin ella. Sin embargo, que la sentencia sea de condena no significa necesariamente que se haya dictado con perspectiva de género, pues, vgr., la sentencia puede realizar una valoración estereotipada de la prueba y emplear un lenguaje sexista. Del mismo modo, que sea absolutoria tampoco significa que se haya dictado sin esa perspectiva, pues una misma absolución puede deberse a un déficit de perspectiva de género, pero también fundamentarse de forma constitucionalmente irreprochable. Veamos un ejemplo.

Un hombre y una mujer se conocen una noche en una discoteca. Deciden dar un paseo por una playa cercana. La mujer denuncia al hombre por haberla agredido sexualmente. El hombre objeta que la relación fue consentida. La prueba practicada en juicio oral se ha reducido a las declaraciones de una y otro. No existen elementos probatorios de ningún tipo que corroboren la declaración de la mujer ni que evidencien la presencia de asimetría en la concreta configuración de la relación, aun episódica, entre ambos. La sentencia podría fundar la absolución en diversos estereotipos: a) si, tras salir de una discoteca, una mujer decide dar un paseo nocturno por la playa con un hombre que acaba de conocer, es signo de que prestó un consentimiento anticipado al acto sexual subsiguiente; b) existe una inclinación por parte de las mujeres en denunciar de forma infundada el haber sido víctimas de violencia sexual; o; c) la falta de petición de ayuda o auxilio por la víctima a terceros constituye otro signo de existencia de consentimiento. Ahora bien, la resolución absolutoria podría afirmar que no se pone en duda el testimonio, pero que el mismo no es prueba suficiente para condenar en un modelo basado la presunción de inocencia como regla de juicio.

Con todo, supuestos como el examinado no siempre y en todo caso deben abocar a la absolución, lo que permite rechazar la crítica de la impunidad de estas conductas, pues es probable que una investigación realizada con perspectiva de género hubiera permitido hacer acopio de datos probatorios o de contexto relevantes que, posteriormente, podrían haberse empleado en el juicio oral como elementos de corroboración del testimonio único. Así, v.gr. podría haberse recibido declaración testifical a las personas que hubieran tenido contacto con ambos implicados antes y después de los hechos, con la finalidad de contrastar sus propósitos y estado, en el primer caso, y sus reacciones, en el segundo. Aquí, deben abandonarse las valoraciones estereotípicas, pues no todas las víctimas reaccionan del mismo modo, pero ha de reiterarse la importancia de completar el medio probatorio, insuficiente por sí solo, con información proveniente de otras fuentes. Además, es conveniente deshacer un equívoco: el dato de que la víctima no presente signos de haber sido agredida no significa que hubiera prestado el consentimiento sexual, pero tampoco que no lo hubiera prestado. En este sentido, es un dato neutro. Sin embargo, la presencia de signos de agresión es un dato orientado a la hipótesis inculpatoria, de peso corroborador.

Por último, Arena se refiere a la interesante tesis de los sesgos implícitos defendida por Saul. En línea con lo que apunta creo que los trabajos de David Eagleman y otros neurocientíficos sobre la actividad no consciente del cerebro¹⁸ están abriendo nuevas líneas de estudio que, sin duda, nos obligarán a reflexionar sobre las bases de nuestros sistemas jurídicos. Esta reflexión, necesaria e inaplazable, deberá hacerse con especial cuidado para identificar qué elementos e instituciones deben ser transformadas, cuáles convendría suprimir y qué otras deben ser conservadas a toda costa.

¹⁸ Es recomendable el sitio www.scilaw.org/team

5. RÉPLICA A OLGA FUENTES SORIANO

La coincidencia con Olga Fuentes Soriano es prácticamente plena. Como ella misma indica, más que refutar mi texto, lo matiza, planteando un aspecto complementario que me parece particularmente interesante. Así, estima que en la afirmación de que la perspectiva de género nada puede aportar en el enjuiciamiento para suplir las deficiencias probatorias del testimonio único no corroborado subyace una visión estática del fenómeno. Intentando ser lo más fiel posible a su argumento creo que el principal reproche a mi tesis sería que se centraría en la prueba como resultado o punto de llegada, donde solo habría un testimonio no corroborado, momento en el que las máximas de la experiencia o generalizaciones provenientes de la perspectiva de género de poco servirían, olvidando que la prueba también puede concebirse de forma dinámica o como actividad, de tal modo que no solo durante la investigación sino también en el enjuiciamiento tales máximas serían de indudable utilidad para incorporar elementos de corroboración empleables en el razonamiento probatorio. Bajo este ángulo complementario, podría eludirse el resultado indeseado gracias a los aportes de la perspectiva de género, que permitirían hallar elementos de confirmación que, en otro caso, pasarían desapercibidos.

En mi trabajo sostengo que en la práctica judicial sucede con relativa frecuencia que, en juicios por delitos de violencia de género, los medios de prueba propuestos para su práctica en el acto de la vista se reducen a los testimonios de la afirmada víctima y del acusado. Del mismo modo, afirmo que ello no siempre ocurre, frente a lo que suele alegarse, ante la ausencia de otros elementos informativos de interés. Por el contrario, lo cierto es que en tales casos suelen existir numerosos elementos indiciarios de gran valor acreditativo que pueden acumularse a la declaración del testigo que afirma ser víctima y que permiten, por su riqueza y a la luz de la perspectiva de género, una adecuada reconstrucción histórica del suceso de forma respetuosa con las exigencias de la presunción de inocencia. Sin embargo, durante la investigación se suele prescindir de incorporar tales elementos, por lo que en el juicio oral no estarán identificadas las fuentes de prueba que podrían incorporar las informaciones con valor corroborador. La perspectiva de género desempeñaría entonces una función esencial, identificando en la fase preprocesal los datos informativos relevantes que pueden obtenerse de la declaración de la víctima, que hacen coherente y congruente su relato, e indican el camino que ha de seguirse para la búsqueda e incorporación a la causa del material externo que lo corrobore. Intentaré desarrollar esta idea.

En el proceso penal hay dos fases claramente diferenciadas. Una, de investigación, cuya finalidad principal es descubrir el hecho con caracteres de delito y la persona posiblemente responsable de él, y otra, de juicio oral, cuyo objetivo es verificar si la hipótesis provisional que resulta de la investigación acerca del hecho y de su autor queda acreditada más allá de toda duda razonable. Por tanto, la primera fase tiene

como meta construir una hipótesis inculpatoria provisional, y la segunda verificar si tal hipótesis queda o no acreditada con arreglo al estándar acreditativo requerido¹⁹

A tal efecto, en la primera fase se identifican distintas fuentes investigativas (v.gr. las personas que presenciaron el hecho o lo sufrieron, los facultativos que atendieron a quienes sufrieron lesiones, los profesionales que evaluaron el coste económico del desperfecto o daño causado, etc), que se incorporan al expediente procesal preliminar convirtiéndose en medios de la investigación (v.gr. declaraciones de testigos, peritos, etc) y aportan elementos investigativos (v.gr. la información relevante que resulta de cada medio de investigación practicado). En la segunda, de modo equivalente, las fuentes probatorias se introducen en el juicio oral como medios de prueba aportando elementos probatorios. Suele ocurrir que las fuentes investigativas resulten coincidentes con las fuentes de prueba en el juicio oral dada la vocación probatoria de aquellas, pero, aun dándose tal coincidencia, existe una clara diferencia entre los medios investigativos y los medios probatorios, pues los primeros, a salvo las excepciones expresamente previstas en las leyes, carecen de valor probatorio en el juicio oral. Tal distinción obedece a razones tanto epistemológicas como político-constitucionales en las que ahora no me detendré.

Valorar significa construir un razonamiento que nos conduzca desde el elemento de prueba al resultado. Para ello es esencial la elección de la inferencia correcta que lleve de un sitio a otro. Tales inferencias sirven tanto para determinar la relevancia del medio de prueba (v.gr. si las huellas dactilares del acusado se encuentran en el lugar de los hechos, eso significa que en algún momento estuvo en esa lugar) como para determinar su fiabilidad (v.gr: si la madre del acusado declara haberlo visto apuñalar a su propio hijo, la atendibilidad del dato resulta de la atendibilidad del testimonio, que resulta, a su vez, de la inferencia de que no se miente al declarar en contra del interés propio)²⁰, como para conectar los elementos probatorios entre sí y con la hipótesis a probar. Entre esas inferencias se encuentran las máximas de la experiencia, lo que nos remite a los estereotipos y a las expectativas generales sobre el comportamiento de los demás, así como a la tendencia a percibir que la conducta de los demás se ajusta a ellas, aspectos sobre los que la perspectiva de género puede aportar informaciones muy valiosas.

¹⁹ Entre una y otra fase, se inserta lo que la doctrina denomina «fase intermedia», que cumple, básicamente, dos funciones: declarar o no suficiente la instrucción y realizar el llamado «juicio de acusación», en el que el órgano judicial examina el fundamento de la acción penal, y acuerda, bien sobreseer la causa, archivándola, bien abrir el juicio oral, permitiendo el pase a la siguiente fase. A tal efecto, en la fase intermedia, el órgano judicial ha de comprobar la razonable probabilidad de la acusación, con la finalidad de decidir si la causa ha de continuar o cerrarse anticipadamente. En palabras de la Circular 1/1989 de la Fiscalía General del Estado «es necesario arbitrar un momento procesal garantizador de que nadie va a ser sometido a un juicio penal sin fundamento. Momento garantizador o filtro de acusaciones infundadas que, de otra parte, existe en todos los sistemas acusatorios».

²⁰ Ejemplos tomados de IGARTUA SALAVERRIA, 2018.

La valoración presenta aspectos similares durante la fase investigativa y la de enjuiciamiento. Con todo, hay algunas diferencias. Por lo que respecta a la investigación, las diferencias sustanciales radican, en primer lugar, en el carácter dinámico de esta fase, que se encuentra en construcción, por lo que la identificación de los elementos investigativos relevantes y fiables solo puede tener lugar al final de dicha fase y con un carácter de alta provisionalidad. En segundo lugar, y como derivación de lo anterior, el razonamiento es esencialmente abductivo y conjetural, pues a medida en que se van incorporando datos investigativos nuevos se introducen nuevas hipótesis explicativas de esos datos o se revisan las hipótesis precedentes. Mientras que en el juicio oral el juez se enfrenta a una hipótesis probatoria que describe los hechos del conflicto y debe decidir si está o no probada para fundar en ella el fallo, el investigador formula una hipótesis inicial, a la vista de los primeros elementos investigativos de los que dispone y, a continuación, va adecuándola, modificándola o sustituyéndola por otras a medida que avanza la indagación sobre la base del material del que va haciendo acopio. No hay, por tanto, y a diferencia de lo que sucede en el juicio oral, hipótesis cerradas a confirmar o refutar. Es claro, en consecuencia, que, ya se llame policía, fiscal o juez instructor, el investigador tiene un papel activo, por lo que sus propios sesgos indudablemente inciden en el contenido de las hipótesis que decide explorar. Por otro lado, es evidente que el estándar acreditativo propio de esta fase es distinto.

No obstante las sustanciales diferencias entre las fases de investigación y enjuiciamiento, la elección de las inferencias tanto para determinar la relevancia del medio investigativo como para establecer su fiabilidad como para conectar los distintos elementos investigativos entre sí y con la hipótesis de que se trate queda sujeta a las mismas reglas rectoras que las que, en los respectivos ámbitos, afectan a la selección de las inferencias en el juicio oral, por lo que desempeñan un importante papel los estereotipos y las máximas de la experiencia. La ventaja, en la perspectiva de la acusación, radica en el hecho de que durante la investigación se puede completar la insuficiencia acreditativa. Sin embargo, llegado el momento del juicio oral, y enfrentados a un testimonio único no corroborado, las inferencias que puedan hacerse para calibrar la fiabilidad del testimonio, por más que se ajusten a la perspectiva de género, no compensarán la falta de suficiencia de la prueba.

Ciertamente, no existiría óbice alguno en proponer para el juicio oral medios de prueba que pudieran servir como elementos de corroboración que no constaran en el expediente investigativo y que aportaran informaciones valiosas gracias a las aportaciones de la perspectiva de género, pero se trataría de una labor de la acusación, en la que el juzgador no podría tener un papel activo. Ahora bien, si nos colocamos en la posición del juzgador, que no puede incorporar prueba de oficio, solo podrá limitarse a valorar la existente. Y si la existente se redujera, ab initio, a la declaración testifical de la afirmada víctima la perspectiva de género no podría suplir la insuficiencia probatoria.

Distinto es el caso, al que se refiere Fuentes Soriano, en el que ab initio hay diversas fuentes de prueba (v.gr. la declaración testifical de la afirmada víctima, documentos que acreditan la dependencia económica de aquella hacia el acusado, u otras testificales que evidencian la ruptura por la víctima de todas sus relaciones sociales y familiares y contactos con terceros). En tales casos, es indudable que la valoración con perspectiva de género permitiría atribuir valor corroborador a aquellas informaciones que reflejaran una situación de asimetría o dominación en el seno de la pareja, posibilitando la confirmación de extremos de la declaración testifical con relevancia probatoria. Del mismo modo, parece claro que una valoración sin tal perspectiva podría concluir que solo hay un medio de prueba disponible (la declaración de la víctima), en la medida en que los demás no aportarían información de interés. No hay, en consecuencia, divergencias entre su planteamiento y el mío, ya que yo me refería al caso en el que el medio de prueba efectivamente practicado hubiera consistido en la sola declaración de la víctima.

6. ÚLTIMAS OBSERVACIONES

No querría concluir sin reiterar la idea de que en el ámbito de la valoración probatoria plenaria de lo que se trata es de tomar conciencia de la presencia de los estereotipos y de su potencialidad distorsionadora en el proceso de cognición y de interpretación, así como destacar que ello no equivale, sin más, a determinar el sentido de la resolución ni a descartar soluciones probatorias que puedan coincidir casualmente con el estereotipo. Algunos ejemplos pueden aclarar esta idea.

Suele afirmarse que la actitud victimizada de quien dice haber sufrido un hecho delictivo es un signo de fiabilidad de su testimonio y dato corroborador de este. Así, la mujer que sale del despacho en el que trabaja, agitada, nerviosa y llorando al tiempo que afirma que un compañero la ha agredido sexualmente proporcionaría un elemento probatorio atendible para la hipótesis acusatoria (la existencia y autoría de la agresión). Sin embargo, si el resto de elementos probatorios, provenientes de otros medios de prueba igualmente atendibles fueran incompatibles con aquel, no podría estimarse justificada la tesis de la acusación, pese a que hubiéramos partido de la fiabilidad inicial del testimonio. La fuente y el medio probatorio serían atendibles, pero el elemento de prueba que suministran quedaría neutralizado.

En sentido inverso, la perspectiva de género nos enseña que el hecho de que esa misma víctima esperase a la finalización de la jornada laboral para abandonar el lugar de trabajo, y que, tras una reflexión de varios días, decidiera denunciar los hechos, no merma la fiabilidad del testimonio, pues la persona pudo sufrir la reacción de estrés agudo, también conocida como bloqueo emocional, que puede surgir en respuesta a un evento traumático, y que se caracteriza por la falta de respuesta adecuada al estímulo, lo que explicaría que no reaccionara inmediatamente. Ahora bien, si el resto de elementos probatorios provenientes de otros medios, relevantes, atendibles y disponibles, resultara compatible tanto con la información proporcionada por ella como

con la información proporcionada por el acusado (v.gr. la relación fue consentida), la ausencia de reacción inmediata de la afirmada víctima podría ser valorada como sintomática de la situación de bloqueo emocional, pero no permitiría descartar como improbable o implausible la hipótesis enfrentada del consentimiento, lo que impediría dar por justificada la tesis de la acusación. Pero si, yendo un paso más, el resto de elementos probatorios, relevantes, atendibles y numerosos, fuese radicalmente incompatible con el relato de aquella, cabría incluso afirmar que quien dice ser víctima mintió, lo que no significa que en la valoración se hubiera partido del estereotipo machista de la mujer fabuladora o mentirosa. Aquí radica la diferencia entre la aproximación sociológica (o nivel macro) y la procesal (nivel micro), y muchas fuentes de malentendidos: en el nivel procesal no se afirma que las mujeres mientan, sino que en el caso concreto una mujer particular ha podido mentir.

Otra hipótesis: maltrato a la pareja. La perspectiva feminista pone de relieve que la violencia de género es transversal y que, por tanto, una mujer independiente económicamente y con buena posición social se encuentra igualmente expuesta a ella. Su falta de ajuste al estereotipo de víctima económicamente dependiente y de extracción social inferior no puede, por ello, ser tomada en consideración a priori para descartar la atendibilidad inicial de los elementos probatorios resultantes de su declaración. Sin embargo, si el resto de elementos probatorios operase en contra del dato, afirmado por la víctima, de que sufrió maltrato, este elemento o dato informativo quedaría neutralizado, lo que no significa que el testimonio fuera desatendido por su falta de concordancia con el «perfil de mujer maltratada».

En definitiva, en el ámbito del juicio oral la perspectiva de género juega un papel especialmente valioso para no descartar la atendibilidad de la declaración de la afirmada víctima. Ahora bien, no sirve para afirmar, por sí sola, el superior peso probatorio de la narración individualmente considerada, esto es, de los elementos probatorios que suministra. Consecuentemente, también sirve para evaluar la atendibilidad del resto de medios de prueba y, en especial, la declaración de la persona acusada, pero no para calibrar, de forma individual, su fuerza probatoria. Por último, tiene otra utilidad relevante en el momento de la valoración conjunta: permite poner en conexión los elementos probatorios entre sí y determinar su concordancia con la hipótesis a probar pues, a la luz de dicha perspectiva, un conjunto de elementos probatorios en apariencia dispersos puede adquirir cierto sentido, que en otro caso podría haber pasado desapercibido.

7. BIBLIOGRAFÍA

- ANDRÉS IBÁÑEZ, P., 2019: «Principio de presunción de inocencia y principio de victimización: una convivencia imposible». *Revista do Ministério Público*, 160. Lisboa.
- ARDUINO, I., 2017: «Mecanismos de simplificación alternativos al juicio y género en el proceso penal: redefinir la discusión desde la política criminal». En *Género y Justicia Penal*. Buenos Aires: Didot.

- CAMARENA GRAU, S., 2017a: *La Directiva 2016/343. Los retos de la trasposición. Como debe y puede protegerse la inocencia en el proceso penal*. Consejo General del Poder Judicial. Cuadernos digitales de formación n. 8.
- , 2017b: *Prueba testifical y presunción de inocencia. Breve referencia a técnicas de interrogatorio*. Consejo General del Poder Judicial. Cuadernos digitales de formación n.º 55.
- DE PAULA RAMOS, V., 2019: *La prueba testifical. Del subjetivismo al objetivismo, del aislamiento científico al diálogo con la psicología y la epistemología*. Barcelona: Marcial Pons.
- ERICE MARTÍNEZ, E., 2018. «Perspectiva de género y derecho penal», en *Boletín de la Comisión Penal de Juezas y Jueces para la Democracia*. N. 10, volumen I.
- FERRER BELTRÁN, J., 2007: *La valoración racional de la prueba*, Barcelona: Marcial Pons.
- GASCÓN, M., 2010: *Los hechos en el derecho*, Barcelona: Marcial Pons.
- IGARTUA SALAVERRIA, J., 2018: «Modos de valoración de las pruebas y tipos de razonamiento», en *Cuestiones sobre prueba penal y argumentación judicial*. Santiago de Chile: Ediciones Olejnik.
- LÓPEZ ORTEGA, J. J., 2000: «Los principios constitucionales del proceso penal», *Derecho procesal penal salvadoreño*, (Casado Pérez J. M. dir.) Edición de la Corte Suprema de Justicia de El Salvador y de la Agencia Española de Cooperación Internacional, San Salvador.
- MANZANERO, A., 2010: *Memoria de testigos. Obtención y valoración de la prueba testifical*. Madrid: Pirámide.
- RAMÍREZ ORTIZ, J. L., 2019: *Perspectiva de género, prueba y proceso penal: una reflexión crítica*. Valencia: Tirant lo Blanch.
- SAN MIGUEL BERGARECHE, M. N., 2018: «Juzgar y castigar ¿con perspectiva de género?». En *Boletín de la Comisión Penal de Juezas y Jueces para la Democracia*. N. 10, volumen II.
- VIVAS LARRUY, M.A., 2016: «Enjuiciar con perspectiva de género. Perspectiva de género en Derecho penal», en *Impartición de justicia con perspectiva de género en los distintos órdenes jurisdiccionales*. Consejo General del Poder Judicial. Cuadernos Digitales de Formación N.º volumen: 48.
- VÍVES ANTÓN, T., 2011: *Fundamentos del sistema penal*. Valencia: Tirant lo Blanch.
- WATKINS, S., 2018: «¿Qué feminismos?», En *New Left Review*. Marzo-Abril 2018.

CIENCIA PARA
EL PROGRESO

LOS CONOCIMIENTOS DE LA PSICOLOGÍA MÁS ALLÁ DE LA PRUEBA PERICIAL*

Carmen Herrero

Universidad de Salamanca

cherrero@usal.es

RESUMEN: El presente trabajo plantea la necesidad de que el Derecho y los tribunales utilicen el conocimiento científico procedente de la psicología. Más concretamente, sugiere el papel de distintos procesos y fenómenos psicológicos y sociales en la obtención, interpretación y valoración de la prueba. Para ello, aborda la problemática del testimonio del experto y la admisibilidad de la prueba aportando investigaciones empíricas que cuestionan la distintividad del razonamiento legal. Se constata la diferente consideración o estatus otorgado a la psicología frente a otras disciplinas forenses, se muestra la utilización por parte de los operadores jurídicos de creencias de sentido común no validadas científicamente y se ilustran sus efectos negativos en distintos ámbitos y en ciertas decisiones de la Sala Segunda de lo Penal del Tribunal Supremo Español. En particular, las referidas a los criterios orientativos para valorar la credibilidad de la declaración de las víctimas.

Además, se proponen distintas áreas en las que el Derecho puede beneficiarse de los conocimientos de la psicología jurídica y se sugiere una aproximación más amplia al rol del perito. Por otra parte, se aplicará el conocimiento de la psicología, especialmente de la psicología cognitiva, para explicar los errores en las pruebas forenses. Se analizan también las implicaciones para la prueba de los hechos y para la calidad de la toma de decisiones. Finalmente, se aboga por una mayor interacción entre académicos/profesionales de la psicología y el derecho, así como por la necesidad de estudios empíricos con los profesionales del sistema de justicia.

PALABRAS CLAVE: Psicología Jurídica; testimonio del experto; sentido común; prueba; sesgos en las ciencias forenses.

* La autora agradece a los revisores que han contribuido a mejorar la calidad de este trabajo.

CONTRIBUTIONS OF PSYCHOLOGY BEYOND THE EXPERT EVIDENCE

ABSTRACT: This paper raises the need for the law and the courts to use scientific knowledge from psychology. More specifically, it suggests the role of different psychological and social processes in obtaining, interpreting and evaluating the evidence. To do this, it addresses the problem of the expert's testimony and the admissibility of the evidence by providing empirical research that dispute the distinctiveness of legal reasoning. The different consideration or status granted to psychology compared to other forensic disciplines is verified, as well the use by legal operators of not scientifically validated common-sense beliefs and its adverse effects, as it occurred in some decisions of the Spanish Supreme Court. In particular, those referring to guiding criteria to assess the credibility of the victims.

Additionally, different areas in which the law can benefit from the knowledge of legal psychology are proposed, and a broader approach to the role of the expert is recommended. On the other hand, knowledge of psychology, especially cognitive psychology, will be applied to explain errors in forensic tests. The implications for the finding of facts and the quality of decision-making are also analysed. Finally, it is advocated for more significant interaction between academics/professionals of psychology and law, as well as the need for empirical studies based on the professionals of the justice system as participants.

KEYWORDS: legal psychology; expert testimony; common-sense; evidence; bias in forensic science .

SUMARIO: 1. INTRODUCCIÓN. ALGUNOS CASOS Y PREMISAS BÁSICAS.— 2. SOBRE REGLAS Y CRITERIOS DE ADMISIBILIDAD DE LA PRUEBA (CIENTÍFICA?).— 3. CIENCIA, PSICOLOGÍA VS SENTIDO COMÚN: 3.1. Sobre la diferente consideración de la ciencia psicológica: alguna constatación. 3. 2. Sobre la diferente consideración: el sentido común como explicación.— 4. CIENCIA PSICOLÓGICA E INTERVENCIÓN DEL EXPERTO: 4. 1 ¿Es necesario el testimonio del experto en psicología, en psicología jurídica? 4.2. contribuciones e implicaciones del conocimiento experto en psicología jurídica.— 5. CIENCIA PSICOLÓGICA VS CIENCIAS FORENSES: 5. 1. Sobre el cambio de paradigma en las ciencias forenses: la contribución de la psicología.— 6. CONCLUSIONES.— 7. REFERENCIAS.

RECOMMENDED CITATION: HERRERO, CARMEN: «Los conocimientos de la psicología más allá de la prueba pericial», in *Quaestio facti*, 2: 363-408. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22533

1. INTRODUCCIÓN. ALGUNOS CASOS COMO PUNTO DE PARTIDA Y PREMISAS BÁSICAS

Caso 1: Choque frontal de dos trenes cerca de los Ángeles. Uno de los accidentes de mayor gravedad en el que murieron 25 personas en septiembre de 2008. Con demandas millonarias en juego, las autoridades iniciaron una extensa investigación una de cuyas claves era determinar si el conductor, que había fallecido, se había saltado un semáforo en rojo por ir distraído enviando mensajes de texto con su teléfono. Si

era culpable, lo ocurrido tendría importantes consecuencias para la empresa ferroviaria para la que trabajaba. Si el problema había estado en un mal funcionamiento de las señales, otras empresas diferentes tendrían problemas. Cuatro testigos afirmaron con toda seguridad que el conductor pasó la señal en verde, no se saltó el semáforo. La investigación realizada, sin embargo, mostró lo contrario. ¿Cómo es posible que cuatro personas estuvieran tan equivocadas en un detalle tan importante? (FRENDA, NICHOLS Y LOFTUS, 2011).

Caso 2: Franklin, 51 años. En 1990 condenado por asesinar a una compañera de juegos de su hija 20 años antes. La principal prueba: la denuncia de su hija, Eileen, que tenía 8 años cuando asesinaron a su amiga. Eileen supuestamente había mantenido «reprimido» el recuerdo del asesinato durante 20 años, hasta que en un momento dado comenzó a tener recuerdos detallados y expuestos con un elevado grado de confianza (LOFTUS, 1993).

Caso 3: Tammy, Canadá. 9 de junio de 2011. Retiran los cargos a una mujer tras haber pasado 14 años en la cárcel acusada erróneamente de matar a su hijo de 2 años, acusación que siempre negó. Prueba clave: el testimonio del experto patólogo que determinó que la asfixia o el estrangulamiento eran la causa del fallecimiento. Sus informes fueron calificados posteriormente de “ilógicos y en absoluto científicos”, descubriéndose además otros dictámenes erróneos en su historial como patólogo en casos de muerte infantil. La defensa, por su parte, no presentó otros testimonios que contradijeran la causa de la muerte, aunque argumentó que podía haberse producido por un ataque epiléptico (ROECH, 2012).

Caso 4: Sally Clark. Condenada por el asesinato de sus dos hijos pequeños en 1999, fue liberada en 2003. Investigación adicional reveló que la probabilidad de muerte súbita había sido calculada erróneamente. Múltiples casos documentados en los que los veredictos se han visto influidos por un razonamiento científico-estadístico incorrecto en los tribunales (FENTON, 2011).

Caso 5: Richardson Canadá: Liberado en 2013. Condenado a 30 años por matar a una chica de 19 años, golpeada y estrangulada. Pasó 20 años en prisión. Prueba fundamental: las marcas de mordeduras en víctima (coincidentes supuestamente con dentadura imputado). En el juicio el novio de la chica declaró por primera vez haber oído a Richardson amenazar de muerte a su novia. Pero admitió haber relacionado a Richardson con el caso solo después de que la policía le informara de que las mordeduras coincidían con su dentadura (innocenceproject.org).

Caso 6: Donald Thompson, psicólogo australiano, experto en temas de testimonio. Detenido y acusado de violación por la identificación inequívoca de la víctima. Fue liberado después de ser sometido a una rueda de reconocimiento. Coartada perfecta: cuando se produjo la violación estaba en un debate de TV, con jefe de policía y otros contertulios tratando, precisamente, el tema del testimonio. ¿Qué había hecho que la víctima reconociera a Thompson pero que le atribuyera conductas que no había realizado? (SCHACTER, 2001).

Caso 7: Abrrazak M. en prisión desde 1991 acusado de varios delitos, entre ellos algunas violaciones. Seis años más tarde el ADN demostró claramente que al menos en 3 de las violaciones no era el culpable. Lo era Antonio G., ya condenado por otras violaciones. No fue liberado, murió en la cárcel en el 2000, sin dejar de reclamar que era inocente (SOBRAL, 2008).

Caso 8: M. P, habitual en el «menudeo» de la droga y de las comisarías. Identificado como el autor de un atraco a mano armada a una joyería. Testigo: describió al autor como «pelirrojo, de estatura mediana y piel clara». Ese testigo reconoció al autor «*sin ninguna duda*» en rueda. ¿caso resuelto? ¿Cómo se compuso la rueda? Todos los distractores eran morenos menos uno que era calvo. Solo había un pelirrojo. M. P. no era autor del atraco que se le imputaba (SOBRAL, 2008).

¿Qué ponen de manifiesto estos y otros casos, también los aparecidos en nuestro país?, ¿qué tienen en común? Entre otras cosas, el papel fundamental y, más aún, decisivo, de distintos elementos de juicio, fuentes o medios de prueba; información y conocimiento sustanciales que se aportan al proceso sobre el quién, el qué, dónde, cuándo, cómo, o el por qué y que pretende, por lo tanto, contribuir a la «reconstrucción» de unos hechos y con ella, a la de la verdad sobre los mismos. Pero ¿de qué verdad se está hablando? ¿Es posible seguir utilizando la manida verdad judicial o procesal como refieren otros?

Una imprescindible aproximación racional a la prueba (FERRER 2005, 2007; GASCÓN, 2004, 2010, NIEVA, 2010; TARUFFO, 2005a, 2005b; 2010; VÁZQUEZ, 2013, 2015b), alejada de una concepción como mero acto de persuasión (FERRER, 2017) y convencimiento en el contexto de su libre valoración, requiere abandonar, al menos como posicionamiento de partida, la manida alusión a la verdad procesal o judicial —distinguiéndola de la material o más ajustada a lo real— para referirnos, parafraseando a TARUFFO (2010), a «simplemente la verdad», pero dentro de un marco de garantías fundamentales (RAMÍREZ ORTIZ, 2014).

Los tribunales y, más exactamente, los procesos en distintos órdenes jurisdiccionales no tienen simplemente el objetivo de resolver una controversia o conflicto entre partes y diferentes pretensiones, y hacerlo de acuerdo con unas normas procesales, sino el nada fácil encargo de la averiguación de la verdad. Esto es, el proceso tiene que aspirar a una reconstrucción verídica de los hechos de la causa (TARUFFO, 2005 a, 2005b), a la determinación de una verdad fáctica de calidad (ANDRÉS IBÁÑEZ, 2020) que será la base en la que se asienten, entre otras, la legalidad de las decisiones judiciales y su calidad. Solo así estaremos más próximos a la consecución de la Justicia y al fortalecimiento de la legitimidad de sus instituciones (HERRERO, 2020).

Los casos referidos más arriba muestran, por otra parte, que la obtención de conocimiento válido falló en algún momento, bien de la investigación, o en la admisibilidad de las fuentes de prueba en el proceso, o en la práctica y valoración de estas, o en todos ellos, comprometiendo seriamente la decisión con la que se cerró el procedimiento o el proceso correspondiente. El análisis de lo sucedido, al menos en algunos

de ellos, permite especular que el problema estuvo mucho más relacionado con los hechos, su determinación, interpretación o validez, que con «las reglas de juego» utilizadas para alcanzar la decisión última, o con el variado conjunto de códigos, normativas o articulado que se les aplicaron, aunque sin la connivencia, e implicación activa, de estos elementos hubiera sido imposible concebir tales fatídicos resultados. Es más, no son pocos los que consideran que muchas de las reglas y procedimientos asociados con los procesos legales, junto con otras metas diferentes a la búsqueda de la verdad, así como ciertas carencias en el conocimiento y en los procesos de pensamiento de varios actores jurídicos, se convierten realmente en obstáculos e impiden la consecución de la verdad en el proceso (KOEHLER Y MEIXNER, 2015; SCHAUER Y SPELLMAN, 2013).

Por lo expuesto, es esencial insistir en la calidad de la información, del conocimiento que se obtenga de todos los posibles elementos de prueba por medios lícitos, de la interpretación, utilización y valoración que se haga del mismo. En definitiva, la calidad de todas las actividades y fases relacionadas con la determinación y prueba de los hechos. Afrontar tales requerimientos supone, por un lado, ir más allá de las cuestiones de Derecho, aunque los juristas parecen haberse centrado más en la problemática de la interpretación de las normas y han prestado menos atención a la interpretación y al juicio sobre los hechos (GASCÓN, 2004). Por otro, implica, aunque no exclusivamente, contar con distintos profesionales y académicos no juristas, expertos en materias variadas, a los que se solicita sus conocimientos científicos, técnicos, o de otro tipo, y su testimonio para ayudar, entre otras cosas, a tomar decisiones al juzgador, al tribunal, al jurado en su caso. También de la calidad, fiabilidad y validez de tales contribuciones, de la adecuada interpretación que se haga de ellas por parte de los actores jurídicos, así como de la valoración que de todo ello realice el juzgador, dependerán en gran medida las decisiones que se alcancen en los tribunales. Por lo tanto, el testimonio de los expertos se convierte en uno de los instrumentos fundamentales en el proceso, concretado de diversas formas en los distintos sistemas jurídicos (VÁZQUEZ, 2015a, 2015b, 2016; GASCÓN, 2013) y puede ir más allá de la realización de un peritaje o la emisión de un dictamen.

Desatender más de lo deseable las cuestiones de hecho, «el maltratamiento» de las mismas como refiere ANDRÉS IBÁÑEZ (2017) respecto a la formación de los jueces, el circunscribir, cuando no reducir, la prueba a un mero suceder de normas procesales puede tener consecuencias e implicaciones desafortunadas. Dejando al margen las de carácter jurídico legal, quisiera señalar algunas: a) ausencia de la psicología/psicología jurídica; b) la socialización de los juristas y otros profesionales; c) menor aprecio por la investigación empírica. La primera, ha servido como excusa implícita para dejar fuera de los asuntos jurídico-legales en general, y de los que tienen en mayor medida que ver con el proceso en particular, los conocimientos y avances científicos de ciertas disciplinas como la psicología o, una de sus denominadas aplicaciones, la psicología jurídica que, supuestamente, poco tendrían que aportar a la aplicación de unas normas (a unos hechos con significación y relevancia legales). Los avan-

ces de otras disciplinas también habrían quedado excluidos, salvo los considerados más imprescindibles de las denominadas «ciencias duras», siempre muy presentes y apreciadas en el contexto forense (ej. ciencias forenses). Pero ¿no contribuirían todas ellas, aun en diferente grado si se prefiere, a la determinación de unos hechos, a la averiguación de la verdad? Si los hechos quedan al margen, las disciplinas que los explican y permiten entenderlos también, y en particular si tienen que ver con el comportamiento humano, para cuya interpretación cualquier ciudadano puede ser «psicólogo ingenuo» o lego.

Sin pretender caer en un burdo «psicologismo», podríamos afirmar que la psicología¹ permea el derecho. Los legisladores y operadores jurídicos actúan como psicólogos aplicados ingenuos, que manejan continuamente suposiciones y conjeturas

¹ La delimitación o definición de los distintos ámbitos de la Psicología que se mencionan excede lógicamente de los objetivos del presente trabajo. No obstante, puede ser de utilidad alguna breve aclaración sobre ellos, si bien efectuada a grandes rasgos y sin los matices que probablemente fueran necesarios. Me referiré a la *psicología*, en general, cuando sean múltiples y muy variados los temas y perspectivas implicados sobre la mente y el comportamiento humanos, pero todos ellos abordados científicamente. La referencia a la *psicología cognitiva*, como otros ámbitos disciplinares, supone una perspectiva y no únicamente unas temáticas concretas, nos sitúa ante el estudio científico de los procesos mentales por los que las personas perciben, almacenan, procesan, recuperan y utilizan la información.

La *psicología social*, y su perspectiva psicosocial, pone el acento en la interacción recíproca y el estudio de las relaciones entre los procesos psicológicos individuales y los sociales. La literatura se refiere a ella como el estudio científico de cómo los pensamientos, sentimientos y conductas de las personas están influidas por la presencia real, imaginada o implícita de otros, del contexto y de las situaciones en las que se encuentran los sujetos. Por ello se habla también de *cognición social* para poner el acento en el mismo tipo de procesos referidos previamente pero aplicados fundamentalmente al mundo social. Esto es, cómo pensamos sobre nosotros mismos y los demás y los comportamientos, así como sobre las relaciones sociales, y cómo damos sentido y construimos todo ello. De hecho mucha de la literatura psicosocial diría que toda cognición es necesariamente social, que nuestra mente está socialmente estructurada. Mientras que la *psicología social* se centra en mayor medida en las semejanzas entre los sujetos fruto, entre otras cosas, del contexto, de las situaciones sociales y cómo se construyen, la *psicología de la personalidad* presta más atención al estudio de las peculiaridades y diferencias individuales entre las personas.

Por su parte, la *psicología jurídica* se considera un campo aplicado que estudia todas las temáticas y procesos relevantes en la interacción entre la Psicología y el Derecho, incluidos todos los que afectan a los actores implicados. El campo se nutre principalmente de los conocimientos de la *psicología básica*, de la *psicología cognitiva* o de la *psicología social*, pero también de la psicología del desarrollo, entre otras. Cuando el ámbito de aplicación se concreta aún más y se centra en el estudio, por ejemplo, de la exactitud y credibilidad de las declaraciones, así como de todo lo que ello implica, se habla de *psicología del testimonio*. Por otra parte, cuando los profesionales de la psicología auxilian al sistema de justicia, generalmente como peritos, podríamos hablar de *psicología forense*.

Todos los ámbitos disciplinares mencionados en el presente trabajo tienen un aspecto común que debe subrayarse. Obtienen sus conocimientos a través de la investigación utilizando distintas herramientas del método científico y requieren la realización de trabajo empírico para poner a prueba las hipótesis de las que se trate. Dependiendo de los condicionantes del objeto de estudio concreto y de contenidos implicados, se realizan experimentos, que permiten un mayor control sobre las variables y establecer relaciones de causalidad, cuasi experimentos o estudios correlacionales, e incluso cualitativos, y no solo en «el laboratorio» sino en contextos reales, con mayor dificultad de control, pero no por ello menos relevantes o necesariamente menos fiables. Finalmente, todos los ámbitos mencionados en el trabajo aportan conocimientos significativos y relevantes en el contexto jurídico.

sobre cómo las personas piensan y actúan. El sistema legal, incluso, sería considerado una institución de cognición social (SPELLMAN Y SCHAUER, 2012a, 2012b; SPELLMAN Y WEAVER, 2020). Su relevancia no se agota con los asuntos del testimonio, los testigos y su credibilidad, por muy importante que esta área haya sido y siga siéndolo. De hecho, probablemente constituya el ámbito en el que se han producido varias de las contribuciones más consolidadas (GONZÁLEZ Y MANZANERO, 2018; MAZZONI, 2019; NATIONAL RESEARCH COUNCIL, 2014; WELLS., KOVERA, DOUGLASS, BREWER, MEISSNER, y WIXTED, 2020), alguna de las cuales ha llegado, o empieza a hacerlo cada vez más, también a los juristas en nuestro país. Y no se agota, además, porque hay otros variados procesos psicológicos y psicosociales implicados en el contexto jurídico legal que son significativos, incluso, más allá de los tribunales, no se sitúan únicamente en la jurisdicción penal, al tiempo que la aportación de los profesionales y académicos no se reduciría necesariamente a la realización de algunos peritajes (BRANK, 2019; ROBBENNOLT Y HANS, 2016; ROWELL Y BILZ, 202; SAKS Y SPELLMAN, 2016; STERN, S. M. Y LEWINSOHN-ZAMIR, 2020; STERNLIGHT Y ROBBENNOLT, J. K., 2008). Reparar en todos ellos no solo amplía la perspectiva en relación con la obtención e interpretación de los hechos en el proceso, sino que la sitúa más allá del mismo, al tiempo que refleja la necesidad de una mayor interacción entre juristas y otros ámbitos del saber. No hacerlo contribuye a la separación de mundos que deberían ir de la mano.

La segunda de las implicaciones de relegar el asunto de los hechos referido más arriba tendría que ver con su potencial efecto en lo que podríamos denominar la socialización, y con ella en la práctica, de los futuros juristas y profesionales del derecho, así como en los de otras disciplinas relacionadas o con relevancia en el sistema de justicia. Y de hecho, ayuda a explicar distintas creencias mantenidas por ellos. Como ilustración, al menos parcial, de esta afirmación puede acudir a las falsas creencias, mitos e ideas pseudocientíficas ampliamente compartidas por el alumnado del último curso de derecho (HERRERO, 2016; HERRERO, MASIP Y PICADO, 2016) y que venimos recogiendo los últimos años a partir de un proyecto de innovación docente. Solo me referiré en este momento a tres de ellas: un alto porcentaje (70-92 %, dependiendo de la afirmación de que se trate) consideran que los jueces están más capacitados, y disponen de mayores habilidades para pensar y razonar, que

No se puede afirmar que uno de ellos como «un todo» es más fiable que otro, dependerá del grado, medida y adecuación con la que se utilicen las distintas herramientas del método. De este modo, las afirmaciones y propuestas científicas deben basarse no solo en buenos métodos sino en la obtención de evidencia suficiente y fiable. Como señala REISBERG (2014), por muy cuidadosos que sean los investigadores en la utilización de procedimientos adecuados, una afirmación, propuesta, teoría o explicación de un fenómeno no puede tenerse en cuenta si la evidencia es escasa o débil. En este sentido, el que un ámbito disciplinar sea ampliamente aceptado y considerado fiable por la utilización del método científico o la solidez de sus teorías no significa que cualquier propuesta, afirmación o conocimiento concreto que se deriva de él tenga que ser igualmente aceptado y fiable. En este sentido, hay afirmaciones sobre la mente y comportamiento humanos, de relevancia también en el contexto jurídico, que han obtenido más sólida y mejor evidencia que otras, lo que habrá de analizarse y considerarse caso a caso.

otras personas sin formación jurídica, a la hora de interpretar una prueba que un perito/experto² presenta en el proceso. Estiman que la práctica y experiencia profesional de jueces, abogados o fiscales son determinantes para conocer cuándo un testigo o un imputado está diciendo la verdad; o consideran que los jueces toman mejores decisiones que los jurados. ¿En qué desempeño de rol respecto a los jueces, por ejemplo, están pensando exactamente? ¿Tal vez, y en mayor medida, en la interpretación jurídica? Los resultados, que no permiten por ahora resolver esta cuestión y habrán de colocarnos en el terreno de la mera hipótesis, son muy similares en muestras de alumnado de criminología, en menor medida de los de psicología, pero con igual o más fuerza en una muestra de profesionales de la policía (con años de experiencia y sin ella) y en algún pequeño grupo de operadores jurídicos (jueces, abogados, fiscales, litigantes). Interesante, además, es la seguridad que manifiestan en lo acertado de sus creencias.

La tercera, y última de las consecuencias, alude al poco aprecio que suele tenerse por los estudios empíricos y su utilidad en el contexto jurídico (EDMOND y HAMER, 2012; SCHAUER, 2006; SPELLMAN, 2007; SCHAUER, 2010), en especial si provienen del ámbito no nacional y no se circunscriben a nuestro sistema jurídico concreto. Y no deja de ser un tanto desconcertante dado que, aunque no en exclusiva, la acumulación de conocimiento a través de la investigación empírica es lo que permite poner a prueba aquello que en el derecho solo son hipótesis pendientes de corroboración. Este último aspecto sobre los estudios empíricos nos conduce a la siguiente idea que quisiera formular y permite retomar, además, el problema del control de la calidad de la información, de los elementos de prueba que entran en juego en el proceso, generalmente a través del testimonio de los expertos y peritos.

La utilización de la ciencia en el proceso como medio (método) de averiguación de la verdad, y no solo respecto a cuestiones de admisibilidad, es un asunto controvertido (GASCÓN, 2004, 2013; FAIGMAN, 2006; FAIGMAN et al., 2016, TARUFO, 2005a, 2013; SCHAUER, 2010b; VÁZQUEZ, 2015a) en el que este trabajo no tiene intención de entrar. Pero si quisiera comentar alguno de los aspectos que le circundan. Acudir a los conocimientos científicos no es sinónimo de hallar certezas absolutas, pero sí algunas informaciones relevantes, sobre los hechos o los enunciados de los hechos. Es inherente a la empresa científica moverse en el ámbito probabilístico (BROOMELL y KANE, 2017), lo que supone manejar la incertidumbre. Y esto no suele admitirse de buen grado ni por el público en general, ni por los profesionales de múltiples ámbitos en particular. El razonamiento científico no es sencillo (EDMOND y HAMER, 2012;

² Es necesario aclarar que el hecho de que en ocasiones en el texto se utilicen simultáneamente los términos de perito y experto no significa que se les esté otorgando el mismo significado. Los peritos son expertos o especialistas pero los expertos no necesariamente son peritos. Un experto se convierte en perito, ya sea de parte o de confianza del juez, cuando este lo designa como tal. Por ello, dependiendo del contexto en el que se estén utilizando será más adecuado la conceptualización del experto como perito, y se utilizará preferiblemente ese término, mientras que en otras ocasiones aludiremos al experto más allá de su rol pericial. De hecho, en el trabajo se propone precisamente no limitar la labor y actuación del experto a la realización de un peritaje.

SHAH et al., 2017), como no lo es adaptar y aplicar los fenómenos generales que trata de comprender la ciencia a los casos particulares (FAIGMAN, MONAHAN y SLOBOGIN, 2014; FISHER et al., 2015). Y este último es, en particular, uno de los problemas sustanciales que continuamente habremos de sortear en las relaciones entre la psicología y el derecho. De este modo, la acumulación de conocimiento a través de estudios empíricos permite alejarnos del «todo vale», de que cualquier hipótesis, teoría o resultado sea igualmente válido. Y cuando no lo permite, como plantea REISBERG (2014), tal vez es mejor no decir nada. Tan sencillo como es esto desde la posición de académicos e investigadores, es un imposible en el proceso. Hay que alcanzar decisiones y hacerlo en un tiempo limitado.

Abogar por la intervención de la ciencia en la prueba tampoco es sinónimo de cientifismo (HAACK, 2012, cit. VÁZQUEZ, 2015a), ni de objetivismo ingenuo, aunque el «envoltorio» pudiera parecerlo. Simplemente, o no tanto, se trata de evitar que se tomen decisiones utilizando la pseudociencia, admitiendo resultados e interpretaciones que no son válidas científicamente, incluso aunque vengan de disciplinas consideradas como tales. Esto implica que el conocimiento del experto, en sí mismo, y en particular cuando se concreta en la función de perito, no es necesariamente fiable ni válido. Habrá de determinar cómo se obtuvo y qué permite afirmar. Y para ello, probablemente no sea suficiente acudir a sus credenciales.

Finalmente, no se está proponiendo que la investigación científica sustituya a la jurídica, ni que el testimonio del perito, o del experto en general, lo haga respecto del juzgador o, por ejemplo, del legislador. Tampoco que las probabilidades de error aceptables para algún conocimiento científico determinado sean las que deban asumirse en el contexto jurídico, aunque debieran tenerse en cuenta. Se trata, al menos, de que las «máximas de la lógica y la experiencia» no sean absolutamente incompatibles con las de la ciencia. Porque si no es así ¿qué nos queda?

Estamos abocados al «sentido común», a la «cultura promedio» del «hombre razonable» ... al «acervo de conocimientos» personal que cualquier miembro de la sociedad tiene en su mente y su memoria (TARUFFO, 2013: 205).

La mayor contribución que pueden hacer los psicólogos al sistema jurídico no está en poner este sistema «bajo el microscopio», sino en hacer rápidamente disponible y accesible el conocimiento alcanzado sobre los procesos cognitivos humanos, sobre la naturaleza humana, de manera que, finalmente, sea indudable para todos como «sentido común» (KING, 1984: 79, cit. en GARRIDO Y HERRERO, 2006)

Con las premisas expuestas y el espíritu expresado por King en mente, el presente trabajo tiene los siguientes objetivos. Demostrar la necesidad de que el Derecho, el proceso y los operadores jurídicos acudan al conocimiento científico obtenido por la psicología, en particular a los de la psicología jurídica. Acudiendo a una variedad de estudios empíricos, pretende constatar la utilización de la mera intuición, de creencias de sentido común en el ámbito jurídico e identificar algunos de los efectos perniciosos que provoca alejarse del conocimiento acumulado a través de la investigación. Se pondrá de manifiesto, igualmente, la intervención de supuestos y procesos psicológicos o psicosociales en variados asuntos jurídicos, incluidos varios que atañen a

aspectos cruciales de la prueba. Al hacerlo, se sugiere la pertinencia de que el conocimiento experto «psicojurídico» intervenga y se tenga en cuenta más allá de la prueba pericial, lo que implica que el experto en estas materias no limite sus aportaciones y actuaciones al rol de perito en el proceso. Está más allá del alcance del presente trabajo y de las competencias de la psicología jurídica establecer y especificar las variadas formas en las que se pueda concretar el papel del experto en nuestro sistema jurídico o si las normas procesales que rigen en él deban revisarse e incluso modificarse, o cómo hacerse si fuera el caso. Sí es competencia, sin embargo, advertir de los efectos de no tener en cuenta sus aportaciones u ofrecer el referido conocimiento experto para mejorar la calidad y el funcionamiento de nuestro sistema legal. De modo que en las páginas que siguen, después de una sucinta referencia a las Reglas Federales de Prueba, abordaremos ciertas problemáticas con la admisibilidad de la prueba o con el testimonio del experto y en las que están implicados procesos psicológicos o psicosociales que cuestionan la distintividad del llamado razonamiento legal. Posteriormente constataremos la diferente consideración o estatus otorgado a la psicología frente a otras disciplinas forenses, mostrando la utilización de creencias de sentido común no validadas científicamente por parte, entre otros, de los operadores jurídicos. Se ilustrarán sus efectos negativos revisando críticamente ciertas sentencias de la Sala 2ª del Tribunal Supremo. En concreto, las referidas a criterios orientativos para valorar la credibilidad de las víctimas. Se ponen sobre la mesa, además, distintas áreas o problemas en las que el Derecho y las decisiones en el ámbito jurídico podrían beneficiarse de los conocimientos de la psicología jurídica sugiriendo una aproximación más amplia al rol del perito. Finalmente, se aplicará el conocimiento de la psicología cognitiva y de la psicología social, para explicar los errores en las pruebas forenses.

2. SOBRE REGLAS Y CRITERIOS DE ADMISIBILIDAD DE LA PRUEBA (CIENTÍFICA?)

Uno de los modos en los que el conocimiento científico, también el de la psicología, puede ser utilizado en el sistema de justicia es a través del testimonio del experto (perito) en los tribunales. Pero ¿cuándo y bajo qué circunstancias es admitida su intervención? Responder tal cuestión pasa por la referencia obligada a las Reglas Federales de Prueba americanas, en particular la 702, y a algunas resoluciones emitidas por la Corte Suprema estadounidense en los años noventa. No es intención de este trabajo entrar en el análisis detallado de las mismas. Se abordarán selectivamente ciertos aspectos, los más relacionados con cuestiones «psicojurídicas». Han sido muchos los autores que se han ocupado de ello (FAIGMAN, 2013; GATOWSKI et al., 2001; GROSCUP, J., 2004; VÁZQUEZ, 2015a, 2016), señalando las supuestas ventajas, los inconvenientes y la controversia sobre los factores o criterios que hayan de determinar la admisibilidad y cómo hacerlo. A pesar de las críticas (HAACK, 2020; NEUFELD, 2005; RISINGER et al., 2002, SCHAUER y SPELLMAM, 2013; entre otras), supusieron

un hito y un cambio de perspectiva que ha asomado también en nuestro país, habiendo al menos diez sentencias que refieren total o parcialmente tales criterios³.

El estándar de admisibilidad se establece a través de la denominada *Trilogía Daubert, Joiner, Kumho* (a partir de 1993), por una sucesión de tres casos que llegaron a la Corte Suprema. ¿Cuáles son sus elementos fundamentales? Los expertos serán sujetos cualificados que ayudan al juez y/o jurado a entender la evidencia aportada o determinar los hechos en cuestión. Cualquier testimonio experto tiene que ser no solo relevante, pertinente, en relación con la cuestión de la que se trate, sino fiable. La trilogía situaba en los jueces la responsabilidad como «puerta de entrada» de la información y elementos de prueba adecuados. Ellos son los encargados de valorar la fiabilidad y la validez de los métodos y principios que subyacen a la evidencia del experto (métodos científicamente válidos). Además, y esto es aun más exigente, suponía escrutar la validez de las inferencias entre análisis y aplicación al caso realizadas por los peritos. De modo que un testigo cualificado como experto podrá dar testimonio si está basado en datos y hechos suficientes, si el testimonio es producto de métodos y principios fiables y si se han aplicado adecuadamente a los hechos del caso en cuestión.

¿Cómo habría entonces de valorarse la fiabilidad y validez? De los criterios propuestos, cuatro son los más aludidos en la literatura: a) Si las teorías o técnicas empleadas por el científico-experto han sido probadas o pueden serlo, esto es, si son falsables: tienen que poder probarse empíricamente; b) revisión por pares o publicación; c) tasa de error conocida o que se puede conocer (de los instrumentos o técnicas empleadas para llegar a las informaciones o conclusiones de las que se trate); d) amplia aceptación en la comunidad experta. Reglas posteriores fueron matizando los criterios ampliando su aplicación no solo a conocimientos científicos sino también técnicos. Estas reglas suponían un filtro a cualquier tipo de conocimiento científico, técnico o artístico, si bien se aconsejaba flexibilidad dependiendo del ámbito. Los referidos factores son aplicables, por lo tanto, a una variada gama de disciplinas y saberes entre los que se encuentra la ciencia psicológica (FAIGMAN y MONAHAN, 2009; McAULIFF y GROSCUP, 2009; REISBERG, 2014), las ciencias forenses, la neurociencia, etc. y con ellos se sustituye supuestamente el estándar previo *Frye*, basado en la aceptación general de la comunidad de la proviniera el experto, y que se incorporó en cierto modo a la trilogía con ese matiz importante de «una amplia aceptación de la comunidad experta».

En Europa, en general, no existiría un control tan específico sobre la validez científica del testimonio experto, ni unas reglas similares de exclusión. Los criterios de

³ Las sentencias aludidas pueden solicitarse a la autora del trabajo. SAP CU 10/2011, de 14 de junio; SAP MU 20/2016, de 12 de enero; SAP V, 701/2014, de 28 de julio; AAP SE 87/2017, de 19 de enero; SAP M 216/2017, de 2 de junio; SAP M 419/2017, de 13 de diciembre; AAP CU 125/2018, de 17 de abril; SAN 28/2018, de 25 de junio; STS 2/2019, de 1 de febrero; STSJ M 264/2019, de 10 de diciembre.

relevancia, pertinencia, utilidad, necesidad o legalidad serían los habituales y se deja la cuestión de la fiabilidad hasta el momento de la valoración junto al resto de medios de prueba. Si la evidencia no es fiable, carecerá de todo valor probatorio. Ello ha llevado a cuestionar en qué medida puede confundirse fiabilidad científica con valor probatorio (CHAMPOD y VUILLE, 2010). En otros casos, se subraya la intervención del experto cuando el asunto en cuestión excede el conocimiento ordinario y la experiencia de quien decide, del juzgador (regla *Turner* o Law Commission, 2009/2011, en Inglaterra y Gales, especificando algunos criterios similares a *Daubert*).

Gran parte de las de las diferencias (ej. EE.UU-Europa) tienen que ver con las habidas entre sistemas *Common Law* y *Civil Law* y mucha de la literatura referida se ha ocupado de analizar sus distintas facetas. Entre ellas, el papel diferencial entre jueces y jurados (cuestiones de derecho/cuestiones de hecho). Y si bien no significa en absoluto que los criterios *Daubert* y otras reglas de exclusión desaparezcan cuando es el juez o el tribunal profesional el que decide sobre los hechos, hay algunos trabajos que constatan su flexibilización (SCHAUER, 2006; SAKS y SPELLMAN, 2016; SCHAUER y SPELLMAN, 2013). La literatura jurídica ha recogido distintas explicaciones para la necesidad u obligación de filtrar el testimonio de los peritos en el sistema estadounidense. Así, entre las que no tendrían que ver con la presencia del jurado, estaría el propio diseño normativo que distingue el «pretrial» del «trial» (VÁZQUEZ, 2016). Otros sugieren también que, al menos parcialmente, tal necesidad se basa en una asunción psicológica relevante (EDMOND y HAMER, 2012; KOVERA, y LEVETT, 2015; SPELLMAN, 2007; SPELLMAN y SCHAUER, 2012a b; WISSLER et al., 2013): la sobrevaloración del testimonio del experto por parte de los legos y el riesgo de que sean influenciados más allá de «lo razonable». Implícitamente, por lo tanto, se estaría asumiendo la extendida creencia en la distintividad y superioridad de los que tienen formación jurídico legal a la hora de decidir sobre los hechos y no ser vulnerables a influencias indeseables. Así, la existencia de reglas de admisibilidad diferentes, sean las que fueren, no tendrían tanto que ver con el hecho de que la decisión descansa en los jueces o en los jurados sino más exactamente con la necesidad de poner límites y controlar los excesos de los abogados respecto a la prueba en un sistema adversarial. Esto habría conducido a tratar de forma diferente el testimonio de los peritos frente a cualquier otro, lo que es un asunto controvertido en la literatura. Pero ¿habría razones para pensar que legos y profesionales fueran tan distintos?, y si lo son ¿en qué exactamente?

La investigación empírica ha tratado de responder a estas cuestiones. Sus resultados cuestionan muy claramente la supuesta superioridad al menos cuando se trata de determinar los hechos de un caso (KOVERA, y LEVETT, 2015; SPELLMAN, 2007; SPELLMAN y SCHAUER, 2012a y 2012b). Uno de los asuntos en los que la ausencia de diferencias ha sido la tónica, se refiere a la capacidad de unos y otros para no verse influidos por información no relevante: información o contenido del testimonio del perito, por ejemplo, que sería inadmisibile desde las reglas de prueba o procesales correspondientes. En este sentido, la investigación empírica ha mostrado que los jueces

resultan tan influenciados como los legos y ello afecta a sus decisiones (DEMAINE, 2008, 2012; STEBLAY, N. et al., 2006; SPELLMAN y SAKS, 2016; WISTRICH, GUTHRIE y RACHLINSKI, 2005). Por otra parte, y aunque los estudios con muestras de jueces no han proliferado, contamos con ciertas líneas de trabajo relevantes que vuelven a mostrar la vulnerabilidad de los profesionales a una gama variada de sesgos e influencias y que no son menos susceptibles a ellos que los jurados (CHARMAN et al., 2019; GARRIDO y HERRERO, 1997; GUTHRIE, RACHLINSKI y WISTRICH, 2002, 2007; OEBERST y GOECKENJAN, 2016; RACHLINSKI., 1998; 2000; RACHLINSKI, WISTRICH, y GUTHRIE, C, 2015; REDFIELD, 2017; Robbennolt y Hans, 2016; entre otros)⁴. Así mismo, no podemos olvidar a los abogados puesto que también tienen responsabilidad y protagonismo en los asuntos de prueba, incluso en nuestro sistema, y cuyo estudio empírico ha sido descuidado (DESPADOVA, KUKUCKA Y HILEY, 2020; FOX y BIRKE, 2002; STERNLIGHT y ROBBENNOLT, 2008; WECHSLER et al, 2015). Futuras investigaciones debieran profundizar en lo que todo ello implica.

Finalmente, por lo que a las reglas de admisibilidad se refiere, quisiera mencionar dos aspectos no menos controvertidos. El primero supone preguntarse si los jueces tienen la capacidad para asumir la responsabilidad que les otorga *Daubert*. ¿Pueden apreciar la fiabilidad y la validez del conocimiento científico? El segundo, alude a cuál sea el momento más adecuado para ejercer el control, el filtro. Los estudios de GATOWSKI et al., (2001), GROSCUP (2004), TADEI et al., (2016), o la línea de trabajo de KOVERA y McAULIFF (CUTLER y KOVERA, 2011; KOVERA y LEVETT, 2015; McAULIFF y GROSCUP, 2009; McAULIFF et al., 2009) muestran algunas realidades que no conducen al optimismo. Jueces y abogados no parecen detectar fácilmente problemas metodológicos que amenazan la validez o la fiabilidad de las conclusiones presentadas en los tribunales por los expertos. Las salvaguardas legales, tales como la contradicción, el contrainterrogatorio, etc. no parecen ser tan efectivas como podría suponerse⁵. Con estos, y otros resultados en mente, cabe reparar en la necesidad de

⁴ Debe recogerse un matiz que señala SCHAUER (2010a, nota 3 de su capítulo) y respecto del que no le falta razón. Los estudios empíricos disponibles se han centrado casi exclusivamente en el papel de los jueces, o en su caso jurados, en el momento de determinar lo que sucedió basándose solamente en las pruebas presentadas en el tribunal por las partes («fact-finding») y no en otras responsabilidades. Por tanto, sigue siendo válida la hipótesis de una potencial distintividad del razonamiento legal y el autor afirma que

...the most plausible version of the claim to distinctiveness in legal reasoning would be the comparative statistical claim that some of the methods of reasoning that are located outside of legal reasoning—arguments from precedent, reasoning from rules, and reliance on authority, most prominently—are more highly concentrated in legal argument and decision making than in ordinary reasoning and decision making, the difference being sufficiently great as to support the conclusion that legal reasoning is, in the aggregate, substantially unlike the kind of reasoning that takes place in other decision making domains (pp. 107).

⁵ Prueba de ello son los resultados emergidos recientemente en el estudio de NEAL y COLAB. (2019). En él realizaron una revisión de 364 herramientas de evaluación psicológica que los profesionales manifestaban haber utilizado en los tribunales. Además, efectuaron un análisis del grado en el que se rebatía la admisibilidad de las evaluaciones psicológicas. Los datos mostraron que el 90% de las herramientas de

aproximarnos al problema no solo desde las carencias mostradas por los operadores jurídicos, sino situarlo igualmente en los expertos y sus estrategias o modos de comunicar los hallazgos de sus disciplinas y aplicarlos a los casos (GARRET y MITCHELL, 2013). Y en este aspecto hay un amplio margen de mejora.

Sobre cuál sea el momento más adecuado para realizar el control sobre la calidad científica de la prueba que se aporta en el proceso existen en la literatura dos posturas contrapuestas. Por ejemplo, VÁZQUEZ (2015b, 2016) parece abogar por efectuarlo en el momento de la valoración. Por su parte GASCÓN (2013) ofrece una distinción interesante: habla de admisibilidad procesal y admisibilidad científica. Ambas resultarían necesarias y, además, la científica debiera aplicarse en un momento diferente al de la valoración de la prueba. Pues bien, dejando por un momento al margen los aspectos jurídicos o procesales implicados ¿nos permitirían las investigaciones en psicología realizar alguna sugerencia al respecto? Acudiendo, entre otros, a los estudios aludidos sobre el modo de pensar y razonar de legos y profesionales⁶, habría que sugerir que el control se realizara antes y separado de la valoración de la prueba. De no hacerlo ¿podríamos estar seguros de que la información científicamente cuestionable no tuviera algún peso en la valoración con consecuencias no predecibles? Esta debiera ser, probablemente, otra de las cuestiones que habría que poner a prueba.

En cualquier caso, los criterios Daubert, con todas sus ventajas e inconvenientes, son aplicables a todo tipo de conocimiento, también al de la ciencia psicológica y al de la ciencia forense. Para cualquiera de ellas, las afirmaciones científicas, como señala REISBERG (2014), deben basarse en buenos métodos y en buenas evidencias

evaluación utilizadas habían sido sometidas a prueba empírica, pero solo el 67% se consideraban como aceptadas por la comunidad científica de referencia. Así mismo, solo el 40% tenían revisiones favorables respecto a sus propiedades psicométricas y técnicas. Los datos respecto a la admisión mostraban que el cuestionamiento de la evidencia presentada era muy infrecuente, los abogados, además, rara vez rebatían las valoraciones de los expertos.

Por su parte GAROWSKI et al., (2001) puso de manifiesto que los jueces, independientemente del estándar de admisibilidad seguido en su estado no entendían la falsabilidad como criterio para admitir prueba científica. A pesar del consenso respecto a la utilidad de examinar las tasas de error, sus respuestas revelaban una falta general de entendimiento del significado científico de tasa de error (cuando se les preguntaba cómo aplicarían el concepto de error para la admisibilidad, solo un 4% de los jueces mostraron su comprensión). Sin embargo, había un claro entendimiento del proceso de revisión por pares. Finalmente, el factor de aceptación general, además de considerarlo útil o muy útil como los criterios previos, era claramente entendido por el 82% de los jueces cuando se les pidió que discutieran cómo lo aplicarían. Por su parte GROSCUP (2004) constató que los jueces no utilizaban los criterios Daubert, lo que no significaba que se inhibieran de evaluar la fiabilidad del experto en sus decisiones de admisión. Sencillamente acudían a otros indicios como las cualificaciones del perito o cómo su testimonio se adecuaba a los hechos del caso. Sin embargo, es primordial orientar la valoración no hacia la persona del perito y sus credenciales, sino a las afirmaciones vertidas en la pericia que esté realizando

⁶ En particular me refiero a los estudios sobre la incapacidad para no tener en cuenta información que no es fiable o que no estaría permitida pero que se filtra, a veces intencionadamente por parte de los abogados, en la práctica de la prueba (ej. determinadas preguntas, la introducción de valoraciones o calificaciones, o dejar que aparezcan algunas respuestas de peritos y otros testigos, etc.).

de esas afirmaciones. Una afirmación no puede tomarse en serio si los procedimientos no son los adecuados o la evidencia es escasa. Sin embargo, no parece que la ciencia psicológica sea percibida del mismo modo que la forense. Y aún asumiendo las diferencias, a veces significativas, entre disciplinas científicas ¿están justificadas las que parecen suponerse respecto de la psicología únicamente por sus avances? Esta es precisamente una de las cuestiones que abordaremos a continuación.

3. CIENCIA, PSICOLOGÍA VS SENTIDO COMÚN

3.1. Sobre la diferente consideración de la Ciencia Psicológica: alguna constatación

Se constatan desde distintos ámbitos y fuentes dudas respecto a la aplicabilidad de estándares científicos al conocimiento procedente de la psicología, de la psicología jurídica u otras de sus ramas (TARUFFO, 2013). Se habla, para referirse a ella, de «soft science» frente a «hard science» (MUNRO Y MUNRO, 2014), o de su pertenencia a las «ciencias de la comprensión». Los abogados prefieren expertos y pruebas del ámbito forense⁷ que de las ciencias sociales o del comportamiento (WECHLER et al., 2015). Algunos profesionales de la salud mental no ven posible utilizar nada parecido a los criterios Daubert o similares en su trabajo cotidiano, incluso cuando son llamados como expertos a los tribunales (SHAPIRO et al., 2015). Al tiempo, se constata el escepticismo público sobre ella (LILIENFIELD, 2012). En nuestro país, uno de los últimos estudios sobre la *Percepción Social de la Ciencia y la Tecnología* (FECYT, 2019), tampoco resultaba muy halagüeño respecto a la científicidad que se les atribuye a las ciencias sociales o a la psicología. Siendo una de las titulaciones más demandadas, casi el 50% de los encuestados la consideran como entre nada y algo científica. ¿En qué psicología se está pensando? ¿Por qué ese escepticismo? ¿Se está admitiendo ciencia de mala calidad o pseudociencia? Aunque los avatares administrativos relativamente recientes posicionaran a la psicología entre las ciencias de la salud, es también una ciencia social con múltiples facetas y, sí, también la denominada «clínica», que coloca al psicólogo/terapeuta en interacción cara a cara con el paciente/cliente para ayudar en distintas problemáticas. ¿Se está tal vez tomando el todo por la parte? Porque hay otras psicologías y otros profesionales/académicos, que también se dedican a la investigación utilizando el método científico. Es más, la práctica clínica no está exenta de la utilización de estrategias y tratamientos validados científicamente. En este contexto, y quizás más que en otros, tampoco «todo vale». Bien es cierto que los

⁷ Con ámbito forense se hace referencia en este momento a los resultados investigaciones y análisis realizados por disciplinas de las denominadas ciencias forenses y que pueden auxiliar a los tribunales en diferentes causas penales, civiles, laborales, etc. En el ámbito penal, se suele utilizar la denominación criminalística. Esta aclaración puede resultar necesaria dado que los profesionales de la psicología también realizan una labor forense en múltiples contextos aunque se haya venido asociando en menor medida esta disciplina a las ciencias forenses.

psicólogos, de cualquier clase, académicos o profesionales, hemos podido contribuir a esa visión con nuestras (malas) prácticas.

Alguna constatación más de la diferente consideración podríamos encontrarla en dos de los más reputados investigadores en psicología jurídica cuando consideran que el sistema legal se preocupa mucho y toma precauciones para evitar contaminar las trazas de evidencia física presentes, por ejemplo, en la escena de un delito (por ej., sangre, fibras), pero no se toman precauciones similares para evitar la contaminación de los recuerdos de los testigos. Aluden al gran abismo existente entre lo que la ciencia psicológica recomienda para recoger las declaraciones de los testigos y las prácticas reales de los investigadores de los delitos (WELLS y LOFTUS, 2003).

Sin embargo, ¿alguien pensaría en la «contaminación» cognitiva cuando se habla, por ejemplo, de ciencias forenses? ¿Cuál es el instrumento de análisis en muchas de las prácticas que sus profesionales y científicos realizan? Sí, algunos instrumentos tecnológicos. Pero uno fundamental: el ser humano y todo lo que ello implica. La psicología, como veremos en la última parte de este trabajo, nos muestra la implicación de importantes y complejos procesos cognitivos que permiten explicar las decisiones de los analistas y expertos que, como otros profesionales, testifican en los tribunales.

3.2. Sobre la diferente consideración: el sentido común como explicación

Plantearse la razón de esa percepción social sobre la psicología supone tener en cuenta otro tipo de conocimiento, diferente del científico y de las teorías o prácticas científicas, que ejerce su influencia en muy diversos contextos: el conocimiento de sentido común, las teorías implícitas sobre el ser humano, la psicología popular y sus mitos (LILIENTFIELD et al., 2010). Estamos ante una disciplina que estudia y explica fenómenos, procesos individuales y sociales, en los que todo ser humano está implicado, de los que todas las personas pueden tener una experiencia relevante. Y no hay evidencia «más científica» que la obtenida de lo que me ocurre a mí o a los cercanos, donde lo vivido tiene mucho más peso que cualquier información objetiva sobre el juicio humano (NISBETT y BORGIDA, 1970; KAHNEMAN y TVERSKY, 1973). Por ello, muchos contenidos psicológicos resultan familiares, se ha oído hablar de ellos o se ha indagado para satisfacer alguna curiosidad, cuando no se ha acudido a ciertos libros de (no) autoayuda en busca de la receta perfecta. Cuando las personas perciben que lo que piensan sobre un fenómeno coincide con lo que la ciencia ha establecido sobre él puede generarse la sensación de que la psicología, sus estudios o experimentos muestran cosas que ya se conocen, o que se estudia lo obvio. Cuando no coincide, no es extraño mostrar cierto escepticismo e incluso no creerse lo que descubren las investigaciones. Se buscarán en el entorno próximo evidencias, anecdóticas y no sistemáticas, pero muy efectivas, para confirmar que estamos en lo cierto. Todo ello ha contribuido a excluir más de lo deseable a la psicología como conocimiento científ-

fico, y también a alejarla del derecho, del sistema legal o de los tribunales durante largo tiempo. El escaso grado de socialización general en una cultura y razonamiento científicos, también juegan su papel. Esa deferencia al experto de otras ciencias (GASCÓN, 2013), que se considera tan problemática en el ámbito probatorio, parece dejar de serlo, con algunas relevantes excepciones, cuando a contenidos psicológicos concierne.

El problema se acrecienta, sin embargo, cuando se necesita la intervención de un perito, por ejemplo, en los tribunales o la participación de un experto (no necesariamente perito) en otros ámbitos que mejorarían la calidad de las decisiones en diversos contextos jurídicos. En el caso del perito su misión es asistir al juzgador para entender, interpretar o determinar unos hechos. Pero si el propio juzgador, o las partes implicadas, estima que ya se dispone del «acervo de conocimientos» que tendría cualquier «persona media o razonable» sobre el comportamiento humano, y que ello es suficiente para contribuir a resolver el objeto de controversia, fácilmente se prescindirá de sus servicios. Esto no supondría graves consecuencias si el sentido común de los implicados coincidiera con lo que la ciencia haya estipulado sobre la cuestión de la que se trate, pero las investigaciones nos muestran un panorama bastante diferente. Por otra parte, el testimonio experto del perito es mucho más fácilmente admisible cuando se abordan asuntos que parecen exceder el conocimiento común del juzgador o, como decíamos en las páginas iniciales, «la cultura promedio» de cualquier persona. Afortunadamente, son cada vez más las causas en las que se requiere la intervención del psicólogo (jurídico forense), pero suele ser casi siempre en una misma dirección: realización de peritajes, o emisión de informes según los casos, en los que hay que evaluar o diagnosticar psicológicamente a los sujetos, valorar daños o secuelas psíquicas, de competencia para declarar, cuando se trata de menores, de abusos, de adopciones, del impacto de determinados delitos sobre víctimas, o informes de los equipos psicosociales sobre casos de custodia o, menos frecuentemente, pueden ser requeridos en el ámbito del derecho privado o laboral.

Las limitaciones y el contenido de las creencias de sentido común sobre temas variados de psicología en general o de psicología jurídica en particular, han sido investigadas a través de estudios empíricos. Se ha podido comparar las creencias y conocimientos de distintos grupos con lo que descubierto por la psicología, así como analizar las diferencias entre distintas muestras relevantes para el sistema de justicia. ¿Qué nos dicen tales estudios? En ciertos casos, como no podría ser de otro modo, el sentido común del público en general y de los operadores jurídicos en particular, coincide con la ciencia. Pero, en general, indican que los abogados, jueces, jurados o potenciales jurados y la policía utilizan sus creencias, consideradas de sentido común, no validadas científicamente respecto a múltiples aspectos relacionados con la memoria, con el grado de exactitud de los testigos y sus testimonios en adultos y menores o con la detección del engaño, el desempeño profesional de los juristas, o con la percepción del testimonio del experto (BENTON et al., 2006; DESMARAIIS and RED, 2011; GRANHAG et al., 2005; HERRERO et al., 2016; HOUSTON et al., 2013;

KNUTSSON y ALLWOOD, 2014; KASSIN et al., 2001, 2018; MAGNUSSEN et al., 2008, 2010, 2013; OTGAAR et al., WISE et al., 2004, 201, 2014) y otras cuestiones que implican procesos psicológicos y del comportamiento humano. Además, se detecta el alto grado de certeza con el que las mantienen. Sin entrar pormenorizadamente en cada una de sus aportaciones, debe subrayarse que, en especial cuando se trata de jueces, se encuentra un conocimiento razonable sobre cuestiones generales de testimonio, pero también una amplia variabilidad en la consistencia con las opiniones de los expertos que no se encuentra en otras muestras (GRANHAG et al., 2005; WISER y SAFER, 2004). Prueba de ese conocimiento general sobre testimonio se aprecia en el trabajo de HOUSTON ET AL. (2013) cuyos resultados muestran que en el 67% de las ocasiones los jueces coincidían con el conocimiento de los expertos. Alguna diferencia se ha hallado también respecto al conocimiento sobre los indicios para detectar el engaño (BROWNSSELL, A. y BULL, 2011) de modo que los jueces mostraban de nuevo mayor variabilidad y mantenían creencias menos estereotipadas e incorrectas sobre los indicios no verbales.

Otros resultados relevantes, sin embargo, ponen de manifiesto que el conocimiento más detallado sobre otros múltiples temas es bastante más escaso, con lo que mantienen creencias cuya generalización es sesgada e incompleta (ej., los testigos son fallibles o los menores son más sugestionables, etc.). De hecho, algún trabajo identifica (TUPPER et al., 2017) el mantenimiento de al menos cinco mitos o falsas creencias sobre el testimonio: a) todos los testigos son poco fiables; b) todo el mundo conoce cuándo los testigos no son fiables; c) la consistencia es sello distintivo de los testigos que son fiables; d) ¡cuanto más mejor! Si varios testigos lo dicen debe ser verdad; e) «esto es solo sentido común»! Sobre temas que no se relacionan con el testimonio, las creencias que no son acordes con el consenso de los expertos están bastante más extendidas y aparecen en todas las muestras estudiadas. Es necesario, sin embargo, obtener más evidencia empírica, en especial en nuestro país, donde los estudios con jueces y operadores jurídicos son aún más escasos. Y hacerlo es relevante porque no se trata simplemente de creencias, sino del efecto que generan en la práctica profesional como tendremos ocasión de constar en las páginas siguientes.

4. CIENCIA PSICOLÓGICA E INTERVENCIÓN DEL EXPERTO

4.1. ¿Es necesario el testimonio del experto en psicología, en Psicología Jurídica?

Después de lo expuesto parecen claras las limitaciones del sentido común y que las investigaciones en psicología jurídica ofrecen mejoras sustanciales al mismo. Por otra parte, ya hemos visto varias de las carencias y asunciones psicológicas (en ocasiones, sin fundamento) que intervienen en distintos momentos con relación a la admisibilidad de la prueba y a la determinación de los hechos. Muchas de las pro-

blemáticas emergidas no parecen solucionarse necesariamente con las salvaguardas procesales previstas en los distintos sistemas jurídicos.

Pero añadamos algún resultado más. Como se muestra en los trabajos de KOVERA y colaboradores (CUTLER y KOVERA, 2011; KOVERA y LEVETT, 2015, McAULIFF y GROSCUP, 2009, entre otros), el testimonio del perito ejerce influencia sobre los jueces y jurados, sobre todo si es explícito, si el experto tiene buen currículum y credenciales. El testimonio basado solo sobre juicios clínicos o juicios (subjetivos) del profesional tiene más influencia que el basado en datos e investigación empírica y actuariales o estadísticos. O el que presenta una descripción cualitativa del grado en que dos muestras forenses coinciden es más persuasivo que el testimonio que se presenta en términos cuantitativos. Igualmente, las reacciones de los testigos, más o menos emocionales, así como si son en vivo o realizadas por videoconferencia, afectan a los juicios de credibilidad (MELINDER, BURREL, ERIKSEN, MAGNUSSEN y WESSEL, 2016). Así mismo, tal como señalan SAKS y SPELLMAN (2016), los seres humanos no venimos bien equipados para evaluar y utilizar la evidencia científica y estadística. Los operadores jurídicos pueden tener dificultades para comprender las probabilidades y el razonamiento científico. Pero también psicólogos, los médicos y otros profesionales con labores forenses, o sin ellas. Sin una formación específica, e incluso con ella, se tendrá dificultades. No obstante, algunos de los problemas asociados a esos tipos de razonamiento mejoran cuando se sustituyen las probabilidades por frecuencias naturales (GIGERENZER, 2015). Ello, como ejemplo sencillo, abunda en la idea de replantearse cómo se comunican y se dan informaciones sobre hallazgos científicos.

Así mismo, para gran parte de los sesgos y heurísticos que pueden aparecer en el proceso de obtención, análisis e interpretación, los jueces, abogados o fiscales no se diferencian mucho de los legos, como ya hemos señalado. Además de la dificultad para no tener en cuenta información inadmisibile aparecen otras (ej. efecto «hind-sight», efectos de anclaje en las decisiones judiciales, falacias varias, sesgo confirmatorio, escepticismo asimétrico, «blind spot», etc.) (ENGLISH, y MUSSWEILER, 2001; GARRIDO y HERRERO, 1997; GUTHRIE et al., 2001; 2007; OEBERST y GOECKENJAN, 2016; RACHLINSKI, 1998; RACHLINSKI et al., 2015). Estudios específicos realizados con abogados nos muestran otras problemáticas relacionadas (DESPADOVA et al., 2020; FOX y BIRKE, 2002; STOLZENBERG y LYON, 2014, entre otros).

Después de lo expuesto cabría preguntarse en qué medida es necesaria la intervención o las aportaciones del experto en psicología jurídica y en qué se concreta su participación. ¿Es posible no reducirla a la realización de un peritaje para un caso concreto? Los resultados ofrecidos parecen sugerir la pertinencia del experto, pero probablemente no avancemos lo deseable reduciendo a este únicamente al rol de perito, al menos tal como es concebido.

4.2. Contribuciones e implicaciones del conocimiento experto en psicología jurídica

Más allá de los asuntos clásicos y otros más novedosos implicados en diversos peritajes, o del importante papel que los psicólogos forenses realizan en distintos ámbitos de la administración de justicia, se muestran en este epígrafe ciertos ejemplos que, desde mi punto de vista, darían pie a la discusión sobre la ampliación referida en el epígrafe anterior. Así mismo, permiten ilustrar los efectos de la utilización de creencias de sentido común, no validadas científicamente o, sencillamente, los de no acudir a la evidencia empírica aportada por psicología. Sobre algunos me he ocupado brevemente en otra parte (HERRERO, 2019). Los retomaré y se añadirán otros pero su abordaje pormenorizado excede las pretensiones del presente trabajo así como el espacio disponible con lo que me limitaré a enunciarlos sucintamente.

a) Sobre las necesidades de las víctimas, testigos y ciudadanos en general: la psicología social de la justicia es probablemente una de las aproximaciones más relevantes cuando hablamos de víctimas, de sus necesidades y del tratamiento que reciben de su entorno, así como de los profesionales que tienen algún contacto con ellas (HERRERO, 2019). Sin embargo, aludir al tratamiento de las víctimas en contextos jurídicos suele entenderse principalmente como la posibilidad de aplicar atención, terapia o realizar intervención psicológica con las personas afectadas por el delito. Uno de los resultados más sólidos de esta línea de trabajo pone de manifiesto que sentirse tratado justamente tiene más que ver con los *procedimientos* y/o procesos empleados para tomar decisiones que afectan a las personas (víctimas, testigos, ciudadanos, delinquentes, menores, etc.) y a la solución de sus casos que con los *resultados* obtenidos a través de tales procedimientos. Ello supone que puede ser tanto o más importante el «cómo» que el «qué», al menos más de lo que el mito del autointerés nos llevaría a predecir y tiene implicaciones relevantes en el contexto jurídico. Este enfoque nos permite ir más allá de la mera concienciación o sensibilización de los profesionales y nos facilita el diseño de acciones concretas derivadas de las distintas dimensiones y elementos de la justicia procedimental y que suelen ser desconocidas para los juristas.

b) Sobre la valoración-gestión del riesgo de violencia y reincidencia: esta cuestión constituye uno de los tres aspectos más relevantes en la prevención de la violencia (ej., de género). Los otros serían su detección (proactiva, temprana) y la denominada intervención de «los espectadores» (Herrero, 2019). Sin embargo, distintas afirmaciones públicas realizadas por operadores jurídicos, las críticas vertidas sobre su eficacia así como lo escrito en algunos autos o sentencias⁸, sugieren que pueda estar ope-

⁸ Una prueba de la problemática con la referida valoración puede ilustrarse en las siguientes afirmaciones de distintos operadores jurídicos

... En este diagnóstico de peligrosidad creo que es de vital importancia la apreciación, aunque sea subjetiva, del Fiscal y del Juez, debiéndose respetar escrupulosamente el principio de inmediatez, pues de esta forma, en ocasiones, constatamos no solo la agresividad del hombre y su especial ensañamiento e intención de humillar, sino también esa enfermiza relación de dependencia con la

rando una visión desvirtuada de sus características o banalizando la complejidad de su realización (GÓMEZ-HERMOSO, MUÑOZ-VICENTE... et al., 2012). Además, existe probablemente un desajuste entre la utilización del concepto legal de *peligrosidad* y el *valoración-gestión del riesgo* haciéndolos equivalentes cuando no lo son (ANDRÉS PUEYO, 2012; 2014).

c) Sobre cómo interrogar, preguntar, interactuar con las víctimas, testigos, peritos/expertos o acusados: la psicología puede resultar un conocimiento indispensable cuando se trata de plantear buenas prácticas para obtener información de calidad de víctimas, testigos, o acusados a través de entrevistas e interrogatorios, y que sirva entre otras cosas a fines probatorios. La información obtenida no depende únicamente de lo vivido por la persona sino del tipo y el modo en que se formulen las cuestiones. Especialmente cuando estamos ante menores o sujetos vulnerables se acrecienta, más si cabe, la necesidad de prácticas basadas en la evidencia. Afortunadamente contamos ya con protocolos y modalidades de entrevista con buenos resultados (BULL, 2010; COOPER et al, 2018; DICKINSON et al., 2019; GONZÁLEZ Y MANZANERO, 2018, para revisión). Algunas de ellas se han desarrollado especialmente en contextos de investigación policial. No obstante, los principios psicológicos que los sustentan permiten que su esencia se aplique a otros ámbitos, tales como a los profesionales de la justicia, poniendo su eficacia a prueba en futuras investigaciones. Además, contamos con un reducido número de estudios que ilustran los problemas que los operadores jurídicos tienen cuando toman declaración o hacen preguntas, pero se han circunscrito casi exclusivamente en el ámbito de menores y para delitos sexuales.

Sorprende, sin embargo, que en nuestro país no sea habitual la interacción de profesionales y académicos/investigadores de la psicología jurídica y el derecho a la

víctima, obstinándose en mantener la relación a toda costa, oponiéndose a que se adopte la orden de protección, así como justificando y minimizando los hechos... (A. G. Fiscal Violencia de Género, 2013, la negrita es añadida)

«No podemos descubrir a un potencial asesino de su pareja de 40 preguntas» R. B., juez de violencia de género, reclama medios de control para los denunciados y califica la gestión de la violencia machista de «coladero» (La Voz de Galicia, 1 de octubre de 2018)

... estamos ante un criterio de prevención especial sustentado en el concepto de peligrosidad del imputado. Se exige al juez que realice una prognosis a partir del hecho concreto y de las circunstancias conocidas del imputado y que se especule sobre una posible conducta futura... La peligrosidad es un pronóstico de futuras conductas a partir de inferencias sobre algunos datos de hecho del pasado. El margen de error e incertidumbre es muy elevado al fundarse sobre bases poco rigurosas y criterios poco asentados (aún los científicos, cuando existen). Los criterios científico clínico, estadístico o el criminológica... no pueden ser aplicados por cuanto se necesitan medios personales y materiales... que en este juzgado no existen; además cuando existan estos medios se necesitará tiempo de estudio y reflexión por los expertos. Así las cosas, el criterio de decisión sobre la peligrosidad queda confiado exclusivamente a la simple intención del juez a partir de unos datos muy escasos cuando existen (v. gr. los antecedentes penales) y el juez debe resolver con una urgencia tal que la decisión puede quedar huérfana de la necesaria reflexión, estudio y rigor, en la mayoría de los supuestos... pese a estas carencias esenciales la decisión puede fundarse en este endeble e infundado (la mayor parte de las veces) criterio de la peligrosidad (Auto de denegación de medidas cautelares en las diligencias en un supuesto caso de violencia de género. Salamanca, octubre de 2016)

hora, por ejemplo, de confeccionar guías o «libros blancos» que sirvan de pauta en la práctica de los operadores jurídicos⁹, con las ventajas que ello supondría.

d) Sobre el plazo de prescripción de algunos delitos: en Francia recientemente se amplió el plazo de prescripción de los delitos de abusos sexuales a menores. La justificación de la modificación fue la existencia de la represión (inconsciente) de los recuerdos (rebautizado por otros como amnesia disociativa) para experiencias traumáticas, y su posterior recuperación incluso muchos años después. Sin embargo, tal fenómeno, de nuevo, es más bien una creencia extendida en contextos clínicos, jurídicos o académicos que no viene apoyada por la evidencia científica (DODIER y TOMAS, 2019; OTGAAR et al., 2019; OTGAAR et al., 2020; PATIHIS et al., 2014). En España, con la aprobación en junio del presente año del Proyecto de Ley Orgánica de protección integral a la infancia y adolescencia frente a la violencia se extiende también el tiempo de prescripción de tales delitos. Este comenzará a contar a partir de que la víctima haya cumplido 30 años lo que supone que los delitos más graves serán perseguibles hasta que sus víctimas alcancen los 45 años. En este caso, no aparece alusión explícita alguna a la existencia de la represión de los recuerdos del abuso. Se indica que la medida «evita la existencia de espacios de impunidad en delitos que estadísticamente se han probado de lenta asimilación en las víctimas en el plano psicológico y, muchas veces, de tardía detección». No obstante, algunas de las enmiendas presentadas y recogidas en el *Boletín Oficial de las Cortes Generales*, de 5 de noviembre de 2020, elevan más el plazo empezando a contar a partir de los 40 años.

Cabe preguntarse, además, si con estos cambios normativos se mejora realmente la posición de las víctimas o tendrán los efectos deseados, esto es, acabar con la impunidad de los agresores y que, por lo tanto, aumenten el número de personas de incluso más de cuarenta años que denuncian los abusos sufridos en la infancia.

e) Sobre la eliminación de la dispensa de la obligación de declarar (art. 416 LE-Crim): asunto controvertido donde los haya cuando se trata, por ejemplo, de víctimas de violencia de género. Se ha defendido ampliamente su eliminación¹⁰ y parece

⁹ En este sentido puede constatarse las deficiencias de algunas de las recomendaciones planteadas en la «Guía de buenas prácticas para la toma de declaración de las víctimas de violencia de género» (2018), aprobada por el grupo de expertas y expertos en violencia doméstica y de género del Consejo General del Poder Judicial. Especialmente las referidas a las preguntas sobre la determinación de factores de riesgo, que acentúan la creencia no solo sobre lo inadecuado de la formulación y formato de las preguntas, sino sobre los problemas en la concepción sobre la valoración y predicción del riesgo de violencia. Así mismo, sorprende la seguramente bien intencionada recomendación respecto a cuando no es posible que sea el juez/a el que tome declaración a la víctima por estar realizando otras diligencias. En ese caso se considera «una buena práctica el especializar a alguno de los/las funcionarios/as para tal cometido, desarrollando un guión o modelo de denuncia, a fin de evitar, cuando el procedimiento tenga que ser inhibido a otro juzgado, una nueva declaración de la víctima con objeto de completar los extremos de la denuncia».

¹⁰ Véanse las medidas adoptadas por el Pacto de Estado contra la violencia de Género, en concreto la 142 del texto refundido (2019), así como el contenido de las comparecencias de los distintos expertos ante el Congreso y Senado.

haberse optado por ella, aunque no es pacífico el acuerdo sobre sus potenciales beneficios (GONZÁLEZ-MONJE, 2019). Sus partidarios la justifican por el elevado número de mujeres que se acogen a ella lo que conduce supuestamente a la impunidad de los maltratadores. Así mismo, parece asumirse con facilidad que las presiones de los agresores o sus familias o el miedo a las represalias están en el origen de la renuncia a seguir en el procedimiento. ¿Pero es esto realmente así? Asumiendo los desajustes e incompatibilidades que tal artículo ocasiona con la propia Ley Orgánica 1/2004 y algunas de sus modificaciones, o los múltiples argumentos jurídicos implicados, sería recomendable que decisiones de tanto calado se tomaran conociendo exactamente cuántas mujeres se acogen a la dispensa, pero sobre todo, cuáles son las razones para hacerlo, bajo qué circunstancias y en qué momento o fase del proceso por el que atraviesan las víctimas es más probable que ocurra, cómo interactúan las variables, y entre ellas, cuáles constituyen auténticos predictores de la renuncia. Solo así se podrá intervenir en el problema. En este sentido, son esclarecedores los trabajos de Cala y colab. (CALA et al., 2012; CALA, TRIGO y SAAVEDRA, 2016; GARCÍA-JIMÉNEZ, et al., 2018) cuando ponen de manifiesto las variables sociodemográficas, psicosociales y jurídicas que predicen la renuncia a continuar en la causa o el papel que juega el proceso de lo que denominan separación psicosocial por el que atraviesan las víctimas (GARCÍA-JIMÉNEZ, DURFEE, CALA, y TRIGO, 2020). Y de todo ello es imprescindible que tengan conocimiento los operadores jurídicos, pero sobre todo los legisladores. De nuevo, con la eliminación de la dispensa ¿se mejora la posición de la víctima o la calidad de sus declaraciones? ¿Cuáles son los efectos esperados?

f) Sobre intermediación, credibilidad y valoración de las declaraciones de las víctimas: especial referencia a las sentencias y orientaciones de la Sala Segunda de lo Penal del Tribunal Supremo (se aborda en el apartado 4.2.1).

g) Sobre psicología y derecho de daños: otro de los contenidos relevantes en los que confluyen derecho y psicología tiene que ver con los daños, su valoración y la asignación de responsabilidad. Como muestra el reciente trabajo de ROBBENNOLT y HANS (2016) el derecho de daños implica aspectos o contenidos sobre el comportamiento humano con larga tradición en la teoría e investigación psicosocial. Desde qué motiva a las personas a presentar una demanda, determinar si una conducta particular es accidental o intencional, razonable o no razonable, cómo piensa la gente sobre los factores causales del daño, cómo se asigna culpa o responsabilidad, la determinación de compensaciones, son todos áreas de importante confluencia.

h) Sobre el impacto y manejo de herramientas visuales en la admisibilidad, presentación e interpretación de la prueba: los avances tecnológicos experimentados en las sociedades han llegado también al ámbito jurídico y a los tribunales. Ello ha supuesto, entre otras cosas, la utilización cada vez más frecuente de tecnologías de comunicación e información, digitalización o presentación de información a través de medios visuales y no únicamente como apoyo a la presentación de información por parte de peritos, abogados u otros expertos y profesionales. El uso de videgrabaciones, videoconferencias, grabaciones de interrogatorios y declaraciones, reconstruc-

ciones 3D, grabaciones de las cámaras corporales de la policía, además las más comunes presentaciones PowerPoint, fotografías, infografías, gráficas, etc., son algunas de las posibilidades (FEIGENSON, 2010; FEIGENSON y SPIESEL, 2011; GRANOT et al., 2018; HENDERSON y EPSTEIN, 2011, entre otros). ¿Qué influencia tiene la utilización de este tipo de herramientas en el contexto jurídico, en el proceso? Distintos estudios empíricos han puesto de manifiesto ciertos efectos positivos pero también otros potencialmente contraproducentes. De modo que puede optarse por limitar normativamente su utilización, por ejemplo, como apoyo en la presentación de información e incluso como elementos de prueba útiles y relevantes. Pero entonces ¿nos privamos de sus ventajas o, por el contrario, tenemos en cuenta las condiciones que permiten acentuar sus beneficios al tiempo que se eliminan o atenúan los inconvenientes? La investigación nos da algunas respuestas en esta segunda dirección.

i) Sobre los sesgos en la obtención e interpretación de la prueba y la toma de decisiones en el contexto jurídico: otra de las líneas más sólidas en psicología social y cognitiva tiene que ver con el conocimiento de los potenciales sesgos, cognitivos o motivacionales, que pueden aparecer en cualquier persona como sujeto activo al percibir, obtener, procesar, interpretar, etc. información, también en el contexto jurídico y por parte de los operadores jurídicos. Un claro ejemplo de la utilidad de esta línea se mostrará al explicar los errores en las ciencias forenses.

Sirvan esas sugerencias como pinceladas de las posibilidades en las que puede intervenir el experto en psicología social y cognitiva, en psicología jurídica y que van más allá de la realización de un peritaje concreto. Todos ellos, en cualquier caso, sugieren otros ámbitos de actuación y diversas formas de utilizar su experticia. Al tiempo, ponen sobre la mesa la idea de que los conocimientos que aporta esta disciplina deberían incluso ser tenidos en cuenta en el diseño del proceso a nivel legislativo, así como también en otras normas legales sustantivas o procesales.

4.2.1. *Sobre inmediatez, credibilidad y declaraciones de las víctimas*

La evaluación de la credibilidad de las declaraciones de víctimas, testigos, acusados, así como el análisis de las potenciales estrategias que ayuden a discriminar quién miente o dice la verdad son líneas consolidadas en psicología jurídica y parecen de absoluta pertinencia en los tribunales. Sin embargo, es una de las áreas donde mayores mitos, falsas creencias y prácticas pseudocientíficas surgen dentro y fuera de ellos, como trataré de mostrar (DENAULT y DUNBAR, 2019; DENAULT et al., 2020; HERRERO, MASIP, y PICADO, 2016; MASIP, 2005; VRIJ y FISHER, 2020).

La revisión del contenido tanto de Sentencias de Audiencias Provinciales como de las más recientes de la Sala Segunda de lo Penal del Tribunal Supremo español, (STS 119/2019, de 6 de marzo; STS 184/2019, de 2 de abril; STS 282/2018, de 13 de junio; STS 247/2018, de 24 de mayo, entre muchas otras) ponen de manifiesto justamente la problemática sobre los efectos negativos de la utilización de conoci-

mientos no científicos en el derecho en general y en los tribunales en particular. ¿Por qué? Por la extendida utilización de ideas y creencias de sentido común, intuiciones e impresiones subjetivas, cuando no de mitos, falsas creencias, y ciertas prácticas o concepciones erróneas y sesgadas, al menos cuando de asuntos de contenido psicológico se trata. La experiencia, la lógica y las nociones de sentido común en materia de comportamiento humano suelen estar equivocadas, ser incorrectas y falibles, como se ha visto en epígrafes previos (COOPER, DANDO... BULL, 2018; CHAPLIN y SHAW, 2015; KNUTSSON y ALLWOOD, 2014; LILIENFELD, LYNN, RUSCIO y BEYERSTEIN, 2010; MAGNUSSEN et al., 2010; SPORER, 2007). A esta problemática tampoco son ajenas otras disciplinas como la psicología y sus profesionales.

Dos de las creencias más compartidas, y mantenidas con mayor grado de certeza, entre el alumnado de derecho son las referidas a la intermediación como condición imprescindible para apreciar y valorar la credibilidad de los testigos (con la que cerca del 90% está de acuerdo), así como la que alude a la superioridad de la información/comunicación no verbal sobre la verbal para saber si alguien está mintiendo o diciendo la verdad (92%) (HERRERO et al., 2016; HERRERO, 2016). Porcentajes similares se encuentran en muestras de criminólogos, profesionales juristas y policías, si bien para estos dos últimos grupos el referido a la intermediación desciende hasta el 76%). No resulta sorprendente si en el análisis de sentencias que relacionan intermediación y credibilidad, más veces de las deseables, aparecen afirmaciones como las siguientes:

«... Juzgador de instancia quien se halla en condiciones óptimas para decidir sobre la credibilidad que ha de darse a lo oído y visto en el juicio oral, pues cuando el medio de prueba es una persona, la convicción judicial se forma también por los gestos, expresión facial, tono de voz firmeza, duda de las manifestaciones, inseguridad o incoherencia en las mismas etc., que el Juzgador puede apreciar y valorar en consecuencia...» (Sentencia de la Audiencia Provincial de Madrid, 29 septiembre de 2005).

«... apreciando personal y directamente, sobre todo en la prueba de testigos, su expresión, comportamiento, rectificaciones, dudas, vacilaciones, seguridad, coherencia, y en definitiva, todo lo que afecta a su modo de narrar los hechos...» (Sentencia de la Audiencia Provincial de Salamanca, 12 de julio de 2011).

«... el nerviosismo, la dinámica del gesto, del silencio, del tono de voz, de los mismos titubeos, vacilaciones y contradicciones, interacciones entre las partes, amén de otros datos esenciales, en orden a comprobar la credibilidad del sujeto compareciente ante el órgano judicial, como integrantes de la psicología del testimonio, son factores de los que solo puede disponer el juez o magistrado que vive el desarrollo del juicio en instancia...» (Sentencia de la Audiencia Provincial de Salamanca, 11 de junio 2009).

«... Las primeras aparecen constituidas por los datos probatorios estrechamente ligados a la intermediación, tales como el lenguaje gestual, la expresividad de las manifestaciones, nerviosismo o azoramiento en las declaraciones, titubeo o contundencia en las respuestas, rectificaciones o linealidad en su exposición, tono de voz y tiempos de silencio, etc... Junto a la anterior hay una zona franca y accesible de las declaraciones integrada por los aspectos relativos a la estructura racional del propio contenido de la prueba, que al resultar ajenos a la estricta percepción sensorial del juzgador "a quo" sí pueden ser fiscalizados a través de las reglas de la lógica, la experiencia y los conocimientos científicos.» (Sentencia de la Audiencia Provincial de Madrid, 13 junio 2013).

Obsérvese cómo se recurre a la intermediación como medio de apreciar la credibilidad y esta, además, parece sustentarse en elementos de lenguaje o comunicación no verbal (nervios, sonrojos, azoramiento, expresividad facial, tono de voz, seguridad e inseguridad, etc.) solo disponibles al juzgador «a quo». Sorprende, sin embargo, que sea en las Sentencias de la Sala 2ª del Tribunal Supremo referidas (entre muchas otras) donde tales elementos aparecen en todo su esplendor. Más aún si se trata de resolver asuntos sobre violencia de género o, si a través de ellas, parece pretenderse establecer «precedente» sobre los criterios orientativos en la valoración de la credibilidad de las víctimas, en particular cuando supuestamente no hay muchos más elementos de prueba. Unido a ello, y no con menor intensidad, se encuentran creencias sobre los estados psicológicos y su contribución en una declaración o sobre el papel diferencial y privilegiado de la víctima, no como mero testigo, sino como «observador» cualificado por haber vivido unos hechos, en especial cuando su declaración es el único medio de prueba disponible para desvirtuar la presunción de inocencia. Así, en una de las más relevantes, pero no novedosa, se indica que

...recordemos que es posible que el Tribunal avale su convicción en la versión de la víctima, ya que la credibilidad y verosimilitud de su declaración se enmarca en la apreciación de una serie de factores a tener en cuenta en el proceso valorativo del Tribunal. Y así podemos citar los siguientes: 1-seguridad en la declaración ante el Tribunal por el interrogatorio del Ministerio Fiscal, letrado/a de la acusación particular y de la defensa; 2-concreción en el relato de los hechos ocurridos objeto de la causa; 3-claridad expositiva; 4- «lenguaje gestual» de convicción. Este elemento es de gran importancia y se caracteriza por la forma en la que la víctima se expresa desde el punto de vista de los «gestos» con los que se acompaña en su declaración; 5-seriedad expositiva que aleja la creencia del Tribunal de un relato figurado, con fabulaciones, o poco creíble; 6-expresividad descriptiva en el relato de los hechos; 7-ausencia de contradicciones y concordancia del iter relatado de los hechos; 8-Ausencias de lagunas en el relato de exposición que pueda llevar a dudas de su credibilidad; 9-la declaración no debe ser fragmentada; 10-debe desprenderse un relato íntegro de los hechos y no fraccionado acerca de lo que le interese declarar y ocultar lo que le beneficie acerca de lo ocurrido; 11-debe contar lo que a ella y su posición beneficia como lo que le perjudica...

Pues bien, solo podemos calificar tal propuesta de inadmisibles. Mucho de su contenido, en lo que a contenidos psicológicos se refiere, está en contra de la evidencia empírica más consolidada sobre evaluación de credibilidad y detección de la mentira (DENAULT y DUNBAR, 2019; DENAULT et al., 2019; GARRIDO, MASIP y HERRERO, 2006; KASSIN et al., 2001, 2018; MASIP, 2017; MASIP y HERRERO, 2015, 2015b, c; MASIP, HERRERO, GARRIDO y BARBA, 2011; MANZANERO y MUÑOZ, 2011; SPORER, 2007; VRIJ, 2018, VRIJ, HARTWIG y GRANHAG, 2019). Y ello porque acuden a la necesidad de apreciación de factores que tienen que ver básicamente con indicadores de comportamiento no verbal o a la consideración de cómo se declara (seguridad en la declaración, expresividad descriptiva, seriedad expositiva, ausencia de lagunas, entre otras). De hecho, se resalta expresamente la importancia del «lenguaje gestual» de convicción. Cuestiones como estas son las que hacen plantear si la credibilidad y su prueba, o la convicción judicial particularmente en el caso de la testifical, no son

acaso un mero asunto de persuasión alejado de la racionalidad que a renglón seguido esas mismas sentencias dicen profesar para alcanzar las decisiones.

Otro problema añadido debe mencionarse, a saber: parece que se esté pervirtiendo en toda regla el sentido fundamental de la intermediación como medio de producción de la prueba. Si, como no podría ser de otro modo, es un principio imprescindible, no debiera serlo porque a través de ella el juzgador pueda percibir, en el caso de las pruebas personales, supuestos indicadores que lleven a inferir conocimiento empírico válido (probatorio) alguno (ANDRÉS IBÁÑEZ, 2003; FERRER, 2017). Por lo tanto, es preciso que los operadores jurídicos sepan que no existe ningún indicador conductual que puedan observar que, por sí mismo o en una combinación prefijada con otros, permita inferir con mínimas garantías que alguien está mintiendo o diciendo la verdad (DENAULT et al., 2019; MASIP, ALONSO Y HERRERO, 2006; DE PAULO et al., 2003 BOND Y DEPAULO, 2006; SPORER, 2007). Son pocos los indicadores que existen y están moderados por múltiples variables y contextos. Además, somos malos detectores, y peores aun si nos centramos en conductas o indicios no verbales, dada la superioridad de los verbales (DE PAULO et al., 2003) y el entrenamiento, como ha puesto de manifiesto la investigación, no mejora la detección (HAUCH et al., 2016). Sin embargo, las personas perseveran en sus ideas y estas no son fácilmente modificables. Pero cuando lo que nos ocupa son asuntos relacionados con el sistema de justicia, sus instancias y decisiones, la difusión de ideas pseudocientíficas como las expresadas en las sentencias, aparentemente inocua, debiera restringirse. Y aun no teniendo efecto directo, o sí, sobre el resultado de un proceso concreto, puede tenerlos de otro tipo. Sirva como una ilustración más lo que ocurre con las desacreditadas (BURGOON, 2018; JORDAN et al., 2019; KASSIN et al., 2018) «micro-expresiones» como medio para detectar el engaño (EKMAN, 2009, 2011), otra faceta más del «lenguaje gestual de convicción». Tal aproximación se asume acríticamente, incluso por parte de los profesionales (NIETO, 2019) y ligada de nuevo a la intermediación. O «se cuele» en las sentencias aunque sea para, en su valoración, no otorgarle afortunadamente peso alguno (Sentencia de la Audiencia Provincial de Barcelona, 335/2019, 8 de mayo; Sentencia de la Audiencia Provincial de Barcelona, 1158/2018, 12 de diciembre).

Respecto a lo otra cuestión implicada, la cualidad de testigo privilegiado de la que afirma ser víctima, a las críticas y argumentos de calado esgrimidos desde el ámbito jurídico (ANDRÉS IBÁÑEZ, 2020; RAMÍREZ ORTIZ, 2020) habría que añadir lo que REISBERG (2014) formula para las denominadas «flashbulb memories», producto de hechos emocionalmente impactantes. El mensaje clave no es que sean erróneas o inexactas. Lo fundamental es que no podemos dar por hecho su exactitud. La idea de que hay cosas que se quedan «grabadas a fuego» en nuestra mente, y no se olvidan jamás, es errónea. De ahí la necesidad de corroboración externa. Así mismo, las afirmaciones y criterios que refieren y «exigen» cómo debe ser la declaración de las víctimas (seria, no fragmentada, relato íntegro, ausencia de lagunas, claridad expositiva, etc.) contravienen igualmente la evidencia sobre el funcionamiento de nuestros recuerdos, como lo hacen las alusiones a la apreciación de la seguridad con la que

se declara, aspecto este que, salvo excepciones (WIXTED y WELLS, 2017), no puede utilizarse como indicador de exactitud ni de credibilidad.

Una cuestión final respecto a los indicadores verbales. A pesar de su mayor precisión, hay que ser muy cautos cuando se acude a ciertas herramientas (SVA, uno de cuyos elementos constituyentes, pero no el único, es el CBCA¹¹) que se han popularizado entre los profesionales. La forma de entenderlas, su contexto adecuado de utilización, su objetivo y alcance, así como su complejidad pueden estar también desvirtuándose. Y ello cuando se rechaza como instrumento adecuado en el ámbito forense, como cuando se acepta acríticamente sin ser conscientes de sus limitaciones (KÖHNKEN, MANZANERO y SCOTT, 2015; MANZANERO y MUÑOZ, 2011; MASIP y ALONSO, 2011; RAMÍREZ ORTIZ, 2018). O cuando se tratan de aplicar aisladamente alguno de sus criterios.

5. CIENCIA PSICOLÓGICA VS. CIENCIAS FORENSES

Hablábamos más arriba de la diferente consideración de la psicología respecto a otras disciplinas. Retomemos parte de esta idea tomando como punto de partida los factores fundamentales implicados en las condenas erróneas. Tanto las cifras de Proyecto Inocencia (innocenceproject.org) como del Registro de Exoneraciones Americano (Universidad del Estado de Michigan. Facultad de Derecho, <https://www.law.umich.edu/special/exoneration/Pages/mission.aspx>), así como muy diversos trabajos (BENFORADO, 2015; DROR YBUCHT, 2012; GARRETT, 2019; KOEN Y BOWERS, 2018; KUKUCKA, 2018; LUPÁRIA, 2015; NEUFELD, 2005; NEUFELD Y SCHECK, 2010, entre otros) indican que las dos causas principales son los errores en el testimonio de los testigos (69 % de los casos) y la mala aplicación de las ciencias forenses (45%).

Respecto a los testigos y sus declaraciones, no resulta sorprendente desde el punto de vista de la psicología jurídica. La demostración de condenas erróneas por esa causa confirmó lo que la disciplina venía planteando desde hacía mucho tiempo respecto a la falibilidad de los testigos y sus identificaciones (TECHNICAL WORKING GROUP FOR

¹¹ SVA: *Statement Validity Assessment* (Evaluación de la validez de las declaraciones); CBCA: *Criteria-Based Content Analysis* (Análisis de Contenido basado en Criterios). SVA es un protocolo o procedimiento de generación y falsación de hipótesis sobre el origen de una declaración, y con él se pretende recoger y evaluar la información de las declaraciones de las víctimas. Consta de tres elementos: 1) un protocolo semi-estructurado de entrevista para recopilar información de modo no sesgado que se realiza con unos parámetros determinados para no alterar los recuerdos o la declaración de la persona entrevistado y que habrá de transcribirse posteriormente; 2) el CBCA, una lista de 19 criterios con el que se evalúa la entrevista; y 3) la denominada lista de comprobación de la validez. A pesar de la complejidad y de los distintos elementos del protocolo, prácticamente ha sido el CBCA el que se ha popularizado como si se tratase simplemente de «chequear» la aparición o no de unos criterios para inferir la credibilidad. Una revisión de esta herramienta así como de sus limitaciones y advertencias de usos inadecuados pueden encontrarse en MASIP y ALONSO (2011) en M LAMEIRAS (Coords.), *Violencia de género*. Tirant o en KÖHNKEN, MANZANERO y SCOTT (2015), en *Anuario de Psicología Jurídica*, 25: 13-19.

EYEWITNESS EVIDENCE, 1999; WELLS et al, 2000, 2020). Pero no fue hasta que las pruebas de ADN lo permitieron y se constató que tal falibilidad causaba decisiones erróneas, cuando se concedió crédito a los hallazgos y demandas de la ciencia psicológica.

Más sorprendente puede resultar la segunda de las causas, la que tiene que ver con las ciencias forenses, consideradas la máxima expresión de la ciencia en los tribunales y prácticamente sin posibilidad de error. Y sorprende acaso porque igual se está siendo vulnerable al denominado «efecto CSI» o, al menos, a alguna de sus facetas. Supuestamente tal efecto influye aumentando las expectativas, por ejemplo, de los jurados debido a las avanzadas técnicas y pruebas científicas que aparecen en los variados CSI del mercado. Fruto de tales expectativas, abogados, jueces, fiscales o periodistas reclamaban que ser asiduo a estos programas genera absoluciones erróneas cuando en los procesos reales no aparecen suficientes o determinadas pruebas forenses, o son menos impactantes que las de tales series. Puede resultar ilustrativo que los resultados del estudio de creencias referido previamente (HERRERO et al, 2016) muestra que en torno al 50% del alumnado de derecho y del grupo de profesionales juristas, cerca de 300 encuestados, considera que los errores en las ciencias forenses (análisis de huellas, marcas dentales, de calzado o sangre, análisis capilares, etc.) son una de las principales causas por las que se puede condenar a personas inocentes. La cifra en una muestra de 140 policías disminuye hasta el 28%. Así mismo, entre los primeros, el 45% cree en la existencia del Efecto CSI, mientras que entre los segundos asciende al 56%. Por sugerente que pudiera parecer el referido efecto, y a pesar de la controversia generada, parece estar de nuevo ante una ilusión, una falsa creencia originada de la observación anecdótica y en absoluto sistemática realizada por periodistas, fiscales y otros actores legales para explicar las absoluciones que consideran desconcertantes (COLE y DIOSO-VILLA, 2011; COURT REVIEW, 2011; KLENTZ et al., 2020; TYLER, 2006; SCHANZ y SALFATI, 2016; SMITH et al., 2011).

Volvamos en este punto a recordar el caso Richardson (caso 5 de la introducción) y añadamos tres más: el primero, el de S. Cowans Liberado después de pasar 6 años en prisión: Condenado erróneamente por disparar a un policía. Fue el primer caso de exoneración posterior en el que estuvieron implicadas las huellas que aparecieron en un vaso utilizado supuestamente por Cowans, además de la declaración de dos testigos. Pero ¿por qué sus huellas «no match» con las del vaso hasta después de ser identificado como sospechoso y detenido? El segundo, el del español O.S.F. absuelto después de pasar 600 días en prisión en Italia condenado por narcotráfico y como jefe de una banda mafiosa. La policía española consideró que podía ser un caso de suplantación de identidad. La prueba crítica resultó ser la identificación por voz realizada a partir de unas escuchas. Y, por último, el ya famoso caso de Brandon Mayfield (KASSIN, DROR y KUKUCKA, 2013; SIMONCELLI, 2016) que nos sitúa en el atentado del 11-M, de marzo de 2004, en el que explotaron varias bombas en trenes de Madrid causando cerca de 200 muertos e innumerables heridos. Dos de las huellas recuperadas de alguna de las mochilas que aparecieron en la investigación fueron

compartidas por la Policía española con el FBI, que las introdujo en el sistema automatizado de huellas (IAFIS). El cuarto en la lista de potenciales candidatos generados por el sistema fue Mayfield. Un abogado de Oregon, americano musulmán, casado con una egipcia y que había representado en alguna ocasión a personas vinculadas al terrorismo. Siguiendo supuestamente el protocolo, varios analistas independientes identificaron la huella.

La policía española no estuvo nunca de acuerdo con tal identificación y, de hecho, denunció presiones por parte del FBI. Este, sin embargo, insistió en la certeza del hallazgo. No había ninguna evidencia más que vinculara al abogado con los atentados de Madrid, a pesar de las escuchas, seguimientos y registros que se realizaron. Mayfield fue detenido a principios de mayo y pasó en prisión tres semanas hasta que, finalmente, la policía española asoció la huella al argelino O. Daoud. El error generó una investigación interna en el FBI y el correspondiente informe por parte de la Oficina del Inspector General de más de 300 páginas en las que aparecía *el sesgo confirmatorio* entre las causas que contribuyeron a la identificación (OIG, 2006; 2011).

¿Qué sugieren de nuevo estos casos, además de la falibilidad de pruebas tradicionalmente consideradas infalibles? ¿Simplemente «error humano»? Si así fuera, muy probablemente estuviéramos siendo vulnerables al *Error Fundamental de Atribución* (Ross, 1977) que estudiamos los psicólogos sociales. Volveríamos una vez más a colocar todo el peso en los factores personales subestimando sobremedida los de la situación. Y del mismo modo que ocurre con el testimonio y las variables que afectan a la identificación, a las declaraciones, no podemos dejar de considerar «las variables del sistema» (WELLS, 2020; WELLS y LOFTUS, 2003), aquellas sobre las que el sistema de justicia puede tener control y que son fundamentales no solo para evitar los errores, sino para guiar una práctica de calidad. Por lo tanto ¿Qué más podemos aprender? (LOFTUS, 2004). Las situaciones y los contextos particulares son claves para interpretar y valorar las distintas piezas de evidencia, las fuentes-medios de prueba. Los sistemas y sus procedimientos son primordiales. Así mismo, es claro que no suele haber una única fuente de error aislada si tenemos en cuenta toda una serie de decisiones humanas y de procedimientos respecto a la investigación, a la obtención de la prueba, a los testigos y su elevada confianza, y no únicamente los sesgos de los investigadores, de los analistas, y de todos los actores implicados. Además, podríamos reparar, cosa que no suele hacerse, en el efecto potencialmente pernicioso de las interacciones entre piezas de evidencia aparentemente diferenciadas pudiendo producirse el denominado «*bias snowball effect*» (DROR, 2012). Dicho efecto pone de manifiesto cómo el contexto (o informaciones de diferentes líneas de evidencia o informaciones irrelevantes o sesgadas de un caso) puede afectar a los procesos de decisión de analistas, de investigadores y, posteriormente si no son detectados, a las decisiones y prácticas de los operadores jurídicos. Ocurre cuando la información que se va obteniendo de un caso es percibida como siendo independiente y corroborativa en una misma dirección, cuando en realidad puede haber sido o estado contaminada por fuentes similares de información contextual (KASSIN et al, 2013; DROR et al., 2013; DROR, 2012; 2018; EDMOND et al., 2015, 2017).

Los casos referidos y, sobre todo, los análisis y controversias con ellos generados contribuyeron a que en la primera década de este siglo se comenzara a gestar un importante cambio de paradigma en las ciencias forenses que aún hoy continúa en proceso. Veamos muy sucintamente algunos de los factores más importantes que contribuyeron al cambio.

Más allá de todas las críticas recibidas y de los análisis que han mostrado facetas problemáticas de la trilogía *Daubert, Joiner, Kumho*, no puede negarse que constituyó un hito con el que aumentaba el grado de exigencia y de control en la calidad de la prueba, del conocimiento científico, técnico y/o artístico que se admitía en los tribunales. Se requería un mayor o diferente escrutinio de la fiabilidad y validez que el establecido en la sentencia previa *Frye*. Por otra parte, no puede olvidarse el auge, los avances y mejoras que tuvieron lugar en las pruebas de ADN en esos momentos, que permitieron, a su vez, realizar o repetir en su caso los análisis en personas condenadas que, como se demostró, resultaron ser inocentes. Las cifras que se iban conociendo, así como distintas decisiones controvertidas y algunas publicaciones fueron los detonantes perfectos para que se gestara el cambio. Entre tales decisiones resultó impactante la del juez Pollak que no admitió como prueba de identificación las huellas dactilares puesto que no reunía los criterios *Daubert* (*Science*, 2002, 295), aunque posteriormente rectificó su decisión. Así mismo, un editorial de *Science* (2003, vol. 302) rezaba «Forensic Science: Oxymoron?». FAIGMAN (2002) se preguntaba «Is science different for lawyers?» o hablaba de los jueces como científicos «amateurs» (FAIGMAN, 2006).

Pero sin duda alguna, hubo otros dos factores determinantes. El primero, los informes internacionales así como los trabajos de académicos y, en menor medida, profesionales que exigían un cambio en las ciencias forenses (SAKS y FAIGMAN, 2008; SAKS y KOEHLER, 2005). La evaluación de estas disciplinas y el informe de la National Academy of Science (NAS, 2009) al que dio lugar, *Strengthening Forensic Science in the United States: A Path Forward*, resultó implacable en su diagnóstico: excepto algún tipo de análisis concreto de ADN, el resto de las pruebas forenses carecían de validez científica. Era urgente que las áreas forenses trabajaran para hacer más científicos sus fundamentos así como los de los análisis a los que daban lugar. También en Europa comenzaron los cambios (CHAMPOD, C. y VUILLE; LAW COMMISSION, 2009, 2011). Todo ello generó la necesidad de abandonar el conocido paradigma de la identificación/exclusión, imperante hasta esos momentos, e ir hacia otros más acordes a los conocimientos científicos (GASCÓN y LUCENA, 2010; GASCÓN, LUCENA y GONZÁLEZ, 2010; NEUFELD y SCHECK, 2010; ROYAL STATISTICAL SOCIETY, 2017; SAKS, 2010). El segundo, toda una línea de investigación sobre lo que podríamos llamar el factor humano en las ciencias forenses y que abordamos seguidamente (EDMOND et al., 2019).

5.1. Sobre el cambio de paradigma en las ciencias forenses: la contribución de la psicología

Analizar el papel de psicología en este gran cambio supone reparar en los procesos cognitivos y sociales implicados en la obtención, análisis e interpretación de las informaciones y elementos de prueba efectuadas por los expertos (analistas, investigadores, científicos forenses, policía, psicólogos o psiquiatras en el caso de la evaluación psicológica forense, etc.). Dos aspectos, al menos, tenemos que contemplar. Por una parte, cómo trabaja y procesa información nuestro sistema cognitivo. Y por otra, lo que requieren las tareas de los expertos analistas, los que se encargan de examinar y procesar los materiales de los que se trate. Podría añadirse un tercer elemento que puede jugar algún papel en el tema que nos ocupa: la tecnología, pero siempre en interacción con el ser humano. Y acaso este elemento forma parte de las tareas y requerimientos cognitivos que demandan los análisis correspondientes.

Respecto a la primera cuestión, y de modo muy general, la investigación pone de manifiesto que las personas no somos perceptores pasivos, ni objetivos, de la información que buscamos o recibimos. Tal información interactúa con lo que «ya llevamos en la mochila». Así, las percepciones, las cogniciones que tienen los sujetos se ven muy influidas por la experiencia, las motivaciones, las expectativas, los estereotipos o por el contexto. Esto va a ser determinante para entender con qué facilidad en todos esos procesos pueden aparecer los sesgos (cognitivos) (CHARMAN et al, 2019; DROR, 2012). Respecto a la segunda, y en contra de los que nos muestran las series televisivas, se necesita mucho más que una máquina o una base de datos muy avanzada para aproximarse a la tarea de analizar huellas, marcas de mordeduras, trazas de sangre, cabellos, balas, etc. Un aspecto común en tales análisis tiene que ver con la percepción y comparación de patrones visuales. Pero en cualquier caso, un elemento fundamental es la dosis de interpretación que requieren las tareas. Y es ahí dónde, dependiendo de los procedimientos seguidos en la obtención, análisis e interpretación de información, dependiendo del contexto, el analista puede ser vulnerable a los sesgos. Pensemos en lo siguiente. ¿Qué ocurriría si analizando unas trazas determinadas al sujeto le llega información sobre algún aspecto del caso al que corresponde el material? ¿Afectaría a sus interpretaciones? O por ejemplo, reparemos en lo que ocurre cuando el analista que va a verificar (siempre hay verificación y a veces más de una) un resultado, una interpretación previa de un compañero, llega a conocer la decisión a la que aquel llegó, ¿afectaría a su conclusión? Muy probablemente alguien piense que si realmente es un experto no debiera afectarle. O que si lo hiciera, se debió a la falta de competencia o a alguna otra característica de la persona (Recordemos lo que en páginas precedentes definíamos como el Error fundamental de Atribución, Ross, 1977). Eso nos lleva a la siguiente cuestión.

La investigación indica que precisamente por ser experto, y paradójicamente, se ponen en marcha una serie de procesos cognitivos diferentes respecto a los no expertos que, siendo ventajosos en muchos aspectos, conducen a la mayor vulnerabilidad al sesgo y las interpretaciones inadecuadas. Ser un experto, en cualquier ámbito, no

supone ser más inteligente, trabajar más rápido o más eficiente, sino aproximarse a las cosas de forma diferente (DROR, 2011; MACLEAN, 2016). De este modo, es característico de los expertos, por ejemplo, prestar atención selectiva, procesar la información más automáticamente, fragmentar información o utilizar lo que denominamos un procesamiento «top-down», esto es, ir de la teoría a los datos y no a la inversa. Todas estas características facilitan la ejecución rápida y eficiente, pero también que se pierda información, se atienda a las que no son relevantes, que los esquemas condicionen, de forma inflexible, ir en una única dirección evitando buscar, por ejemplo, teorías o hipótesis alternativas. Todo ello puede generar la denominada «visión túnel» (FINDLEY y SCOTT, 2006). Por lo expuesto, se infiere que cuando en este contexto se habla de sesgos o errores (cognitivos), hay que subrayar que son no intencionales. Por lo tanto, hay que distinguirlos de la interpretación que en ocasiones se otorga al concepto de sesgo en el derecho como unido a la parcialidad que puede producirse en los actores del proceso, por ejemplo, por intereses variados.

El problema más relevante en el origen de los posibles sesgos en el ámbito forense se deriva de la influencia del contexto en las tareas y procesos requeridos para los análisis, interpretaciones y conclusiones forenses. Ocurre cuando un juicio o decisión está influido por información irrelevante o inapropiada que el analista o el que toma la decisión obtiene de su contexto circundante (NATIONAL RESEARCH COUNCIL, 2009). Algunas guías (NATIONAL COMMISSION OF FORENSIC SCIENCE, 2015) han ido definiendo qué se considera información irrelevante para las situaciones y análisis concretos¹². Los efectos del contexto se convirtieron en fuente de preocupación cuando distintos estudios empíricos comenzaron a poner de relieve que, por ejemplo, era menos probable que los examinadores de huellas comunicaran un «ajuste» entre la huella recogida en la escena del delito y la del sospechoso cuando se les decía que el sospechoso tenía una coartada sólida (DROR y CHARLTON, 2006; DROR y ROSENTHAL, 2008; NIST, EXPERT WORKING GROUP ON HUMAN FACTORS IN LATENT PRINT ANALYSIS, 2012). Si bien ha sido el examen de las huellas dactilares el que ha suscitado mayor atención en las investigaciones (STEVENAGEM y BENNETT, 2017), en los últimos años se han publicado un buen número de trabajos sobre el examen de escritura en documentos (OSBORNE, BIRD y STOEL, 2019; NIST, EXPERT WORKING GROUP FOR HUMAN FACTORS IN HANDWRITING EXAMINATION, 2020), sobre

¹² Por ejemplo, y siguiendo la referida guía, en el caso del examen de huellas dactilares latentes que valora la proposición de que la huella sobre alguna superficie u objeto en la escena del delito procede de un sospechoso X, se podría considerar información irrelevante el conocer los antecedentes penales del sospechoso, o que el sospechoso se inculcó o atribuyó el delito, o que algún testigo lo identificó en una rueda de reconocimiento, o que estaba implicado en otra de las piezas de evidencia física recogida de la escena del delito (ej., ADN), o que otro examinador identificó al sospechoso como fuente de una huella encontrada en un objeto o elemento diferente al examinado por el primer analista. No obstante, hay que subrayar que la información que es irrelevante para tareas de análisis, como la comparación de huellas, puede no serlo cuando se realizan otras tareas (ej., determinar qué se recoge de la escena de un delito, o analistas cuya labor de consultar e intercambiar información con policías o abogados para ayudarles a entender el panorama que emerge del examen de múltiples elementos de prueba. Por lo tanto, qué sea o no relevante para una tarea siempre dependerá de la naturaleza de esta.

la información digital de dispositivos variados (SUNDE y DROR, 2019) o sobre antropología forense. Respecto a los sesgos concretos, uno de los que ha recibido mayor atención es el denominado *sesgo confirmatorio* (NICKERSON, 1998), cuyas aplicaciones a múltiples contextos están muy consolidadas en psicología social. En su aplicación a las ciencias forenses (FORENSIC SCIENCE REGULATOR, 2015, 2020; KASSIN et al., 2013; KUKUCKA, 2018; KUKUCKA, HALEY Y KASSIN, 2020) se ha considerado que engloba la clase de efectos a través de los que las creencias preexistentes de un individuo, las expectativas, las motivaciones o el contexto situacional influye la recogida, percepción e interpretación de evidencia durante el curso de una investigación penal (KASSIN et al., 2013). Así mismo, lo expuesto sobre los efectos del contexto o el sesgo confirmatorio es igualmente aplicable a otras actividades forenses, las de los psicólogos y profesionales de la salud mental (NEAL, 2017; NEAL, HIGHT, HOWATT Y HAMZA, 2018), ámbito en el que ha alcanzado gran expansión. Además, sean de una especialidad forense u otra, los propios expertos no reconocen fácilmente la posibilidad de que sus análisis y procesos de pensamiento, así como sus expectativas y otras informaciones que no tienen que ver directamente con el material que examinan condicionen las decisiones que alcanzan o las interpretaciones que realizan. Esto es, estamos de nuevo ante el «*blind spot bias*» (PRONIN, LIN Y ROSS, 2002).

Además de los sesgos mencionados, se ha realizado un estudio más detallado de cuáles son sus fuentes. Así, DROR (2017, 2020), después de años de investigación, propone una taxonomía que permitiría identificar diferentes deficiencias fruto de la contaminación cognitiva de las observaciones y conclusiones forenses. Distingue siete fuentes, aunque en el trabajo de 2020 añade una más. Conceptualizada en forma de pirámide, la estructura de la taxonomía tiene en cuenta, en su base, la estructura cognitiva humana, para pasar en los niveles medios a fuentes que se asocian en mayor medida con el ambiente, la cultura y la experiencia, mientras que en los más altos se relacionan con aspectos del caso específico. Así mismo, permite, junto a toda la investigación consolidada en psicología social y cognitiva, sugerir pautas de actuación para atenuar los efectos de los sesgos o eliminarlos, tarea no sencilla. Para ello, es importante atender a los procedimientos y a los factores de las organizaciones implicadas, pues sin su modificación difícilmente se avanzará en la dirección adecuada. En este camino, puede resultar igualmente útil el trabajo de KOEHLER (2020) pues propone veinte amenazas a la validez de determinadas inferencias forenses, agrupadas también en categorías, alguna de las cuales se superpone a las fuentes de sesgo sugeridas previamente. Todas estas cuestiones tienen implicaciones sobre la interpretación o a la determinación de unos hechos, así como en la valoración de la prueba pericial.

6. CONCLUSIONES

Comenzábamos el trabajo señalando la necesidad de que el Derecho, el proceso y los operadores jurídicos acudan al conocimiento aportado por la psicología, ciencia sobre la que todavía existe cierto escepticismo y a la que se otorga diferente estatus

respecto de otras disciplinas. Sin embargo, hemos constatado que los conocimientos de la psicología social y cognitiva no solo han permitido explicar los sesgos de ciencias supuestamente infalibles, como las forenses, sino que han propuesto procedimientos para atenuarlos e, incluso, eliminarlos. Ello está contribuyendo a un nuevo enfoque de las prácticas forenses. También hemos mostrado los efectos perniciosos que generan las creencias de sentido común y el ignorar los estudios empíricos realizados en contextos jurídicamente relevantes.

La constatación de que se utilizan creencias de sentido común o que aparecen sesgos entre los operadores jurídicos, como ocurre en todos los seres humanos, incluidos los psicólogos jurídicos y forenses, además de ser una obviedad, es insuficiente. Por otra parte los profesionales juristas en ocasiones se quejan de la poca utilidad de la ciencia cuando tienen que tomar una decisión concreta en un tiempo limitado. Este potencial desencuentro no es únicamente responsabilidad de los juristas. La psicología, especialmente de la Jurídica, no debería ser meramente una disciplina interesante y hasta entretenida, sino útil. Así, sería deseable poder instruir a los operadores jurídicos sobre cómo abordar más adecuadamente ciertos asuntos del comportamiento humano y cómo pueden, por ejemplo, evaluar la información que los peritos introducen en los tribunales, o que los expertos aportan más allá del foro. Necesitamos, por lo tanto, crear contextos de trabajo e interacción más efectivos que no se reduzcan a la sala de justicia y a la realización de peritajes. Ello puede que requiera también acortar las distancias entre la academia/investigación «psicojurídica» y su práctica forense.

Todo ello requiere reparar en algún elemento relacionado con la formación de juristas y psicólogos jurídicos. Respecto a la de los primeros, sería aconsejable que se inicie en etapas tempranas, esto es, cuando están en las facultades de derecho. La formación no debe concebirse únicamente como una acumulación de informaciones sobre temas o contenidos puntuales, sino como un proceso que les permita socializarse también en una cultura del razonamiento científico y adquirir una perspectiva más transversal. Para ello no es imprescindible acudir a los conocimientos de la psicología. Otros ámbitos disciplinares pueden ser igual o más provechosos. Asimismo, los psicólogos jurídicos debiéramos elevar nuestra formación respecto a los fines y métodos del derecho, así como nuestros conocimientos sobre la esencia y fundamentos de nuestro sistema jurídico y sus normas.

7. BIBLIOGRAFÍA

- ANDRÉS IBÁÑEZ, P., 2003: «Sobre el valor de la intermediación. (Una aproximación crítica)», en *Jueces para la Democracia. Información y Debate*, 46: 57-66.
- , 2017: «Sobre la pobreza cultural de una práctica (judicial) sin teoría», en *Derecho PUCP*, 79: 111-126.
- , 2020: «En materia de prueba: sobre algunos cuestionables tópicos jurisprudenciales», en *Quaestio facti. Revista Internacional sobre Razonamiento Probatorio*, 1: 75-102.

- ANDRES-PUEYO, A., 2012: «La peligrosidad criminal: análisis crítico de un concepto polisémico», en E. DEMETRIO-CRESPO (Ed.), *Neurociencias y Derecho penal. Nuevas perspectivas en el ámbito de la culpabilidad y tratamiento jurídico-penal del concepto de peligrosidad* (pp. 557–580). Buenos Aires: Bdf & Edisofer.
- ANDRES-PUEYO, A. y ARBACH, K., 2014: «Peligrosidad y valoración del riesgo de violencia en contextos forenses», en E. Garcia (Ed.), *Psicopatología Forense: comportamiento humano y tribunales de justicia* (pp. 1–37). Bogotá, Colombia: El Manual Moderno.
- BARR, M. y BEN-SHAKHAR, G., 2019: *Science, Pseudoscience, non-sense, and critical thinking: why the differences matter*. London: Routledge.
- BENFORADO, A., 2015: *Unfair. The new science of criminal injustice*. New York: Crown Pub Inc
- BENTON, T. R., ROSS, D. F., BRADSHAW, E., THOMAS, W. N., y BRADSHAW, G. S., 2006: «Eyewitness memory is still not common sense: comparing jurors, Judges and Law Enforcement to eyewitness experts», en *Applied Cognitive Psychology*, 20, 115–129.
- BOGARD, G., MEIJER, E. H., VRIJ, A. y MERCKELBACH, H., 2016: «Strong, but wrong: lay people's and police officers' beliefs about verbal and nonverbal cues to deception», en *PLoS ONE*, 11: e0156615. doi: 10.1371/journal.pone.0156615
- BOND, C. F., JR. y DEPAULO, B. M., 2006: «Accuracy of deception judgments», en *Personality and Social Psychology Review*, 10: 214-234.
- BRANK, E., 2019: *Psychology of Family Law*. N. Y: New York University Press.
- BROOMELL, S. B. y KANE, P. B., 2017: «Public perception and communication of scientific uncertainty», en *Journal of Experimental Psychology: General*, 146: 286-304.
- BROWNSSELL, A. y BULL, R., 2011: «Magistrates' beliefs concerning verbal and non-verbal behaviours as indicators of deception», en *The European Journal of Psychology Applied to Legal Context*, 3: 29-46.
- BURGOON, J. K., 2018: «Microexpressions are not the best way to catch a liar», en *Frontiers in Psychology*, 9: 1672.
- CHAMPOD, C. y VUILLE, J., 2011: «Scientific evidence in Europa. Admissibility, evaluation and equality of arms», en *International Commentary on Evidence*, 9: <https://doi.org/10.2202/1554-4567.1123>
- CHARMAN, S., DOUGLASS, A. B. y MOOK, A., 2019: «Cognitive bias in legal decision making», en N. BREWER A. B. DOUGLAS (eds.), *Psychological Science and Law* (pp. 30). Nueva York: Guilford Press.
- COLE, S. A. y BARNO, M., 2020: «Probabilistic reporting in criminal cases in the United States: a baseline study», en *Science and Justice*, 60, 406-414.
- COLE, S. A. y DIOSO-VILLA, R., 2011: «Should judges worry about 'CSI Effect'?, en *Court Review*, 47: 16-27.
- COOPER, P., DANDO, C., ORMEROD, T., MATTISON, M., MARCHANT, R., MILNE, R. y BULL, R., 2018: «One step forward and two steps back? The "20 Principles" for questioning vulnerable witnesses and the lack of an evidence-based approach», en *The International Journal of Evidence and Proof*, 22, 392-410.
- CROZIER, W., KUKUCKA, J. y GARRET, B. L., 2020: «Juror appraisals of forensic evidence: effects of blind proficiency and cross-examination», en *Forensic Science International*, 315: 110433.
- CUTLER, B. L. y KOVERA, M. B., 2011: «Expert psychological testimony», en *Current Directions in Psychological Science*, 20: 53-57.
- DEMAINE, L., 2008: «In search of an anti-elephant: confronting the human inability to forget inadmissible evidence», en *George Mason Law Review*, 16: 99-140.
- , 2012: «Realizing the potential of instructions to disregard», en L. Nadel y W. P. Sinnott-Amstrong (Ed.), *Memory and Law* (pp. 185-212), Oxford: Oxford University Press.
- DENAULT, V. y DUNBAR, N., 2017: «Nonverbal communication in courtrooms: Scientific assessments or modern trials by ordeal?», en *The Advocates' Quarterly*, 47: 280-308.
- , 2019: «Credibility assessment and deception detection in courtrooms: Hazards and challenges for scholars and legal practitioners», en T. Docan-Morgan (Ed.), *The Palgrave handbook of deceptive communication* (pp. 915-936). Basingstoke, United Kingdom: Palgrave Macmillan.
- DENAULT, V., 2020: «Misconceptions about nonverbal cues to deception: a covert threat to the justice system», en *Frontiers in Psychology*, 573460. doi: 10.3389/fpsyg.2020.573460.

- DENAULT, V., PLUSQUELLEC, P., JUPE, L. M., ST-YVES, M., DUNBAR, N. E., HARTWIG, M., ... VAN KOPPEN, P. J., 2020: «The analysis of nonverbal communication: The dangers of pseudoscience in security and justice contexts», en *Anuario de Psicología Jurídica*, 30: 1-12.
- DEPAULO, B. M., LINDSAY, J. J., MALONE, B. E., MUHLENBRUCK, L., CHARLTON, K. y COOPER, H., 2003: «Cues to deception», en *Psychological Bulletin*, 129, 74-118.
- DESMARAIS, S. L. y READ, J. D., 2011: «After 30 years, what do we know about what jurors know? A metanalytic review of lay knowledge regarding eyewitness factors», en *Law and Human Behavior*, 35, 200-210.
- DESPADOVA, N. M., KUKUCKA, J. y HILEY, A., 2020: «Can defense attorneys detect forensic confirmation bias? », en *Zeitschrift für Psychologie*, 228: 216-220.
- DICKINSON, J. J., COMPO, N. S., CAROL, R. N., SCHWART, B. L y McCAULEY, M. R., 2019: *Evidence-based investigative interview: applying cognitive principles*. New York: Routledge.
- DODIER, O. y TOMAS, F., 2019: «When psychological science fails to be heard: the lack of evidence-based arguments in a ministerial report on child sexual abuse», en *Psychiatry, Psychology and Law*, 3: 385-395.
- DROR, I. E. (2012). Cognitive bias in forensic science. *Science & Technology 2012 Yearbook* (pp. 43-45). McGraw-Hill.
- DROR, I. y BUCHT, R. 2012: «Psychological Perspectives on Problems with Forensic Science Evidence», en B. Cutler (Ed.), *Conviction of the Innocent: Lessons From Psychological Research* (pp.257-276). American Psychological Association.
- DROR, I. ROSENTHAL, R., 2008: «Meta-analytically quantifying the reliability and bias ability of forensic experts», en *Journal of Forensic Sciences*, 53: 900-903.
- DROR, I., 2011: «The paradox of human expertise», en Narinder Kapur (Ed.), *The Paradoxical Brain* (pp.177-188). Cambridge University Press.
- , 2013: «The ambition to be scientific: human expert performance and objectivity», en *Science and Justice*, 53: 81-82.
- , 2017: «Human expert performance in forensic decision making: Seven different sources of bias», en *Australian Journal of Forensic Sciences*, 49: 541-547.
- , 2018: «Better science for better justice: a proposal for joint experts», en *Science and Justice*, 58:465-466.
- DROR, I., 2018: «Biases in forensic experts», en *Science*, 360 (6386): 243.
- DROR, I., 2020: «Cognitive and human factors in expert decision making: six fallacies and eight sources of bias», en *Analytical Chemistry*, 92: 7998-8004.
- DROR, I., CHAMPOD, C., LANGENBURG, G., CHARLTON, D, HUNT, H. y ROSENTHAL, R., 2011: «Cognitive issues in fingerprint analysis: Inter- and intra-expert consistency and the effect of a 'target' comparison». *Forensic Science International*, 208: 10–17.
- DROR, I., McCORMACK, B. M. EPSTEIN, J., 2015: «Cognitive bias and its impact on expert witnesses and the court», en *The Judges Journal*, 54, 4
- DROR, I., THOMPSON, W. C., MEISSNER, C. A., KORNFIELD, I., KRANE, D., SAKS, M. y RISINGER, D. M., 2015: «Context management toolbox: A linear sequential unmasking (LSU) approach for minimizing cognitive bias in forensic decision making», en *Journal of Forensic Sciences*, 60: 1111-1112.
- EDMOND, G., 2012: «Is reliability sufficient? The Law Commission and expert evidence in international and interdisciplinary perspective (Part 1)», en *The International Journal of Evidence and Proof*, 16: 30-65. doi:10.1350/ijep.2012.16.1.391.
- EDMOND, G. y HAMER, D., 2012: «Evidence Law», en P. Cane y H. M. Kritzer (Eds.), *The Oxford Handbook of Empirical Legal Research*. Oxford University Press. doi: 10.1093/oxfordhb/9780199542475.013.0028.
- EDMOND, G., TOWLER, A., GROWNS, B., RIBEIRO, G...MARTIRE, K., 2017: «Thinking forensics: Cognitive science for forensic practitioners», en *Science and Justice*, 57 (2):144-154. doi: 10.1016/j.scijus.2016.11.005.
- EKMAN GROUP (2011). F.A.C.E. training: Interactive training by Dr. Paul Ekman. Recuperado de <https://www.paulekman.com/about/about-paul-ekman-group-llc/>

- EKMAN, P., 2009: *Telling lies: Clues to deceit in the marketplace, politics, and marriage*. New York, NY: Norton
- ENGLISH, B. y MUSSWEILER, T., 2001: «Sentencing Under Uncertainty: Anchoring Effects in the Courtroom», en *Journal of Applied Social Psychology*, 31: 1535-1551.
- FAIGMAN, D. L. y MONAHAN, J., 2009: «Standars of legal admissibility and their implications for psychological science», en J. L. SKEEM, K. S. DOUGLAS y S. O. LILIENTFELD (Eds.), *Psychological Science in the Courtroom* (pp. 3-25). Nueva York: Gilford Press.
- FAIGMAN, D. L., 2002: «Is science different for lawyers?», en *Science*, 297:339-340.
- , 2006: «Judges as “amateur scientist”», en *Boston University Law review*, 86: 1207-1225.
- , 2013: «The *Daubert* revolution and the birth of modernity: managing scientific evidence in the age of science», en *University of California, Davis*, 46: 893-930.
- FAIGMAN, D. L., MONAHAN, J. y SLOBOGIN, C. 2014: «Group to Individual (G2i) Inference in Scientific Expert Testimony», en *The University of Chicago Law Review*, 81: 417-480.
- FAIGMAN, D. L., SLOBOGIN, C. y MONAHAN, J., 2016: «Gatekeeping Science: Using the Structure of Scientific Research to Distinguish Between Admissibility and Weight in Expert Testimony», en *Northwestern University Law Review*, 110: 859-904.
- FEDERAL JUDICIAL CENTER, 2011: *Reference manual on scientific evidence (Third Edition)*. Washington, D.C.: The National Academies Press.
- FEIGENSON, N., 2010: «Visual evidence», en *Psychonomic Bulletin and Review*, 17: 149-154.
- FEIGENSON, N. y SPIESEL, C., 2009: *Law on display: The digital transformation of legal persuasion and judgment*. New York, NY: Oxford University Press.
- FENTON, N., 2011: «Improve statistics in court», en *Nature*, 479: 36-37.
- FERRER, J., 2005: *Prueba y verdad en el Derecho*, (2ªed). Barcelona: Marcial Pons
- , 2007: *La valoración racional de la prueba*, Barcelona: Marcial Pons.
- , 2017: «El control de la valoración de la prueba en segunda instancia», en *Journal for Constitutional Theory and Philosophy of Law*, 33: 107-126.
- FINDLEY, K. y SCOTT, M. S., 2006: «The multiple dimensions of tunnel vision in criminal cases», en *Wisconsin Law Review*, 291: 291-397.
- FISHER, C. E., FAIGMAN, D. L. y APPELBAUM, P. S., 2015: «Toward a jurisprudence of psychiatric evidence: examining the challenges of reasoning from group data in psychiatry to individual decisions in the law», en *University of Miami Law Review*, 69: 685-753.
- FORENSIC SCIENCE REGULATOR (2015). *Guidance. Cognitive Bias Effects. Relevant to Forensic Science Examinations*. FSR 1-G-217. Issue 1.
- , (2020). *Guidance. Cognitive Bias Effects. Relevant to Forensic Science Examinations*. FSR 1-G-217. Issue 2.
- FOX, C. R. y BIRKE, R., 2002: «Forecasting Trial Outcomes: Lawyers Assign Higher Probability to Possibilities That Are Described in Greater Detail» en *Law and Human Behavior*, 26: 159- 173.
- FREND, S. J., NICHOLS, R. M. y LOFTUS, E., 2011: «Current issues and advances in misinformation research», en *Current Directions in Psychological Science*, 20: 20-23.
- FUNDACIÓN ESPAÑOLA PARA LA CIENCIA LA TECNOLOGÍA (FECYT), 2019: *Percepción social de la Ciencia y la Tecnología 2018*. Madrid. FECYT. Ministerio de Ciencia e Innovación.
- GARCÍA-JIMÉNEZ, M., DURFEE, A., CALA-CARRILLO, M. J. Y , y TRIGO, M. E., 2020: «Psychosocial separation and women’s disengagement from prosecutions against abusive intimate partners in Spain», en *Journal of Interpersonal Violence*, doi: 10.1177/0886260520984424.
- GARRET, B. L., 2020: «Wrongful convictions», en *Annual Review of Criminology*, 3: 245-259.
- GARRETT, B. L. y MITCHELL, G., 2016: «Forensics and Fallibility: Comparing the Views of Lawyers and Judges», en *West Virginia Law Review*; Virginia Public Law and Legal Theory Research Paper No. 2016-63; *Duke Law School Public Law & Legal Theory Series*. Disponible en SSRN: <https://ssrn.com/abstract=2865246>
- GARRIDO, E. y HERRERO, C., 1997: «Influence of the prosecutor’s plea on the judge’s sentencing in sexual crimes: Hypothesis of the theory of anchoring by Tvesky and Kahneman», en S. Redondo, V.

- Garrido, J. Perez, & R. Barberet (Eds.), *Advances in psychology and law: International contributions* (pp. 215-226). Berlin, Germany: Walter de Gruyter.
- , 2006: «Relaciones entre la psicología y la ley», en E. Garrido, J. Masip y C. Herrero (Coord.), *Psicología Jurídica* (pp. 3-42), Madrid: Pearson.
- GASCÓN, M. y LUCENA, J. J., 2010: «Pruebas científicas. La necesidad de un cambio de paradigma», en *Jueces para la Democracia*, 69: 95-106.
- GASCÓN, M., 2004 (2ª ed): *Los hechos en el derecho. Bases argumentales de la prueba*. Madrid: Marcial Pons.
- , 2013: «Prueba científica. Un mapa de retos», en C. Vázquez (Ed.), *Estándares de prueba y prueba científica. Ensayos de epistemología jurídica* (pp. 203-213). Madrid: Marcial Pons.
- GASCÓN, M., LUCENA, J. J. y GONZÁLEZ, J., 2010: «Razones científico-jurídicas para valorar la prueba científica: una argumentación multidisciplinar», en *Diario La Ley*, 7481, 4 de octubre.
- GATOWSKI, S. I., DOBBIN, S. A., RICHARDSON, J. T., GINSBURG, G. P., MERLINO, M. L. DAHIR, V., 2001: «Asking the gatekeepers: A national survey of judges on judging expert evidence in a post-Daubert world», en *Law and Human Behavior*, 25: 433-458.
- GIGERENZER, G., 2015: «Heurística inteligente», en J. Brockman (Ed), *Las mejores decisiones*, (pp.43-58). Barcelona: Crítica.
- GÓMEZ HERMOSO, M. R., MUÑOZ VICENTE, J. M., VÁZQUEZ MEZQUITA, B., GÓMEZ MARTÍN, R. y MATEOS DE LA CALLE, N., 2012: *Guía de buenas prácticas para la Evaluación Psicológica Forense del Riesgo de Violencia contra la Mujer en las relaciones de pareja (VCMP)*. Madrid: Colegio Oficial de Psicólogos de Madrid.
- GONZÁLEZ, J. L. MANZANERO, A., 2018: *Obtención y valoración del testimonio*. Madrid Pirámide.
- GONZÁLEZ-MONJE, A., 2019: *La dispensa del deber de declarar en violencia de género: problemas planteados y soluciones propuestas*. Thomson Reuters: Aranzadi.
- GRANHAG, P. A., STRÖMWALL, L. A. y Hartwig, M., 2005: «Eyewitness testimony: Tracing the beliefs of Swedish legal professionals», en *Behavioral Sciences & the Law*, 23, 709-727.
- GRANOT, Y., BALCETIS, E., FEIGENSON, N. y TYLER, T., 2018: «In the eyes of the law: perception versus reality in appraisals of video evidence», en *Psychology, Public Policy, and Law*, 24: 93-104.
- GROSCUP, J., 2004: «Judicial decision making about expert testimony in the aftermath of Daubert and Jumbo», en *Journal of Forensic Psychology Practice*, 4: 57-66.
- GUTHRIE, C., RACHLINSKI, J. J. y WISTRICH, A. J., 2001: «Inside the judicial mind», en *Cornell Law Review*, 86, 777-830.
- GUTHRIE, C., RACHLINSKI, J. J. WISTRICH, A. J., 2002: «Judging by heuristic: Cognitive illusions in judicial decision making», *Judicature*, 86, 44-50.
- , 2007: «Blinking on the bench: How judges decide cases», en *Cornell Law Review*, 93, 1-43.
- HAACK, S., 2020: «Judging expert testimony: from verbal formalism to practical advice», en *Quaestio facti*. Revista Internacional sobre Razonamiento Probatorio, 1: 13-30.
- HAUCH, V., SPORER, S. L., MICHAEL, S. W. y MEISSNER, C. A., 2016: «Does training improve detection of deception? A meta-analysis», en *Communication Research*, 43: 283-343.
- HERRERO, C., 2016: *Evaluación de falsas creencias e ideas pseudocientíficas sobre la conducta y la mente en estudiantes de Psicología, Criminología y Derecho*. Memoria Proyecto de Innovación Docente ID-0141. 2015-2016. Disponible en <http://hdl.handle.net/10366/13149>.
- , 2019: «Bases psico-jurídicas para confeccionar medidas y protocolos de actuación respecto al tratamiento de víctimas especialmente vulnerables», en M. del Pozo, L. Bujosa y A. González (Eds.), *Protocolos de actuación con víctimas especialmente vulnerables: Una guía de buenas prácticas* (pp. 23-40). Madrid, Spain: Thomson Reuters Aranzadi.
- , 2020: *Presentación al Curso de Especialización en Derecho: Obtención, interpretación y valoración de la prueba*, Universidad de Salamanca. Cursos de Especialización en Derecho.
- HERRERO, C., MASIP, J. y PICADO, E., 2016: «Mitos y pseudociencia sobre el comportamiento humano en criminólogos, juristas y policía», en *XI Congreso Español de Criminología. Abriendo vías a la reinserción*.

- HENDERSON, C. y EPSTEIN, J. (Eds), 2011: *The future of evidence. How science and technology will change the practice of law*. American Bar Association.
- HOUSTON, K. A., HOPE, L., MEMON, A. y READ, J., 2013: «Expert testimony on eyewitness evidence: in search of common sense», *Behavioral Sciences and the Law*, 31: 637-651.
- JORDAN, S., BRIMBAL, L., WALLACE, D. B., KASSIN, S. M. HARTWIG, M. y STREET, C. N. H., 2019: «A test of micro-expressions training tool: does it improve lie detection?», en *Journal of Investigative Psychology and Offender Profiling*, 16: 222-235.
- JUPE, L. M. y DENAULT, V., 2019: «Science and pseudoscience? A distinction that matters for police officers, lawyers and judges», en *Psychiatry, Psychology and Law*, 26: 753-765.
- KAHNEMAN, D. y TVERSKY, A., 1973: «On the Psychology of Prediction», en *Psychological Review*, 80: 2037-251.
- KASSIN, S. M., REDLICH, A. D., ALCESTE, F. y LUKE, T. J., 2018: «On the general acceptance of confessions research: opinions of scientific community», en *American Psychologist*, 73: 63-80.
- KASSIN, S., TUBB, A., HOSCH, H. M. y MEMON, A., 2001: «On the 'general acceptance' of eyewitness testimony research: A new survey of experts», en *American Psychologist*, 56, 405-416
- KASSIN, S. M, DROR, I. y KUKUCKA, J., 2013: «The forensic confirmation bias: problems, perspectives, and proposed solutions», en *Journal of Applied Research in Memory and Cognition*, 2: 42-52.
- KING, M., 1984: «Understanding the legal system: A job for psychologists?», en D. J. M Müller y D. E. Blackman y A. J. Chapman (Eds.), *Psychology and law* (pp. 67-82). Chichester John Wiley & Sons.
- KLENTZ, B. A., WINTERS, G. M. y CHAPMAN, J. E., 2020: «The CSI effect and the impact of DNA evidence on mock jurors and jury deliberations», en *Psychology, Crime and Law*, 26: 552-570.
- KNUTSSON, J. y ALLWOOD, C. M., 2014: «Opinions of legal professionals: Comparing child and adult witnesses' memory report capabilities». *The European Journal of Psychology Applied to Legal Context*. 10.1016/j.ejpal.2014.06.001.
- KOEHLER, J. y MEIXNER, J., 2015: «Decision making and the law: truth barriers», en G. Karen y G. Wu (Eds.), *The Wiley Blackwell Handbook of Judgment and Decision Making* (pp. 749-774). John Wiley & Sons.
- KOEHLER, J., 2020: «Forensic source conclusions: twenty threats to validity», en *Zeitschrift für Psychologie*, 228: 149-161.
- KOEN, W. J. y BOWERS, C. M. (Eds.) 2018: *The psychology and sociology of wrongful convictions: Forensic Science Reform*, NY: Elsevier-Academic Press.
- KÖHNKEN, G. y MANZANERO, A. L. SCOTT, T., 2015: «Análisis de la validez de las declaraciones: mitos y limitaciones». *Anuario de Psicología Jurídica*, 25: 13-19
- KOVERA, M. B. y LEVETT, L. M., 2015: «Jury decision making», en B. L. Cutler y P. A. Zapf (Eds.), *APA Handbook of Forensic Psychology: vol. 2. Criminal Investigation, adjudication, and sentencing outcomes* (pp. 271-311). American Psychological Association.
- KUKUCKA, J., 2018: «Confirmation bias in the Forensic Sciences: causes, consequences, and countermeasures», en W. J. Koen y C. M. Bowers (Eds.), *The psychology and sociology of wrongful convictions: Forensic Science Reform* (pp. 215-245), NY: Elsevier-Academic Press.
- KUKUCKA, J., HILEY, A. y KASSIN, S. M., 2020: «Forensic confirmation bias: do jurors discount examiners who were exposed to task-irrelevant information», en *Journal of Forensic Sciences*, 65: 1978-1990.
- KUKUCKA, J., KASSIN, S., ZAPF, P. A. y DROR, I., 2017: «Cognitive bias and blindness: a global survey of forensic science examiners», en *Journal of Applied Research in Memory and Cognition*, 6: 452-459.
- LAW COMMISSION, 2009: «*The Admissibility of Expert Evidence in Criminal Proceedings in England and Wales: A New Approach to the Determination of Evidentiary Reliability*», Law Com. Consultation Paper No. 190.
- , 2011: «*Expert Evidence in Criminal Proceedings in England and Wales*», Law Com. Report No. 325 (hereafter 'Report').
- LILIENTHAL, O., 2012: «Public Skepticism of Psychology. Why may people perceive the study of human behavior as unscientific», en *American Psychologist*, 67: 111-129.

- LILIENTFELD, S. O., LYNN, S. J., RUSCIO, J., y BEYERSTEIN, B. L., 2010: *50 great myths of popular psychology: Shattering widespread misconceptions about human behavior*. Chichester, England: Wiley-Blackwell.
- LOFTUS, E., 1993: «The reality of Repressed Memories», en *American Psychologist*, 48: 518-537.
- LOFTUS, E.F. COLE, S. A., 2004: «Contaminated Evidence», en *Science*, 304: 959.
- LUPÁRIA, L., 2015 (Ed.): *Understanding wrongful conviction. The protection of innocent across Europe and America*. Wolters Kluwer.
- MACLEAN, C. L. y DROR, I., 2016: «A primer on Psychology of cognitive bias», en C. Robertson y A. Kesselheim (Eds.), *Blinding as a Solution to Bias: Strengthening Biomedical Science, Forensic Science, and Law* (pp. 13-24). New York: Elsevier, Inc.-Academic Press.
- MACLEAN, N., NEAL, T. M.S., MORGAN, R. D. y MURRIE, D. C., 2019: «Forensic clinicians' understanding of bias», en *Psychology, Public Policy, and Law*, 25: 323-330.
- MAGNUSSEN, S., MELINDER, A., STRIDBECK, U. y RAJA, A. Q., 2010: «Beliefs about factors affecting the reliability of eyewitness testimony: a comparison of judges, jurors and general public». *Applied Cognitive Psychology*, 24, 122-133
- MAGNUSSEN, S., SAFER, M. A., SARTORI, G. WISE, R. A., 2013: «What Italian defense attorneys know about factors affecting eyewitness accuracy: a comparison with U.S. and Norwegian samples», en *Frontiers in Psychiatry*, doi: 10.3389/fpsy.2013.00028.
- MAGNUSSEN, S., WISE, R. A., RAJA, A. Q., SAFER, M. A., PWELLENKO, N. y STRIDBECK, U., 2008: «What judges know about eyewitness testimony: A comparison of Norwegian and US judges», en *Psychology, Crime and Law*, 14: 177-188.
- MAGRO-SERVET, V., 2018: «Análisis de la Doctrina Jurisprudencial reciente en violencia de género», en *Diario La Ley*, N° 9278, Sección Dossier, 15 de Octubre de 2018,
- MASIP, J., 2005: «¿Se pillan antes a un mentiroso que a un cojo? Sabiduría popular frente a conocimiento científico sobre la detección no-verbal del engaño», en *Papeles del Psicólogo*, 78-91.
- , 2017: «Deception detection: state of art and future prospects», en *Psicothema*, 29: 149-159.
- MASIP, J., BARBA, A. y HERRERO, C., 2012: «Behaviour Analysis Interview and common sense. A study with novice and experienced officers», en *Psychiatry, Psychology and Law*, 19: 21-34.
- MASIP, J. y HERRERO, C., 2015: «Nuevas aproximaciones en detección de mentiras I. Antecedentes y Marco teórico», en *Papeles del Psicólogo*, 36: 83-95.
- , 2015b: «Nuevas aproximaciones en detección de mentiras II. Estrategias activas de entrevista e información contextual», en *Papeles del Psicólogo*, 36: 96-108.
- , 2015c: «Police detection of deception: Beliefs about behavioral cues to deception are strong even though contextual evidence is more useful», en *Journal of Communication*, 65, 125-145.
- , 2017: «Examining police officers' response bias in judging veracity», en *Psicothema*, 29: 490-495.
- MASIP, J., HERRERO, C., GARRIDO, E. y BARBA, A., 2011: «Is the behavior Analysis Interview just common sense?», en *Applied Cognitive Psychology*, 25: 593-604.
- MAZZONI, G., 2019: *Psicología del Testimonio*. Trotta.
- MCAULIFF, B. D. y GROSCUP, J. L., 2009: «Daubert and Psychological Science in court: judging validity from bench bar, and jury box», en J. L. SKEEM, K. S. DOUGLAS y S. O. LILIENTFELD (Eds.), *Psychological Science in the Courtroom* (pp. 26-52). Nueva York: Gilford Press.
- MCAULIFF, B. D., KOVERA, M. B. y NUÑEZ, G., 2009: «Can jurors recognize missing control groups, confounds and experimenter bias in psychological science?», en *Law and Human Behavior*, 33: 247-257.
- MELINDER, A., BURREL, L., ERIKSEN, M. O., MAGNUSSEN, S. y WESSEL, E., 2016: «The emotional child witness effect survives presentation mode», en *Behavioral Sciences and the Law*, 34: 113-125.
- MUNRO, G. D. y MUNRO, C. A., 2014: «“Soft” versus “hard” psychological science: biased evaluations of scientific evidence that threatens or supports a strongly-held political identity», en *Basic and Applied Social Psychology*, 36: 533-543.
- NATIONAL COMMISSION ON FORENSIC SCIENCE (2015). Ensuring that forensic analysis is based upon task-relevant information. 2015 Dec. <https://www.justice.gov/archives/ncfs/file/818196/download> (Septiembre 2020).

- NATIONAL RESEARCH COUNCIL, 2009: «Committee on Identifying the Needs of the Forensic Science Community, *Strengthening Forensic Science in the United States: A Path Forward* National Academies Press, 53 ('NRC Report').
- NATIONAL RESEARCH COUNCIL, 2011: *Reference Manual on Scientific Evidence*, Third Edition. Washington, D. C: National Academic Press (NAP).
- , 2014: *Identifying the Culprit: Assessing Eyewitness Identification*. Washington, DC: The National Academies Press (NAP).
- NEAL, T. M. S., 2017: Discerning bias in forensic psychological reports in insanity cases. *Behavioral Sciences & Law*, doi: 10.1002/bsl.2346
- NEAL, T. M. S., HIGHT, M., HOWATT, B. y HAMZA, C., 2018: «The cognitive and social psychological bases of bias in forensic mental health judgments», en M.K. Miller and B.H. Bornstein (Eds), *Advances in Psychology and Law: Volume 3* (pp. 151-175). New York: Springer.
- NEAL, T. M. S., SLOBOGIN, C., SAKS, M. J., FAIGMAN, D. L. y GEISINGER, K. F., 2019: «Psychological assessments in legal context: Are Courts Keeping “Junk Science” Out of the Courtroom?», en *Psychological Science in the Public Interest*, 20: 135-164.
- NEUFELD, P. J., 2005: «The (near) irrelevance of Daubert to criminal justice and some suggestions for reform», en *American Journal of Public Health*, 95, 107-113.
- NEUFELD, P. J. y SCHECK, B., 2010: «Making forensic science more scientific», en *Nature*, 464: 351.
- NICKERSON, R. S., 1998: «Confirmation bias: A ubiquitous phenomenon in many guises», en *Review of General Psychology*, 2: 175–220.
- NIETO, A. J., 2019: «El principio de intermediación, el lenguaje no verbal y gestual y las micro-expresiones faciales: A propósito de la STD 119/2019, de 6 de Marzo», en *Diario La Ley*, nº 9491, Sección Tribuna, 4 de Octubre.
- NEIVA, J., 2010: *La valoración de la prueba*, Barcelona: Marcial Pons.
- NISBETT, R. E. y BORGIDA, E., 1975: «Attribution and the psychology of prediction», en *Journal of Personality and Social Psychology*, 35: 932-943.
- NIST, EXPERT WORKING GROUP FOR HUMAN FACTORS IN HANDWRITING EXAMINATION, 2020. *Forensic Handwriting Examination and Human Factors: Improving the Practice Through a Systems Approach*. U. S. Department of Commerce, National Institute of Justice, National Institute of Standards and Technology (NIST). NISTIR 8282.
- NIST, EXPERT WORKING GROUP ON HUMAN FACTORS IN LATENT PRINT ANALYSIS, 2012: *Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach*. U. S. Department of Commerce, National Institute of Justice, National Institute of Standards and Technology (NIST).
- OBSERVATORIO CONTRA LA VIOLENCIA DOMÉSTICA DE GÉNERO (2018, noviembre). Grupo de expertas y expertos del Consejo General del Poder Judicial en materia de violencia de género. *Guía de buenas prácticas para la toma de declaración de víctimas de violencia de género*. Madrid: Consejo General del Poder Judicial
- OEBERST, A. y GOECKENJAN, I., 2016: «When being wise after the event results in injustice: evidence for hindsight bias in judges' negligence assessments», en *Psychology, Public Policy, and Law*, 22, 271.
- OIG, OFFICE OF THE INSPECTOR GENERAL, 2006: *A review of the FBI's handling of the Brandon Mayfield case*. Office of the Inspector General, Oversight & Review Division, US Department of Justice.
- OIG., 2011: *A review of the FBI's progress in responding to the recommendations in the office of the inspector general report on the fingerprint misidentification in the Brandon Mayfield case*. <<http://www.justice.gov/oig/special/s1105.pdf>>.
- OSBORNE, N., BIRD, C. y STOEL, R., 2019: «Forensic handwriting examination and cognitive bias: recommendations from the NIST expert working group on human factors», en *Australian Journal of Forensic Sciences*, 51: Sup1, 141-144.
- OTGAAR, H., HOWE, M. L., MERCKELBACH, H. y MURIS, P., 2018: «Who es the better eyewitness? Sometimes adults but other times children», en *Current Directions in Psychological Science*, 27, 378-385

- OTGAAR, H., HOWE, M. L., PATIHIS, L., MERCKELBACH, H., LYNN, S. J., LILIENFELD, S. O. y LOFTUS, E., 2019: «The return of the repressed: The persistent and problematic claims of long-forgotten trauma», en *Perspectives on Psychological Science*, 14: 1072-1095.
- OTGAAR, H., WANG, J., DODIER, O., HOWE, M. L., LILIENFELD, S. O., LOFTUS, E. F., LYNN, S. J., MERCKELBACH, H. y PATIHIS, L., 2020: «Skirting the issue: What does believing in repression mean?» en *Journal of Experimental Psychology: General*, 149: 2005–2006. doi: /10.1037/xge0000982
- PATIHIS, L., HO, L. Y., TINGEN, I. W., LILIENFELD, S. O. y LOFTUS, E., 2014: «Are the “memory wars” over? A scientist-practitioner gap in beliefs about repressed memory», en *Psychological Science*, 25: 519-530.
- PORTER, S. y TEN BRINKE, L., 2008: «Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions», en *Psychological Science*, 19: 508-514.
- PRONIN, E., LIN, D. y ROSS, L., 2002: «The bias blind spot: perceptions of bias in self versus others», en *Personality and Social Psychology Bulletin*, 28: 369–381.
- QUIGLEY-McBRIDE, A. WELLS, L., 2018: «Fillers can help control for contextual bias in forensic comparison task», en *Law and Human Behavior*, 42: 295-305.
- RACHLINSKI, J. J., 1998: «A positive psychological theory of judging in hindsight», en *University of Chicago Law Review*, 65, 571-625.
- RACHLINSKI, J. J., WISTRICH, A. J. y GUTHRIE, C., 2015: «Can judges make reliable numeric judgments? Distorted damages and skewed sentences», en *Indiana Law Journal*, 90: 695-739.
- RAMÍREZ ORTIZ, J. L., 2014: «Verdad, proceso y derecho penal (Interrogatorios en la habitación 101)», en *Jueces para la Democracia. Información y debate*, 79: 18-37.
- , 2018: «La prueba en los delitos contra la indemnidad sexual», en *Diario La Ley*, 9199, Sección Doctrina, Editorial Wolters Kluwer.
- , 2020: «El testimonio único de la víctima en el proceso penal desde la perspectiva de género», en *Quaestio facti. Revista Internacional sobre Razonamiento Probatorio*, 1: 201-245.
- REDFIELD, S. E., 2017: *Enhancing Justice. Reducing Bias*. American Bar Association.
- REISBERG, D., 2014: *The Science of perception and memory: a pragmatic guide for Justice System*. Oxford: Oxford University Press
- REPORT TO THE PRESIDENT. PRESIDENT’S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY 2016: «*Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*».
- RIBEIRO, G., TANGEN, J. M. y MCKIMMIE, B. M., 2019: «Beliefs about error rates and human judgment in forensic science», en *Forensic Science International*, 297, 138-147.
- RISINGER, D. M., SAKS, M. J., THOMPSON, W. C. ROSENTHAL, R., 2002: «The Daubert/Kumho implications of observer effects in forensic science: Hidden problems of expectation and suggestion», en *California Law Review*, 90 (1): 1-56.
- ROBBENOLT, J. K. y HANS, V. P., 2016: *The Psychology of Tort Law*. New York: NY University Press.
- ROBERTSON, C. y KESSELHEIM, A., eds., 2016: *Blinding as a Solution to Bias: Strengthening Biomedical Science, Forensic Science, and Law*. New York: Elsevier, Inc.-Academic Press.
- ROECH, K., 2012: «Wrongful convictions in Canada», en *University of Cincinnati Law Review*, 80: 1465-1526.
- ROSS, L., 1977: «The intuitive Psychologist And His Shortcomings: Distortions in the Attribution Process», en L. Berkowitz (Ed.), *Advances in experimental social psychology*, (Vol. 10, pp. 173-220). New York: Academic Press.
- ROWELL, A. y BILZ, K., 2021: *The Psychology of Environmental Law*. New York University Press.
- ROYAL STATISTICAL SOCIETY, 2017: *Statistics and probability for advocates: understanding the use of statistical evidence in courts and tribunals*.
- SAKS, M. J. y FAIGMAN, D. L., 2008: «Failed forensics: how forensic science lost its way and how it might yet find it», en *Annual Review of Law and Social Science*, 4:149–171.
- SAKS, M. J. y KOEHLER, J. J., 2005: «The coming paradigm shift in forensic identification science», en *Science*, 309: 892-895.
- SAKS, M. J. y SPELLMAN, B. A., 2016: *The psychological foundations of Evidence Law*. New York: NY University Press.

- SCHACTER, D. L., 2002: *The seven sins of memory: how the mind forgets and remembers*. Houghton Mifflin (Traducc. Castellano, 2003, 2007 y 2011 *Los siete pecados de la memoria: cómo olvida y recuerda la mente*. Ariel)
- SCHANZ, K. y SALEFATI, C. G., 2016: «The CSI effect and its controversial existence and impact: a mixed methods review», en *Crime Psychology Review*, 2: 60-79.
- SCHAUER, F. y SPELLMAN, B., 2013: «Is expert evidence really different?», en *Notre Dame Law Review*, 89: 1-26.
- SCHAUER, F., 2006: «On the supposed jury-dependence of evidence law», *University of Pennsylvania Law Review*, 155: 165-202.
- , 2010a: «Is There a Psychology of Judging? », en D. Klein y G. Mitchell (Eds.), *The Psychology of Judicial Decision Making* (pp. 103-120). Oxford University Press.
- , 2010b: «Can Bad Science be Good Evidence? », en *Cornell Law Review*, 95: 1191-1220.
- SHAH, P., MICHAL, A., IBRAHIM, A., RHODES, R. y RODRÍGUEZ, F., 2017: «What makes everyday scientific reasoning so challenging? en *Psychology of Learning and Motivation*, 66: 251-299.
- SHAPIRO, D. L., MIXON, L., JACKSON, M. SHOOK, J., 2015: «Psychological expert witness testimony and judicial decision making trends», en *International Journal of Law and Psychiatry*, 42-43: 149-153.
- SIMONCELLI, T., 2016: «Rigor in Forensic Science», en C. Robertson y A. Kesselheim (Eds.), *Blinding as a Solution to Bias: Strengthening Biomedical Science, Forensic Science, and Law* (pp. 129-131). New York: Elsevier, Inc.-Academic Press.
- SMITH, S. M., STINSON, V. y PATRY, M. W., 2011: «Fact or fiction? The myth and reality of the CSI effect», en *Court Review*, 47: 4-7.
- SNOOK, B., MCCARDLE, M. I., FAHMY, W. y HOUSE, J. C., 2017: «Assessing truthfulness on the witness stand: Eradicating deeply rooted pseudoscientific beliefs about credibility assessment by triers of fact», en *Canadian Criminal Law Review*, 22: 297-306.
- SOBRAL, J., 2008: «Psicología, testimonios y testigos: un asunto clave en la investigación criminal», en V. Garrido y J. Sobral, *La psicología aplicada al descubrimiento, captura y condena de los criminales* (pp.371-393), Barcelona: Nabla Ed.
- SPELLMAN, B. y SCHAUER, F. 2012a: «Law and Social Cognition». *Virginia Public Law and Legal Theory Research Paper* No. 2012-10. Disponible en SSRN: <https://ssrn.com/abstract=2000806>.
- 2012b: «Legal Reasoning», en K. J. Holyoak y R. G. Morrison (Eds.), *The Oxford Handbook of thinking and reasoning* (pp.719-735). Oxford: Oxford University Press.
- SPELLMAN, B. y WEAVER, C. A., 2020 (forthcoming): «Memory and the Law, en M. Kahane y A. Wagner (Eds.), *Handbook of Human Memory -- Vol. II: Applications*. Oxford (Prox), en *Virginia Public Law and Legal Theory Research Paper* No. 2020-56, recuperado de SSRN: <https://ssrn.com/abstract=3644394>
- SPELLMAN, B., 2007: «On the supposed expertise of judges in evaluating evidence. In response to Frederick Schauer, On the Supposed Jury-Dependence of Evidence Law, 155 U. PA. L. REV. 165 2006: *University of Pennsylvania Law Review*, 156:1-9.
- SPOERER, S. L., 2007: «Evaluating witness evidence: The fallacies of intuition», en C. Engel y F. Strack (Eds.), *The impact of court procedure on the psychology of judicial decision making* (pp. 111-150). Baden-Baden, Germany: Nomos Verlag.
- STEBLAY, N., HOSCH, H. M., CULHANE, S. E. y MCWETHY, A., 2006: «The impact on juror verdicts of judicial instruction to disregard inadmissible evidence: A meta-analysis», en *Law and Human Behavior*, 30: 469-492.
- STERN, S. M. y LEWINSOHN-ZAMIR, D., 2020: *The Psychology of Property Law*. NY University Press
- STERNLIGHT, J. R. y ROBBENOLT, J. K., 2008: «Good Lawyers Should Be Good Psychologists: Insights for Interviewing and Counseling Clients», en *Ohio State Journal on Dispute Resolution*, 23: 437-
- STEVENAGEM, S. V. y BENNETT, A., 2017: «A biased opinion: demonstration of cognitive bias on fingerprint matching task through knowledge of DNA test results», en *Forensic Science International*, 276: 93-106.

- STOLZENBERG y LYON, T. D., 2014: «How Attorneys Question Children About the Dynamics of Sexual Abuse and Disclosure in Criminal Trials», en *Psychology, Public Policy, and Law*, 20: 19-30.
- STRÖMWALL, L. y GRANHAG, P. A., 2003: «How detect deception? Arresting the beliefs of police officers, prosecutors and judges», en *Psychology, Crime, and Law*, 9: 19-36.
- SUNDE, N. y DROR, I., 2019: «Cognitive and human factors in digital forensic: problems, challenges, and the way forward», en *Digital Investigation*, 29: 101-108.
- TADEI, A., FINNILÄ, K., REITE, A., ANTFOLK, J. y SANTTILA, P., 2016: «Judges' Capacity to Evaluate Psychological and Psychiatric Expert Testimony», en *Nordic Psychology*, 68: 204-217.
- TARUFFO, M., 2002: *La prueba de los hechos*. (trad. cast. de Jordi Ferrer Beltrán, Madrid: Trotta).
- , 2005 a: «Conocimiento Científico y estándares de prueba judicial», en *Boletín Mexicano de Derecho Comparado*, año XXXVIII, 114: 1285-1312
- , 2005 b: «Prólogo», en J. Ferrer, *Prueba y verdad en el Derecho* (2ª ed.). Barcelona: Marcial Pons.
- , 2010: *Simplemente la verdad. El juez y la construcción de los hechos*, Barcelona: Marcial Pons.
- , 2013: «La aplicación de estándares científicos a las ciencias sociales y forenses», en C. Vázquez (ed.), *Estándares de prueba y prueba científica. Ensayos de epistemología jurídica* (pp. 203-213), Madrid: Marcial Pons.
- TECHNICAL WORKING GROUP FOR EYEWITNESS EVIDENCE, 1999: *Eyewitness evidence. A guide for law enforcement*. Washington DC: U.S. Department of Justice.
- THOMPSON, W. C., 2018: *Developing effective methods for addressing contextual bias in Forensic Science*. National Institute of Justice. Report of Grant. Final Summary Overview. Recuperado de <https://www.ncjrs.gov/pdffiles1/nij/grants/252784.pdf>.
- TUPPER, N., SAUERLAND, M., HOPE, L. y MERCKELBACH, H., 2015: «Seeing and believing: common courtroom myths in eyewitness memory», en *The Inquisitive Mind*, 28 <https://www.in-mind.org/article/seeing-and-believing-common-courtroom-myths-in-eyewitness-memory>
- TYLER, T. R., 2006: «Viewing CSI and the threshold of guilt: Managing truth and justice in reality and fiction», en *The Yale Law Journal*, 115: 1050-1085.
- VARIOS AUTORES, 2015: «*Criminal Justice and forensic science evidence: current controversies*. Monográfico British Journal of American Studies 2015.
- VÁZQUEZ, C., 2013: *Estándares de prueba y prueba científica. Ensayos de epistemología jurídica*, Barcelona: Marcial Pons.
- , 2015a: *De la prueba científica a la prueba pericial*, Barcelona: Marcial Pons.
- , 2015b: «La admisibilidad de las pruebas periciales y la racionalidad de las decisiones judiciales», en DOXA, *Cuadernos de Filosofía del Derecho*, 38: 101-130.
- , 2016: «La prueba pericial en la experiencia estadounidense. El caso *Daubert*», en *Jueces para la Democracia. Información y Debate*, 86: 92-112.
- VRIJ, A., 2018: «Verbal lie detection tools from an applied perspective», en J. P. Rosenfeld (Ed.), *Detecting concealed information and deception: Recent developments* (pp. 297-327). San Diego, CA: Elsevier: Academic Press.
- VRIJ, A. y FISHER, R. P., 2020: «Unraveling the misconception about deception and nervous behavior», en *Frontiers in Psychology*, 11: 1377. doi: 10.3389/fpsyg.2020.01377
- VRIJ, A., GRANHAG, P.-A. y PORTER, S., 2010: «Pitfalls and opportunities in nonverbal and verbal lie detection», en *Psychological Science in the Public Interest*, 11: 89-121.
- VRIJ, A., HARTWIG, M. y GRANHAG, P. A., 2019: «Reading lies: Nonverbal Communication and Deception», en *Annual Review of Psychology*, 70: 295-317.
- WECHSLER, H. J., KEHN, A., WISE, R. A. y CRAMER, R. J., 2015: «Attorney beliefs concerning scientific evidence and expert witness credibility», en *International Journal of Law and Psychiatry*, 41, 58-66.
- WELLS, G. L., MALPASS, R. S., LINDSAY, R. C. L., FISHER, R. P., TURTLE, J. W. y FULERO, S. M., 2000: «From the lab to the police station. A successful application of eyewitness research», en *American Psychologist*, 55: 581-598.
- WELLS, G. L., KOVERA, M. B., DOUGLASS, A. B., BREWER, N., MEISSNER, C. A. y WIXTED, J. T., 2020: «Policy and procedure recommendations for collection and preservation of eyewitness identification evidence», en *Law and Human Behavior*, 44: 3-36.

- WELLS, G. y LOFTUS, E., 2003: «Eyewitness memory for people and events», en A. M. Goldstein y I. B. Weiner (Eds.), *Handbook of Psychology, Vol. 1, Forensic Psychology*, (pp. 149-160). . Hoboken, NJ: John Wiley & Sons.
- WISE, R. A. y SAFER, M. A., 2004: «What US Judges know and believe about eyewitness testimony», en *Applied Cognitive Psychology*, 18, 427-443.
- WISE, R. A., SAFER, M. A. y MARO, C. A., 2011: «What U.S. law enforcement officers know and believe about eyewitness factors, eyewitness interviews and identification procedures», en *Applied Cognitive Psychology*, 25, 488-500.
- WISE, R. A., SARTORI, G., MAGNUSSEN, S. y SAFER, M. A., 2014: «An examination of the causes and solutions to eyewitness error», en *Frontiers Psychology*, doi: 10.3389/fpsy.2014.00102.
- WISSLER, R. L., WILLIAMS, K. E. y SAKS, M. J., 2013: «Dual-Processing models of admissibility: how legal test for the admissibility of scientific evidence resemble cognitive science's System 1 and System 2», en *Virginia Journal of Law and Technology*, 17: 356-371.
- WISTRICH, A. J. y RACHLINSKI, J. J., 2017: «Implicit bias in judicial decision making: how it affects judgment and what judges can do about it», en S. E. Redfield (Ed.), *Enhancing Justice. Reducing Bias* (pp. 87-130). Chicago: American Bar Association.
- WISTRICH, A. J., GUTHRIE, C. y RACHLINSKI, J. J., 2005: «Can judges ignore inadmissible information? The difficulty of deliberately disregarding», en *University of Pennsylvania Law Review*, 153: 1251-1345.
- WISTED, J. T y WELLS, G. L., 2017: «The relationship between eyewitness confidence and identification accuracy: a new synthesis», en *Psychological Science in the Public Interest*, 18: 10-65.

IURIS PRUDENTIA

¿DEBERÍA PROBARSE EL ODIO? CONSIDERACIONES EN TORNO A LA SENTENCIA DEL CASO DIANA SACAYÁN*

María Laura Manrique
Investigadora CONICET
laura.manrique@conicet.gov.ar

RESUMEN: En este trabajo analizo *críticamente* la sentencia del «caso Sacayán». Este caso es relevante porque, entre otras cosas, aplica por primera vez en la Argentina el agravante de odio a la identidad de género al homicidio cometido contra una persona travesti, Diana Sacayán. Por otra parte, también sugiero que la distinción entre delitos de odio y delitos odiosos puede ser de utilidad para entender el rechazo que nos generó este hecho pero que ello es independiente de si el hecho fue cometido por odio o no.

PALABRAS CLAVE: «travesticidio»; delito de odio o delitos odiosos.

SHOULD WE PROVE HATE? SOME COMMENTS ON DIANA SACAYÁN'S CASE

ABSTRACT: In this paper, I, firstly, criticize the judicial decision in the Sacayán case. This is a central case because it applied for the first time in Argentina the aggravating hate motive to the identity

* A lo largo de la elaboración de este trabajo me he beneficiado de los comentarios y sugerencias que realizaron Belén Gulli, Hernán Bouiver, IB, Javier Álvarez, Pablo Navarro. Gracias a ellos, el trabajo es mejor de lo que era antes de sus comentarios.

of gender in the killing of a transvestite person, Diana Sacayán. Secondly, I suggest that the distinction between hate crimes and heinous crimes could be useful to understand the disgust that this homicide generated, although this is irrespective of whether the crime was in fact motivated by hate or not.

KEYWORDS: «transvesticide»; hate crimes vs. heinous crimes.

SUMARIO: 1. INTRODUCCIÓN.— 2. EL CONTROL RACIONAL DE LOS FUNDAMENTOS.— 3. EL VOTO DEL MAGISTRADO CALVETE.— 4. EL VOTO DEL MAGISTRADO JULIO BÁEZ.— 5. EL VOTO DE LA MAGISTRADA IVANA BLOCH.— 6. DELITOS DE ODIOS Y DELITOS ODIOSOS.— 7. BALANCE.— 8. BIBLIOGRAFÍA

RECOMMENDED CITATION: MANRIQUE, MARÍA LAURA: «¿Debería probarse el odio? Consideraciones en torno a la sentencia del caso Diana Sacayán», in *Quaestio facti*, 2: 411-441. Madrid: Marcial Pons Ediciones Jurídicas y Sociales. DOI: http://dx.doi.org/10.33115/udg_bib/qf.i2.22494

1. INTRODUCCIÓN

Diana Sacayán fue asesinada en su domicilio el 11 de octubre de 2015. Su cuerpo fue hallado amordazado y atado de pies y manos. En el lugar de los hechos se encontró una tijera, un martillo y un cuchillo ensangrentado, con hoja de 20 cm, que había sido utilizado en el homicidio. La víctima presentaba 27 lesiones, de las cuales 13 eran heridas punzo cortantes. Luego de casi tres años de proceso, el 18 de junio de 2018, Gabriel David Marino fue hallado responsable de la muerte de Sacayán y condenado por el delito de homicidio doblemente agravado por el odio a la identidad de género y por haber mediado violencia de género (art. 80 inc. 4 e inc. 11, respectivamente del Código Penal argentino).

Este caso es trascendente no solo por la muerte de una persona, sino también porque reunió un amplio conjunto de características que generaron una enorme repercusión¹. Por ejemplo, Diana Sacayán, era una reconocida militante del colectivo LGTBI y su muerte —que algunos medios calificaron como una «exhibición de atrocidad»²— fue rápidamente asociada a esa trayectoria personal de la víctima. Desde un punto de vista jurídico, el caso es notablemente relevante porque en la sentencia condenatoria —considerada por muchos agentes sociales como trascendente o

¹ La repercusión mediática (prensa y redes sociales) puede apreciarse fácilmente en internet. Por ejemplo, la expresión «Diana Sacayán» está asociada en Google a más de 30.000 resultados, la mayoría de los cuales fueron publicados entre el momento de su homicidio y los meses posteriores a la condena de Marino.

² Infobae lo identificó como una «exhibición de atrocidad». <https://www.infobae.com/sociedad/2016/11/09/diana-sacayan-los-aberrantes-dichos-de-su-novio-que-ira-a-juicio-por-asesinarla/>

histórica— se aplica por primera vez en Argentina la agravante de odio a la identidad género, acuñándose para estas circunstancias específicas, el término «travesticidio».

La sentencia del Tribunal Oral 4 de la Ciudad de Buenos Aires fue firmada por Adolfo Calvete como presidente del Tribunal y los vocales Julio Báez e Ivana Bloch. Sin embargo, la magistrada Bloch disiente del encuadre jurídico ofrecido por sus colegas, negando que, en este caso, hubiese bases suficientes para aplicar la agravante de odio a la identidad de género³.

El análisis de este extenso fallo promueve importantes interrogantes de naturaleza dogmática⁴. Por ejemplo, ¿qué se entiende por «odio» en el código penal argentino?, ¿qué bienes pretendría proteger el legislador al momento de dictar la norma?, ¿se pueden reprochar los motivos que determinan las acciones en un derecho penal liberal (i.e., centrado en los actos como base de atribución de responsabilidad)?, ¿de qué manera se puede/debe probar el odio?, etc. A su vez, también es posible mencionar preguntas vinculados al razonamiento probatorio y estándares de prueba en el proceso penal. Por ejemplo, ¿cómo deberían identificarse los estados mentales?, ¿se puede establecer un estándar probatorio para que se considere acreditado que un individuo ha actuado movido por el odio?⁵; ¿cuál es el peso que deben tener otras explicaciones alternativas al hecho?, etc.⁶.

La respuesta a cada uno de estos interrogantes merecería un cuidadoso estudio que no se puede realizar aquí. Por esa razón, en este trabajo me centraré en el modo en que la agravante del odio por identidad de género es aplicada en la sentencia del «Caso Sacayán». La estructura de este trabajo es simple: en general, sus diversas secciones ofrecen un análisis de los diferentes votos que fundamentan la sentencia. En la sección 2 señalaré un problema general que afecta al voto mayoritario: las dificultades para un control racional de la argumentación por un uso abusivo de lenguaje emo-

³ Cuando el artículo se encontraba en prensa, el 2 de Octubre de 2020 la Sala I de la Cámara Nacional de Casación en la Criminal y Correccional de la Capital Federal resolvió mantener la condena a Marino pero rechazó el agravante por odio y confirmó la agravante por femicidio (inciso 11 del artículo 80 del código penal argentino). No se analizan en este trabajo los argumentos de ese recurso.

⁴ El fallo posee 414 páginas, un abordaje completo de la sentencia haría el trabajo no solo tedioso sino también inabordable. Por esa razón, aquí solo me centraré en la cuestión que considero relevante para el fundamento de la decisión como un homicidio como un delito de odio. Tampoco trataré todos los problemas teóricos vinculados a los delitos de odio. El presente trabajo no pretende ser un estudio sistemático de este tema. He abordado algunos problemas de los delitos de odio en MANRIQUE, 2019 Y MANRIQUE 2020. Para un estudio general sobre odio véase, por ejemplo, DIAZ LOPEZ, 2013, PERALTA, 2012; HURD Y MOORE, 2004; DUFF Y MARSHALL 2015.

La falta de tratamiento de otros temas centrales tratados en la sentencia no pretende restarle valor, sino que no están vinculados con el interés teórico de este trabajo. Por otro lado, un punto que valdría la pena considerar, pero no abordaré aquí es cuestionarse sobre la necesidad y utilidad de que los fallos posean esa extensión.

⁵ Quien sostiene con énfasis que diferentes estándares probatorios deberían exigirse en distintas etapas del proceso o tipos de procesos es, entre otros, FERRER BELTRÁN, 2013.

⁶ Para el tratamiento de problemas relevantes vinculados a la misma sentencia, véase ÁLVAREZ, 2019: 365-383

tivo y simplificaciones injustificadas. En las secciones 3 y 4 analizaré los principales argumentos de la mayoría, *i.e.*, los votos de los magistrados Calvete y Báez, respectivamente. En la sección 5, revisaré los fundamentos de la disidencia, del voto de la magistrada Bloch, y en la sección 6 señalaré determinados aspectos de los delitos de odio a la identidad de género que necesitan ser mejor reconstruidos para comprender adecuadamente la naturaleza de esta agravante. Allí defenderé dos cosas. Por una parte, que es preciso distinguir cuidadosamente entre delitos de odio y delitos odiosos⁷, y, por otra parte, que aun cuando un delito odioso —como indudablemente fue el homicidio de Sacayán— nos cause un profundo desagrado, nuestro sistema jurídico impide que este sentimiento de repulsión pueda ser una razón para agravar las penas. Por último, en la sección 7 haré un balance de lo dicho a lo largo del artículo.

2. EL CONTROL RACIONAL DE LOS FUNDAMENTOS

En la reconstrucción del caso, en el voto mayoritario se encuentra no solo una presentación de «datos brutos» sino que se realizan algunas afirmaciones que van más allá de la mera descripción de la información sobre la cual se juzgará el caso. Por ello, como un paso previo a analizar sus ideas centrales, subrayaré algunas de esas afirmaciones ya que condicionan todo el enfoque del voto mayoritario.

En la página 3, el juez Calvete sostiene que la víctima⁸: «Presentaba certeros signos de haber sido víctima de un hecho cometido con un alto grado de violencia, lo que fue ratificado, luego, con el informe de autopsia...», o que las lesiones infringidas a Diana Sacayán fueron de «extrema brutalidad, insensibilidad y, por su pluralidad y especificidad, dirigidas a marcar el rasgo específico típico del odio...» (STO Núm. 4, 2018: 175).

A pesar de la contundencia de estas afirmaciones no hay en la descripción de los hechos un criterio que permita determinar el «alto grado» de violencia y tampoco se encuentra en los fundamentos de sus votos nada que explique las razones del homicidio (*i.e.*, el odio), ni mucho menos del estado mental de Gabriel Marino al momento de cometer el delito. Así, por ejemplo, el tribunal no se preocupa en mostrar por qué las lesiones que presentaba Sacayán son muestra de una violencia cualitativamente diferente a otros hechos cometidos con la misma arma (cuchillo). De igual modo, no

⁷ Por delito odioso estoy pensando en los denominados *heinous crime* del mundo anglosajón. Estos son crímenes que nos parecen brutales u odiosos por el hecho en sí y no por el estado mental del imputado. El elemento del odio no se encuentra en el imputado sino en el modo en que la comunidad percibe el crimen. Véase, por ejemplo, REAMER, 2005. Para un resumen de la discusión teórica, véase DÍAZ LÓPEZ, 2013: 68-77.

⁸ Sentencia del Tribunal Oral en lo Criminal y Correccional Núm. 4 de la Capital Federal (Argentina). Sentencia dictada 18/06/2018. Fecha de los fundamentos 06/07/2018. Causa número: 62.162/2015. En adelante: STO NÚM. 4, 2018: 3. Para ver la sentencia completa, véase: <https://fislem.org/wp-content/uploads/2019/04/Fallos-Sacayan.pdf>

queda claro a qué se refiere el tribunal cuando afirma, en p.3, que «las circunstancias del contexto y modo de comisión del hecho permitieron suponer...» que el homicidio estuvo determinado por la condición mujer trans de la víctima.

Esa combinación de lenguaje emotivo y simplificaciones sobre el modo en que ocurrieron los hechos plantea un serio inconveniente: dificulta el control racional de las afirmaciones (CARRIÓ, 1983: 77-79). Este tipo de control es fundamental en un proceso penal que exige una descripción *verdadera* de los hechos como fundamento para imponer una sanción penal. Ello compromete a los actores con un debate en el que se puede aceptar o descartar las proposiciones a la luz de su valor de verdad. Cualquier desviación de ese ideal debería examinarse cuidadosamente y, en lo que se refiere al análisis del caso Sacayán, es necesario subrayar que el modo en que se expresa el voto mayoritario puede generar dos tipos de problemas para el control racional de los argumentos: *a)* Emotividad y cambios de creencias y *b)* entimemas y premisas implícitas.

a) Emotividad y cambios de creencias. Podría ocurrir que frente a un mismo fenómeno, e.g., un homicidio cometido de una cierta manera X, estuviésemos de acuerdo acerca de las características fácticas (i.e., acuerdo en las creencias), pero tuviésemos diferentes respuestas emocionales (desacuerdo en las actitudes) P⁹. En esos casos, es bien conocido que el uso de un lenguaje emotivamente teñido con una alta carga valorativa puede conducir a un cambio en las creencias. Ese cambio se produce porque sería irreconciliable mantener nuestra creencia ante el desagrado que nos provoca un determinado fenómeno. Así, no es infrecuente que un individuo que ha aceptado la proposición emocionalmente teñida «Quien hizo X es un monstruo» se niegue a aceptar que, por ejemplo, su hijo ha sido quien ejecutó X (proposición fáctica). De ese modo, un lenguaje emotivo puede llevar a una resignificación de los hechos sin un control racional de la verdad o falsedad de las afirmaciones (CARRIÓ, 1983: 304-305).

b) Un problema recurrente en la argumentación es pasar por alto las premisas que pueden fundamentar una cierta conclusión. En estos razonamientos entimemáticos, la conclusión no se sigue directamente de las premisas, sino de otras proposiciones que se encuentran presupuestas (COPI Y COHEN, 2013: 41-43)¹⁰. En el mejor de los casos, ello ocurre porque quien desarrolla el argumento está indudablemente convencido de la obviedad de ciertas circunstancias y considera que no es preciso ofrecer mayor y mejor defensa de sus afirmaciones. Pero, las convicciones no aportan ningún dato a favor de la veracidad de las proposiciones y es posible que la audiencia de un determinado discurso tenga diferentes puntos de vista acerca de ciertos hechos, o carecer de un acceso completo a la información relevante¹¹. Por ello, el fundamento

⁹ Para un desarrollo de estas ideas, véanse STEVENSON, 1944: 206-226; ROSS, 1963: 304-305.

¹⁰ También véase, por ejemplo, PÉREZ BARBERÁ Y BOUVIER, 2004.

¹¹ Sobre el error de confundir grados de convicción con grados de justificación epistémica LAUDAN, 2003; LAUDAN, 2005; LAUDAN, 2013b: cap. III.

de cada premisa tiene que identificarse de manera clara y presentarse en un lenguaje neutral (*i.e.*, apto para analizar su valor de verdad). De otro modo, quien estuviese dispuesto a negar la conclusión a la que arriba el tribunal —y no se percatara de las proposiciones implícitas— se ve forzado a impugnar solo aquello que se menciona en el argumento, aun cuando esas premisas, por sí mismas, sean plausibles —pero insuficientes para justificar la decisión¹².

El uso excesivo de adjetivos de alto contenido emotivo dificulta el control racional, pero no implica que las conclusiones sean necesariamente erróneas. Por ello, más allá del vocabulario desafortunado elegido por la mayoría del tribunal, es necesario examinar cuidadosamente el valor de verdad de sus afirmaciones.

3. EL VOTO DEL MAGISTRADO CALVETE

El fundamento de una sentencia será adecuado solo si se exponen y justifican las pruebas o indicios de los que se parte, se especifica la ley general de la que se parte y se muestra que las pruebas o indicios constituyen una instancia particular del antecedente de esa ley general (GASCÓN ABELLÁN, 2004: 220; BELTRÁN, 2011)¹³.

¹² Por supuesto, el contenido de las premisas implícitas puede ser de diferente tipo. Podrían ser probatorias, donde lo implícito, son las máximas de la experiencia o generalizaciones. También debemos recordar que el hecho de que una máxima de experiencia no se explicite no quiere decir de por sí que esté infundada en cuanto máxima. Una máxima implícita podría estar bien fundamentada. (Para un análisis de las máximas de experiencia, sus límites y fundamentos, véanse: TARUFFO, 1992; TARUFFO, 2002; GONZÁLEZ LAGIER, 2013: cap. II; TUZET, 2013: cap. XVI).

Pero las premisas implícitas también podrían ser definiciones encubiertas donde por ejemplo la dogmática o la jurisprudencia establece que, por ejemplo, «odio de género» es igual a «heridas brutales». Por último, y creo que es lo que sucede en este caso, hay premisas implícitas de ambos tipos. Se ha establecido que implícitamente «odio de género» equiparándolo a «heridas brutales» y se tiene por probada que las heridas fueron brutales sobre la base de máximas de experiencia implícitas cuyo fundamento es al menos no determinante. Durante el trabajo se tratará de mostrar estas ideas.

¹³ Un problema importante pero diferente al que se señala aquí es si, a pesar de todo, podría decirse que se satisface un estándar de prueba. Es decir, podría haber justificación epistémica (*i.e. prueba*) sin que exista estándar de suficiencia alguno. Si existiese un estándar y no se lo satisficiera, entonces aquí hay, efectivamente, un problema de justificación adicional. Hay tres cuestiones posibles. (i) El caso en que no hay justificación satisfactoria porque no se alegan las razones epistémicas pertinentes, una falla argumental. Se trata de un problema relativo al caso concreto, a la decisión judicial en cuestión. (ii) El caso en que no hay estándares de prueba, esto es, un umbral de suficiencia, un grado de justificación epistémica a partir del cual un enunciado *queda probado* (BAYÓN, 2009), de modo que las decisiones judiciales (aun cuando no injustificadas a la luz de la primera cuestión) son discrecionales respecto de la suficiencia de las razones epistémicas (pruebas) con que se cuenta. Este es un problema de carácter general, no relativo a la decisión en concreto. (LAUDAN, 2005; LAUDAN y SAUNDERS, 2009; FERRER BELTRÁN, 2010; LAUDAN, 2011; FERRER BELTRÁN, 2013; LAUDAN, 2013a; LAUDAN, 2013b; LAUDAN, 2016; FERRER BELTRÁN, 2018). (iii) Casos donde, habiendo por hipótesis un estándar de prueba no se lo respeta. Casos donde se considera *probado* un enunciado fáctico aun cuando no hay pruebas suficientes a la luz del estándar. Esta cuestión sería relativa a la decisión singular. La discusión relativa al punto (ii) no será considerada aquí.

El problema central del voto mayoritario es que no solo se despreocupa por hacer explícitos los indicios y las máximas de experiencia sobre la base de los cuales infieren que el delito se cometió con odio (la razón que justifica que siempre el tipo de herida cuenta como odio), sino que, además, no se plantean siquiera la posibilidad de que el odio sea un estado mental específico que debe ser probado en el caso particular, i.e. mostrar que Marino odiaba a Diana Sacayán por el género al que pertenecía. Esto último sucede incluso aunque los compromisos teóricos asumidos expresamente por el tribunal afirman que el odio es un elemento subjetivo añadido al dolo y, por tanto, debería probarse (STO NÚM. 4, 2018: 173).

Veamos con cierto detalle los argumentos desarrollados.

El magistrado Calvete comienza caracterizando el odio como una aversión a una persona o grupo de personas; luego caracteriza el género y por último la noción de odio al género o identidad de género. Afirma, por ejemplo:

El devenir de los tiempos y las reformas legislativas, ampliaron las hipótesis relacionados con el concepto de odio, el cual, luego de la modificación de la ley 26.743, que introdujo al homicidio por odio a la identidad de género, su art. 2 nos aportó una definición precisa, según la cual se entiende por identidad de género la vivencia interna e individual del género tal como cada persona la siente, la cual puede corresponder o no con el sexo asignado al momento del nacimiento incluyendo la vivencia personal del cuerpo (STO NÚM. 4, 2018: 171).

También afirma que los homicidios por odio están caracterizados por el odio o aversión que le genera la víctima por su condición de poseer determinada identidad de género. Así, señala que «cuando el sujeto activo mata a la víctima por su aversión a esas condiciones, constituye la motivación individual de corte psicológico...que pone en marcha la acción homicida» (STO NÚM. 4, 2018: 171). Y continúa:

La figura se completa con un elemento subjetivo distinto al dolo, en tanto el autor debe matar al sujeto pasivo por odio hacia algunas de estas circunstancias: el género, la orientación sexual, la identidad de género o la expresión de la identidad de género del paciente [sic]. Es un componente subjetivo diferente del dolo que repara en las motivaciones de la acción¹⁴.

En particular, en lo que respecta a los tipos de homicidios agravados por ser cometidos con odio hacia una persona travesti, denominado tanto por los acusadores como por el tribunal «travesticidio» (STO NÚM. 4, 2018: 174), afirma el magistrado que esta figura —a diferencia de lo que ocurre con el odio racial en que la manifestación de ese odio se suele manifestar de algún modo oral o escrito— esta no suele presentarse con estas manifestaciones sino que,

¹⁴ STO NÚM. 4, 2018: 173. Este párrafo es tomado literalmente del trabajo de Gustavo AROCENA (2014, 19). Tanto Arocena como los miembros del tribunal utilizan el término «paciente» para referirse al sujeto pasivo de la relación. En este sentido Arocena señala «a quien debe matar el sujeto activo es a un ser humano cuya muerte cause sufrimiento a la pareja o ex pareja de aquel. Este es, entonces el paciente en este delito». Dado que el término «paciente» es utilizado usualmente para referirse a una persona que requiere un tratamiento hubiese sido conveniente denominar al sujeto pasivo de la relación delictiva como «padeciente». Que de acuerdo al Diccionario de la RAE significa en su primera acepción «sentir física y corporalmente un daño, dolor, enfermedad, pena o castigo».

...puede presentarse acompañada de otros factores y/o con cierto contenido fáctico, en los que dicha señalización va de la mano con una agresión extrema y sostenida, acompañada de la descalificación de la víctima por su carácter de travesti o trans, con proyección en las especiales lesiones infringidas a esta (STO NÚM. 4,2018: 175).

Es decir, para el magistrado la manifestación más evidente del odio a una persona travesti es mediante las lesiones que se provocan en la víctima. Muestra de este odio es, para el Tribunal, el método elegido para causar la muerte, ya sean los golpes como así también el arma blanca para ocasionarle el deceso. Se resalta el hecho de que Diana recibió trece puñaladas (que, luego como indica la magistrada Bloch son heridas corto punzantes y no puñaladas) en distintas partes del cuerpo y con diferente profundidad, de las cuales dos de ellas fueron mortales. También se destacan los diferentes golpes recibidos en los brazos, en el rostro y en otras partes del cuerpo junto a cierta asfixia generada por las ataduras tipo mordaza. A todo ello, se añade el conocimiento que tenía el imputado de la víctima y

...la circunstancia de que el rasgo evidenciador del odio se explicitó nítidamente en la actitud asumida por Marino durante el evento, entre lo que se encuentra el lugar de asiento de las puñaladas, cuya ubicación no fue al azar sino que la multiplicidad de ellas estuvieron dirigidas a lugares específicos relacionados con la asignación del sexo y sus tributos más definidos como son el rostro, los pechos y los glúteos (STO NÚM. 4,2018: 176).

Para el Tribunal, estos rasgos específicos son la muestra de que el homicidio fue cometido con odio hacia la identidad de género (travesti). A continuación, señalaré mis observaciones críticas a esta argumentación.

a) En primer lugar, está fuera de discusión que la *correcta* aplicación del agravante por odio a la identidad de género exige que el enunciado ‘Marino mató a Sacayán por odio a la identidad de género’ sea verdadero¹⁵. Sin embargo, no hay un análisis que pretenda mostrar que Marino poseía un determinado estado mental al momento de cometer el homicidio. Como muestra de esta despreocupación, por ejemplo, los fundamentos de la sentencia solo se refieren a las pericias psicológicas destinadas a probar que el sujeto comprendía la criminalidad del acto y estaba en condiciones de someterse al juicio (STO NÚM. 4,2018: 191-192). Se extraña un análisis más detallado sobre si Marino experimentaba algún tipo de odio hacia alguna clase de persona. Por supuesto, también podría establecerse que Marino en general odiaba a cierta clase de personas (por ejemplo travestis) y todavía ello sería insuficiente para justificar que Marino mato a Diana Sacayán *por el odio* que poseía a esa clase de personas. Es decir, para que el tribunal pueda aplicar correctamente la norma del 80 inc. 4 del código penal argentino debería mostrar que Marino mató a Sacayán por

¹⁵ Esta frase debe entenderse en un sentido amplio. Existe una extensa discusión sobre si para que exista la aplicación de una norma exige que el enunciado haya sido probado o que este sea verdadero. Aunque es una discusión muy interesante creo que no plantea diferencias para el problema que trataré aquí. Para esta discusión, véanse, por ejemplo: CARACCILO, 2000; SÁNCHEZ BRÍGIDO y SELEME, 2002; BOUZAT y CANTARO, 2003; TARUFFO, 2003, FERRER BELTRÁN, 2011; FERRER BELTRÁN, 2001; FERRER BELTRÁN, 2006. Para un resumen de este debate, véase DEI VECCHI, 2013.

el odio que le tenía y que la razón de este odio estaba ocasionada por la identidad de género de Sacayán. En la sentencia se extraña un análisis más pormenorizado sobre las razones para fundamentar ambas tesis¹⁶. Es probable que, si la explicación de la conducta se debe al odio hacia las personas trans o travestis, este sentimiento se mantenga más allá de la acción individual analizada, es decir incluso cuando se ha dado muerte a uno de los integrantes de la clase odiada. Así, podemos medir la intensidad de una emoción mostrando cuánto del comportamiento de la vida del agente puede explicarse conforme a esa emoción¹⁷. En otras palabras, acreditar que se actúa con odio a la identidad de género tiene que ser compatible con la *historia* del sujeto (VON WRIGHT, 1984: 66)¹⁸. Es preciso insertar la conducta en un relato más amplio que le confiere significado y permite su explicación. Por supuesto, la justificación de esa explicación normalmente requiere incorporar mayores detalles y, por ello, aunque ella aún no muestra que la acción en concreto se haya realizado por el odio, puede tomarse como un síntoma (o indicio tipo) a tener en cuenta. Por ejemplo, quienes pertenecían al Ku Klux Klan seguían odiando a las personas de origen afroamericano, incluso aunque ya hubieran cometido una barbarie hacia alguno de sus miembros. La mera pertenencia a un grupo tampoco justifica por sí sola la imputación, pero ello, aunque es todavía compatible con que el homicidio en particular no haya sido ocasionado por el odio que le tenía a cierta clase de persona, puede empezar a dar pautas que nos ayuden a *explicar* (mostrar por qué la conducta era inevitable a la luz de las creencias y deseos del agente) su conducta¹⁹. El hecho de que, al menos en

¹⁶ También podría suceder que Marino odiase a Diana Sacayán por alguna razón diferente a su identidad de género y que la matase por ello. Esto sería irrelevante para agravar su conducta. Nuestro código penal limita las razones por las cuales el odio es considerado una agravante. Así, por ejemplo, el odio entre Montescos y Capuletos en la tragedia de Shakespeare sería irrelevante para agravar la conducta. Para una discusión de por qué el Estado elige ciertos motivos véanse: DIAZ LOPEZ, 2013; KAHAN, 2001.

¹⁷ Por supuesto, este no es el único criterio para medir la intensidad de una emoción. Hay otros (como la violencia de los cambios físicos) que incluso pueden entrar en conflicto. Sin embargo, este problema no surge en los hechos analizados porque los mismos son ignorados completamente. Para un desarrollo desde el punto de vista filosófico de la intensidad y permanencia de algunas emociones como motivo véase: KENNY, 2004: 24-34.

¹⁸ No estoy afirmando aquí que si no ha habido datos en su historia de vida de conductas que indiquen odio hacia una clase no puede haber de ninguna manera odio. Así, no estoy afirmando que, por ejemplo, el primer acto de un skinhead no puede contar como odio porque no tiene una historia de vida de odio. Podríamos pensar que aunque no ha actuado violentamente con anterioridad su pertenencia a ciertos grupos puede explicar por qué actuó de esa manera. Por supuesto, todo ello es compatible con que en algún caso una persona X que pertenece a un grupo skinhead ha matado a una persona judía y no ha habido odio. Estas son las grandes dificultades que enfrentan quienes investigan y quienes aplican las normas. En síntesis, solo quiero decir que la historia de vida puede ser un indicador más y que, en ocasiones, para entender qué hizo el sujeto en el momento del hecho hay que ampliar el panorama para poner en contexto su comportamiento.

¹⁹ Existe una gran discusión acerca de si los estados mentales deben probarse o si ellos solo deben atribuirse. Esta discusión es central para entender qué debemos probar y el modo de hacerlo. Sin embargo, tal como afirmé más arriba, el Tribunal parece asumir de manera expresa al caracterizar al odio que el mismo debe probarse. Véase GONZÁLEZ LAGIER, 2005: Cap V.

los fundamentos de la sentencia, no surja un claro interés por el estado psicológico de Marino, o por mostrar que él ya había realizado otras conductas explicadas por el odio, da una pauta de que al Tribunal no consideró como relevante el efectivo estado mental del sujeto al momento de actuar. En resumen, aunque está claro que Marino mató a Sacayán, el argumento del tribunal es insuficiente para mostrar que lo hizo por odio a la identidad de género.

b) En segundo lugar, afirma que las manifestaciones de odio, a diferencia de lo que ocurre con el odio racial donde ellas suelen manifestarse de modo oral o escrito (STO NÚM. 4,2018:175) se revela en la extrema violencia con la que se comete el hecho. Sin embargo, no se ofrecen argumentos que permitan corroborar que en general que el odio a la identidad de género no puede manifestarse de manera oral o escrita similar a los delitos a la identidad racial o religiosa. Esto podría parecer una cuestión de detalle carente de importancia, sin embargo, esta idea le sirve al tribunal para descartar que deberían buscarse otras manifestaciones del odio diferentes a la esgrimida en las heridas de Sacayán. El argumento del voto mayoritario sería «Dado que el odio a la identidad de género no se puede manifestar de manera verbal o escrita entonces el único criterio que podemos considerar como manifestación del odio son las heridas infringidas en el cuerpo de Sacayán». Esa afirmación parece destinada a confirmar lo que ya se ha presupuesto (*post hoc, propter hoc*). En otras palabras, como en la reconstrucción de los hechos solo se asume que el hecho fue cometido con «extrema violencia» —y no mediante otras manifestaciones verbales o escritas— se sostiene que la característica de los delitos de odio por identidad de género es esa violencia. Si bien es verdad que frecuentemente el odio religioso o racial se muestra en una simbología característicos como las esvásticas, nada impide que el odio hacia la identidad de género se hubiese manifestado de modo verbal. Así, en contra de lo que sostiene el tribunal, y afirmando que los delitos al odio a la identidad de género *pueden* manifestarse de manera oral o escrita podría señalarse que dada la ausencia de otros elementos verbales o simbólicos característicos, habría que descartar al odio de género como agravante del homicidio.

c) Con respecto a las heridas, vale la pena hacer unas cuantas aclaraciones. El tribunal parece tener varios objetivos conjuntos. Uno, consiste en mostrar que el hecho fue cometido con extrema violencia. Esto es que las heridas punzo cortantes y el lugar en que ellas fueron infligidas hace que el hecho sea identificado como de extrema violencia. Otro objetivo que también persigue el voto mayoritario es señalar que esa extrema violencia es inequívocamente prueba del odio hacia una clase determinada de personas²⁰. Conforme a si las heridas eran del tipo que contarán como «violencia extrema», el magistrado «extrae» el odio del análisis de las heridas infligidas.

²⁰ Podría ser que el tribunal afirmara de manera expresa que el odio a la identidad de género está definido por la extrema violencia manifestada en las heridas. Creo que a fin de cuentas esta es la idea del voto mayoritario del Tribunal, pero ello se ve opacada por los compromisos teóricos que asumen al comienzo y el rastro de querer mostrar que el odio está probado. Para una análisis de cómo el trasfondo teórico impacta en los hechos que pretendemos probar, véase GONZÁLEZ LAGIER, 2007.

Pero, si la naturaleza de las heridas es el criterio utilizado para determinar el odio entonces no alcanza con afirmar que fueron trece puñaladas y en lugares simbólicos del cuerpo. Debería hacer un análisis pormenorizado de cada una de las heridas, su tamaño y el nivel de violencia. Esta tarea es realizada por la magistrada Ivana Bloch, y en ello se ve que las heridas mortales fueron dos y que no hay nada que vaya más allá de la violencia ejercida en homicidios cometidos con apuñalamiento. Es decir, la magistrada realiza un análisis pormenorizado de las heridas para negar la tesis de que las heridas que poseía Diana eran muestra de “violencia extrema”.

En lo que concierne a que la violencia extrema es prueba inequívoca del odio a la identidad de género, creo la fundamentación de la sentencia parece desplazar la discusión sobre la motivación específica de Marino hacia los criterios genéricos para identificar a la clase de delitos cometidos con odio de género. En este caso, ya no se analiza la subsunción individual (Marino mató a Sacayán por odio de género) sino otro enunciado más abstracto (subsunción genérica) «La clase de eventos C cuenta como un caso de odio a la identidad de género»²¹. En la medida en que este segundo enunciado establece una definición completa, suministra las condiciones necesarias y suficientes para delimitar el sentido de la expresión en cuestión. El argumento del voto mayoritario sería:

i) Una regla que establece: Matar a una travesti con violencia extrema *cuenta como* matar con odio a la identidad de género.

ii) La afirmación de hecho de que Marino mató a una travesti con violencia extrema

iii) La conclusión de que Marino mató con odio a la identidad de género.

Mientras que el enunciado individual es una relación entre un objeto y un predicado, que predica la *pertenencia* de un individuo a una clase, el segundo enunciado establece una relación de *inclusión* entre clases. Más allá de esas diferencias formales, el enunciado genérico funciona como una *definición*, *i.e.*, establece un criterio para determinar el valor de verdad del primero ya que nada contará como odio a la identidad de género si no satisface el criterio establecido en el enunciado de subsunción genérica.

Ahora bien, incluso aunque asumamos que la muerte de Diana Sacayán fue ocasionada mediante violencia extrema, y que el odio se manifiesta mediante la violencia extrema no parece haber indicios que permitan considerar a dicha violencia efecto de un odio específico a la identidad de género. Así, es compatible con que el homicidio se haya cometido por el odio hacia la identidad género, por el odio hacia la persona concreta a que se da muerte independientemente de la clase de persona a la que pertenezca, o por el simple hecho de que era su enemiga, etc. Discriminar entre estas

²¹ Para el desarrollo de estas ideas, véase, por ejemplo: ALCHOURRÓN Y BULYGIN, 1991: 267 ; GIMBERNAT ORDEIG, 1999: 36-38.

causas requeriría consideraciones adicionales. Imputarle a Marino la interpretación más gravosa viola el principio de defensa.

Como cuestión adicional se afirma que el método elegido es una muestra del odio, pero no se da una razón para sostener esa idea²². Más aún, el hecho de que el homicidio haya sido cometido con un cuchillo que se encontraba en la casa de Diana es compatible con un curso de acción más improvisado que planificado.

En síntesis, incluso si el tribunal tuviese razón en que las heridas fueron brutales esto siempre debe ser entendida como síntoma (indicio) y no como criterio (establecido en términos de condiciones necesarias y suficientes) para determinar el odio. Es decir, si queremos respetar las garantías del imputado deberían buscarse más datos característicos del odio y no solo asumir que por definición las heridas brutales significan odio²³. De otro modo, esta reconstrucción corre el riesgo de reducirse al ensañamiento²⁴. Sin embargo, esta opción no es la asumida por el código penal argentino ya que ha identificado diferentes situaciones para estos casos, aunque ambas generen idénticas consecuencias normativas. Al incorporar el agravante por odio a la identidad de género no ha querido establecer que los homicidios por odio son generalmente brutales (aunque por hipótesis incluso podamos asumir que lo sean) sino que enfatiza como más grave el motivo por el cual se cometen (PERALTA, 2013: 6).

Hasta aquí he expuesto las razones que da el presidente del tribunal y constituyen el voto mayoritario. También he señalado algunos rasgos de ese voto que podrían cuestionarse²⁵. Al margen de ello, sigamos con la exposición de los motivos de los otros miembros del tribunal. Veamos ahora brevemente cuáles son los argumentos esgrimidos por Julio Báez.

4. EL VOTO DEL MAGISTRADO JULIO BÁEZ

El extenso voto de este magistrado coincide con el del presidente del Tribunal, pero pretende añadir argumentos a los ya sostenidos por Calvete para justificar el homicidio agravado, tanto por el odio como por la violencia de género (art. 80 inc. 4 y 11 del Código penal argentino). La tesis de Báez es la siguiente:

²² Luego, a través del voto de la magistrada Bloch se logra reconstruir por qué les parece relevante.

²³ Para la distinción entre síntoma y criterio, véase, por ejemplo, GARCÍA CARPINTERO, 1996: 477.

²⁴ El tribunal también afirma que las heridas no muestran que haya habido ensañamiento porque ellas no reflejan la intensidad y la duración de la acción propia del ensañamiento (STO NÚM. 4, 2018: 189 y ss)

²⁵ Hay una opción teórica que no he considerado aquí y es la cuestión de quienes niegan por distintas razones que los estados mentales deban probarse. Este es un problema interesante que he tratado de criticar en otros trabajos, siguiendo las ideas de González Lagier. Véanse: GONZÁLEZ LAGIER, 2003 y MANRIQUE, 2012.

Sin embargo, por el modo en el que argumenta el presidente del Tribunal no creo que esta sea la opción teórica asumida por él.

En el examen concreto de la conducta de Marino, se desprende de los hechos probados que la conducta abusiva es, inequívocamente, un hecho de violencia de género, así definido por la normativa internacional y nacional y que la muerte de Sacayán se presenta directamente determinada por ese acto (STO Núm. 4, 2018: 258-259)²⁶.

Este magistrado concuerda con el voto del juez Calvete en que la muestra del odio hacia Diana Sacayán se encuentra en el tipo de heridas que les fueron provocadas (STO Núm. 4, 2018: 198-199).

...estimo que la latitud de las lesiones alojadas en zonas tan sensibles para la mujer, con indiferencia en su incidencia letal o no- es un claro indicador artero que denota en Marino un irrefutable odio a la identidad del género de su agredida (STO Núm. 4, 2018: 198-199)²⁷.

Y continúa,

La *crueldad del ataque* se enderezó o acometer directamente la identidad transexual de Sacayán. Las lesiones alojadas en dichas partes, aunadas a que las mismas fueran producida no solo en la morada de esta, sino que *también se gestaron en el dormitorio* del predio hablan a las claras del contorno sexual y el odio que poseía Marino respecto de esta situación que excedía la relación binaria; *el conjunto homicida habla a las claras que el anhelo de finiquitar la vida de Sacayán tuvo de manera tangencial su calidad de militante en un colectivo definido* (STO Núm. 4, 2018: 199).

Marino y su *consorte conocían de manera extendida la calidad transexual de la occisa*, su notoria militancia y, teniendo en cuenta ello, no evidenciaron una posición terrecida, por el contrario, acometieron contra ella, de la forma brutal explicada en detalle por el colega que lleva la voz (STO Núm. 4, 2018: 199)²⁸.

De este análisis surgen interrogantes, parcialmente diferentes a los que motivaron el análisis del voto de Calvete.

a) En primer lugar, ¿cuáles son los hechos que reflejan que el ataque fue cruel? Ya he señalado anteriormente (sección 2), los peligros que conlleva el uso del lenguaje emotivo para el control racional de los argumentos. Ahora solo añadiré que no hay, al igual que sucedía con el voto del magistrado Calvete, una explicación de qué convierte en 'cruel' la provocación de una determinada herida (es decir, no se hace explícito el criterio necesario para la aplicación de ese predicado).

Por supuesto, esta insuficiencia, por sí misma, no muestra ausencia de crueldad, pero al igual que ocurre con predicados que conllevan un agravamiento en la atri-

²⁶ Al igual que el voto en disidencia de la magistrada Ivana Bloch coincido en que no se realizó ningún esfuerzo por probar que hubo alguna acción que muestre que se mató a Diana Sacayán por el hecho de ser mujer. Además, un punto con el que ningún miembro del tribunal se ocupa de tratar para aplicar el agravante de violencia de género es preguntarse si Diana se auto percibía como mujer o como de identidad trans. Si el caso fuera este último la agravante por femicidio tal como lo entendió el tribunal no se aplicaría, ya que abarca solo a las mujeres (y no a otro colectivo de personas igual o más vulnerable).

²⁷ STO Núm. 4, 18/06/2018, p. 198.

²⁸ Llama la atención que el magistrado le impute «al consorte» de Marino el *conocimiento* del carácter transexual (*sic*, travesti) de Sacayán cuando esta persona no solamente no fue juzgada, sino que tampoco imputada porque ni siquiera se sabe quién es.

bución de responsabilidad, el tribunal debería haberse preocupado especialmente en hacer explícitos los criterios de aplicación de ese concepto.

b) En segundo lugar, para este magistrado, el tipo de heridas ocasionadas y que se haya matado a Sacayán en su hogar y el lugar de la casa donde estas fueron realizadas no solo muestra que hubo odio en la conducta de Marino, sino que también prueba que él quería matar a Diana por pertenecer y militar en el colectivo travesti/trans.

En párrafos anteriores hice referencia a las heridas y aquí solo haré hincapié en la segunda parte de su argumento. Conforme al argumento de Báez, el odio de género se muestra no solo en que el homicidio fue cometido en el hogar de la víctima, sino que le añade valor negativo que el hecho se haya producido, al menos en parte, en la habitación de la víctima. Esta afirmación resulta, curiosa ya que no hay una explicación específica que intente sostener tal idea. Existen dos problemas vinculados entre sí y relacionados con las máximas de experiencia que parece haber incorporado el vocal en su argumento. Estas máximas realizan la conexión entre el hecho de la muerte y el motivo por odio (TUZET, 2013: cap XVI).

En primer lugar, no existe una explicitación de las máximas de la experiencia tenidas en cuenta²⁹. Esta carencia hace más difícil el control de su razonamiento. Dado que esas máximas pueden fundarse en leyes científicas, leyes probabilísticas, meras generalizaciones o prejuicios de diferente clase, sin una explicitación de ellas no podemos estar seguros de a qué ámbito pertenecen las utilizadas por el magistrado e identificar el mejor modo de controlarlas o corregirlas. (Los medios de control y corrección pueden ser la verificación de la fiabilidad de las máximas o la verificación de la aceptación intersubjetiva de ellas. Véase TUZET, 2013: cap. XVI).

En segundo lugar, si asumimos que la máxima de la experiencia es algo así como «aquellos homicidios que se cometen en el hogar de la víctima, y en particular en su habitación suelen reflejar un odio a la identidad de género» no parece haber razones epistémicas que fundamenten la máxima. Más bien, si se incorpora ese dato como relevante, podrían elaborarse una serie de conjeturas alternativas plausibles (*e.g.*, que había una relación mínima de confianza entre víctima y victimario), pero nada de ello fue evaluado y descartado en el debate. Por otro lado, si tomamos en consideración los rasgos facilitados por el informe de la CIDH sobre violencia contra las personas LGTBI puede fundarse una máxima opuesta que afirma que ciertos homicidios de odio suelen producirse en descampados y contra personas que el victimario no conocía.

De la reconstrucción que ofrece el magistrado Báez no se desprende que Diana haya sido torturada, sus partes amputadas, su cara incinerada, su cuerpo descuartizado o arrojado con desprecio en algún lugar, que son los indicios que recoge el Informe de la Comisión Interamericana sobre violencia contra las personas LGTBI,

²⁹ Las máximas de la experiencia planteas problemas o discusiones centrales por sí mismas. No me detendré aquí en este punto. Para ello véase TUZET, 2013.

en particular, los párrafos 107 y 108.³⁰ Este informe se centra en los actos de violencia contra personas que manifiestan diferentes identidades, orientaciones sexuales u expresiones de género diversas a los estándares habituales de masculino y femenino en América (STO Núm. 4, 2018:199-200)³¹. Si bien este informe recoge muchos casos donde personas que pertenecían al colectivo LGBTI son víctimas de delitos de odio, el informe por sí solo no muestra nada en el caso en concreto. Este puede ser una guía útil —de hecho, creo que es muy útil—, pero se requieren argumentos que conecten la utilidad del informe con el caso de Diana Sacayán.

Báez afirma que está convencido de que el odio a la identidad de género se puede encontrar en «su abominación sobre el cuerpo transexual de Sacayán». Como prueba de su tesis ofrece un excursus sobre el cuerpo como entidad social, cultural y política, afirmaciones sobre el modo en que la sociedad patriarcal influye e incorpora sus normas, estereotipos y expectativas a través de los cuerpos de hombres y mujeres, etc. En este panorama, la cosificación del cuerpo femenino y transexual forman parte de las relaciones de dominio que propone el patriarcado (STO Núm. 4, 2018:200-204)³².

Para el magistrado, Marino formaba parte de este entorno ‘cosificador’ porque se abastecía de manera cordial con quienes le suministran dinero, pero reaccionaba con odio ante la negativa a otorgarle ese suministro y este odio se potencia por la pertenencia de Diana Sacayán al colectivo donde militaba (STO Núm. 4, 2018:203). Báez afirma que Marino, de alguna manera, cosifica a las personas porque se relaciona con ella con el interés de beneficiarse económicamente. Esto parece verosímil, según surge de los testimonios incorporados en la causa. Sin embargo, ello muestra que Marino instrumentalizaba a las personas en general, y no solo a aquellas que pertenecen a un colectivo determinado; tampoco muestra que odiase ni a las personas en general ni a un colectivo en particular. Pero incluso si asumimos que ello es así, ¿qué hechos muestran que Marino mató a Diana Sacayán *porque* pertenecía al grupo LGBTI y/o por su liderazgo como militante en el grupo? En el voto de este magistrado, al igual que sucede con Calvete, no hay un dato independiente además de las heridas infligidas que explique de manera unívoca el odio de Marino hacia la identidad de género de Diana Sacayán.

³⁰ El párrafo 107 habla de los altos niveles de crueldad de crímenes contra personas pertenecientes al colectivo LGTBI. El párrafo 108 afirma: «En el Registro de Violencia, referido anteriormente, existen numerosos ejemplos de homicidios particularmente crueles, incluyendo casos de personas lapidadas, decapitadas, quemadas y empaladas. Muchas víctimas son repetidamente apuñaladas o golpeadas hasta la muerte con martillos u objetos contundentes. Otras reciben puñetazos o patadas hasta su muerte, les arrojan ácido o son asfixiadas. Algunas de las víctimas en el Registro fueron reiteradamente atropelladas por carros, mutiladas o incineradas. En muchos casos, las víctimas fueron asesinadas luego de ser sometidas a horribles actos de tortura, tratos inhumanos o degradantes, y múltiples formas de extrema humillación, degradación, y violación».

³¹ Este informe es reconstruido en el alegato del Fiscal General, STO Núm. 4, 2018: 65-70.

³² De nuevo, creo que las afirmaciones de este vocal incluso cuando son verdaderas ellas son irrelevantes para el cuerpo de una sentencia.

El magistrado continúa sus afirmaciones generales. Afirma:

La precariedad en las normas de género se ve de manera palmaria en la tirria que Marino evidenció por Sacayán al momento del suceso. Este extremo se corresponde con el burdo y petiso escalón de la deficiencia en que viven las integrantes del colectivo transexual quienes están expuestos a un elevado riesgo de maltrato, de patologización y violencia (Butler...)

La actividad de Marino —perfilada en su inquina de género— nos enfrenta a otro dilema cruel: a algunos los reconocer como seres humanos y a otros, de la condición de Sacayán, los percibe como si no lo fuera. (Butler...).(STO Núm. 4, 2018:203)³³.

Para Báez fue obvio que el ataque no se dirigió a una persona que consideraba un igual, sino que Marino «vio la luz respecto de una mujer, líder de un colectivo desprotegido; este colectivo, observa dicha precariedad no solo en la discriminación desigual entre la población sino, también, respecto de un segmento de esta cuyas vidas se ven expuestas con mayor facilidad (Butler...)» (STO Núm. 4, 2018:204).

El juez concluye sus argumentos afirmando que:

La inmediatez permitió demostrar —en función de la prueba relevada por el colega— que Marino efectivizó un ataque compartido contra una mujer que —aunque reconocida y líder de una agrupación— se afiliaba en un segmento de la población de vida precaria y subyugada, donde el compromiso institucional de cuidado se ha vuelto laxo, desatendido y con una exposición a la muerte o a la desaparición prematura que puede asemejarse a una suerte de racismo sistemático o abandono calculado (Butler...). (STO Núm. 4, 2018:204)³⁴.

Tal vez, algunas de estas asunciones pueden explicarse (aunque no justificarse)³⁵ por el objetivo de proteger a un colectivo que ha sido históricamente y, sigue siendo, muy vulnerable. Es un hecho que las personas trans/travestis cuando mueren asesinadas mueren más jóvenes que el resto de personas y también son asesinadas brutalmente más a menudo que el resto³⁶. Estas serían dos máximas de la experiencia

³³ . Es curioso que para fundamentar sus afirmaciones acerca de Marino cite a Butler. Ello es curioso, no porque Butler no sea un referente importante en la discusión sobre género sino porque no tiene nada que ver con el caso en concreto. Para mostrar la conexión el magistrado debería usar las ideas de Butler pero mostrar hechos/información *verdadera* recogidos en el debate.

³⁴ STO Núm. 4, 18/06/2018, p. 204. El voto del magistrado en lo que respecta al fundamento del agravante por odio de género continua algunas páginas más. Sin embargo, no creo que el contenido de las páginas siguientes añada algún dato positivo para justificar la sentencia. Así, por ejemplo, no es relevante las acciones de eugenesia que realizaban los nazis, o si el magistrado escribió un artículo afirmando la igualdad de los derechos y tolerancia (sic) que debemos predicar hacia el colectivo trans, o si hay un derecho procesal de dos velocidades, o la utilidad del lenguaje natural.

³⁵ Para la diferencia entre explicación y justificación, véase HEMPEL, 1996: 332 y ss.

³⁶ Existe una idea generalizada de que las mujeres trans/travestis tienen un promedio de vida de 35 años. Sin embargo, ello es equivocado. La confusión surgió del informe de la CIDH en la que se analiza los casos de 594 homicidios cometidos contra algún miembro (o se percibió como miembro) del LGBTI en donde 282 de los homicidios eran sobre personas trans/travestis y el 80 por 100 de ellas tenía menos de 35 años al momento de morir. Ello muestra que las mujeres trans/travestis, cuando son asesinadas en general tienen menos de 35 años pero no que el promedio de vida general sea de 35 años. Véase el informe de la CIDH párrafo 16. Al resaltar este malentendido no quiero negar de ninguna manera que el colectivo LGBTI, y en particular las personas trans/travestis, sean especialmente vulne-

basadas en estereotipos descriptivos asumidos por el magistrado. Por supuesto, no niego en absoluto que esta sea un estereotipo con base cognitiva, sino que pretendo enfatizar que esa información no sustituye una tarea fundamental para la justificación de una decisión: el juez tiene que mostrar que esa generalización también es verdadera para el caso individual de la muerte de Diana Sacayán³⁷. Sin embargo, el argumento parece asumir que la generalización encierra una verdad innegable. Por ello, se argumenta del siguiente modo: Las personas trans/travestis mueren jóvenes. Cuando son asesinadas (que es una importante razón por la cual mueren jóvenes), ello es por cuestiones de odio. Dado que Sacayán era una persona travesti, entonces murió asesinada por cuestiones de odio.

En la medida en que las generalizaciones no representan verdades universales, el argumento es falaz. Es decir, esta pretensión de querer proteger al colectivo, no muestra que Sacayán haya sido una persona en extrema vulnerabilidad³⁸. En resumen, la utilización de este estereotipo es perjudicial para el imputado porque transforma en un sinsentido (una contradicción en sus propios términos) a la proposición «Marino mató a Sacayán por ensañamiento, pero no movido por el odio de género». Más allá de la verdad o falsedad de esa afirmación, la misma tiene un sentido completo y nos conduce a analizar datos empíricos que pueden probar adecuadamente qué hizo un cierto sujeto y cuáles fueron sus motivaciones. Por el contrario, nada de eso tendría sentido si se asume una conexión analítica entre «matar a una persona trans/travesti» y «matar por odio de género».

5. EL VOTO DE LA MAGISTRADA IVANA BLOCH

Como se comentó más arriba la magistrada realiza un voto en disidencia. Si bien coincide en que el hecho fue agravado y por eso acuerda en la atribución de responsabilidad penal, la magistrada encuadra el hecho en el agravante por la relación de pareja (artículo 80 inc. 1 del código penal argentino)³⁹. En otras palabras, la magistrada Bloch niega que se haya probado que en el homicidio de Sacayán haya habido odio a la identidad de género y también niega que existiera o mediare en el hecho violencia de género i.e. inaplicabilidad del inc. 11 del artículo 80⁴⁰.

rables. Para el desarrollo de esta idea, véase: <https://magnet.xataka.com/preguntas-no-tan-frecuentes/personas-trans-no-tienen-esperanza-vida-inferior-a-35-anos-digan-medios>.

La discusión sobre prueba estadística es inmensa y no pretendo reconstruirla aquí más allá de este comentario intuitivo. Para ello, véanse: COHEN, 1977; SHAFER, 1988; TILLERS y GREEN, 1988; COHEN, 1989.

³⁷ Para la distinción y su diferente dirección de ajuste de estereotipos descriptivos y normativos, véanse, ARENA, 2016: 51-75; ARENA, 2019: 11-44 o ARENA, 2018: 5.

³⁸ Sobre la irrelevancia de generalizaciones estadísticas para el caso individual véanse: COHEN, 1977: 74; COHEN, 1998 (1989): 58; FERRER BELTRÁN, 2007: 98; GASCÓN ABELLÁN, 2004: 170.

³⁹ Creo que también se podría discutir la atribución de esta agravante, sin embargo, no me detendré aquí en ello.

⁴⁰ No me detendré aquí en el desacuerdo interpretativo entre Bloch y la acusación acerca del fundamento de la agravante, o la cuestión sustantiva de si reprochar el odio es castigar el carácter de una perso-

En general, coincido con lo afirmado por la magistrada sobre la agravante de violencia de género. Por ello, en esta sección me dedicaré a reconstruir su argumento con respecto al odio a la identidad de género ya que a partir de ese análisis se puede construir una hoja de ruta para analizar la aplicación de esa agravante.

El desacuerdo general con el tribunal no se encuentra en cuáles son los indicadores que podrían ayudar a entender si la conducta fue cometida con odio, sino más bien en la evaluación que hace el tribunal de esos indicadores. Así en palabras de la magistrada:

...en primer lugar debo remarcar que coincido con la elección de todos los indicadores que en abstracto mencionaron los acusadores para poder determinar la relación entre el motivo y el hecho. En lo que disiento es en la valoración concreta que se ha realizado de estos indicadores en relación con el supuesto de hecho. En segundo término, señalaré que así como no está probado el elemento 'odio a la identidad de género' también puede pensarse en hipótesis rivales —a las que luego aludiré— con igual o mayor peso que la planteada por la acusación, situación que impide nada menos que dar por acreditado un elemento que agrava la figura básica y la conmina con la mayor pena. Esto no puede ser tomado livianamente (STO NÚM. 4, 2018:329-330).

Dado que el voto del tribunal insiste en el peso o el valor que poseían las heridas de Diana Sacayán como muestra del odio con que fue asesinada, la magistrada insiste en el testimonio de la experta Amaranta Gómez Regalado y en el informe contra la violencia contra LGBTI de la CIDH para mostrar que no está probado que Marino haya actuado con odio a la identidad de género. Aunque tanto la acusación como el tribunal tuvo en cuenta el informe de la CIDH, según la magistrada, no tuvieron en cuenta alguna información relevante de ese documento (STO NÚM. 4, 2018:330-331).

Recordemos, un rasgo que utiliza el tribunal como indicador del odio es, siguiendo el informe de la CIDH, el grado de violencia con el que fue cometido el homicidio. En particular, utilizan el párrafo 108 que afirma que: «*Muchas víctimas son repetidamente apuñaladas o golpeadas hasta la muerte con martillos u objetos contundentes. Otras reciben puñetazos o patadas hasta su muerte, les arrojan ácido o son asfixiadas...*».

La magistrada va a negar que en la muerte de Sacayán exista ese tipo de violencia. En primer lugar, muestra que desde la elevación a juicio se viene repitiendo el error de que Marino le dio entre 11 y 13 puñaladas a Diana Sacayán. Sin embargo, la magistrada se esfuerza en mostrar que en realidad Diana recibió dos heridas profundas y que el resto de heridas cortantes y golpes son compatibles con el actuar defensivo. Por supuesto, la magistrada no intenta negar que fue un hecho violento, más bien enfatiza que no tuvo el nivel de violencia característicos de los homicidios por odio⁴¹.

na. Estos son temas relevantes y van a incidir en el modo en que se exige o no determinada prueba, pero no es allí donde se manifiesta el desacuerdo con el resto de los jueces. STO NÚM. 4, 2018: 285- 324.

⁴¹ Tal como lo afirma la magistrada siguiendo el informe de la CIDH cuando se habla de este tipo de indicadores se está pensando en hechos donde se torturó a la persona, se la decapitó, quemó o empaló. Para el desarrollo y muestra de este punto véase, STO NÚM. 4, 2018:333-335.

Para ello también cita varios casos paradigmáticos de delitos por odio, como el caso Zamudio en Chile⁴². Sostiene que Diana Sacayán no fue torturada y que la mordaza era un mecanismo para silenciarla y no de tortura.

Sostiene que el contexto tampoco forma parte de los delitos de odio, como podría ser, por ejemplo, la muerte de una persona trans a manos de un desconocido en un callejón oscuro porque estaba vestida de determinada manera, sino que es compatible con un contexto en el que se comete un delito contra la propiedad, una pelea por drogas, etc. Por supuesto, ello no excluye que Marino pudiese cometer un delito de odio pero las circunstancias de su ejecución parecen compatible también con otro tipo de explicaciones (STO NÚM. 4,2018:338-339).

La magistrada concluye:

Lo que simplemente se dice es que cuando los acusadores citan el informe para afirmar que un hecho como el presente es de aquellos en los que allí se consideran como casos inequívocos de odio: o bien no hacen explícita la envergadura de los hechos considerados allí como especialmente crueles o bien cambian las circunstancias de hecho del presente caso para que sí se adecue (STO N° 4,2018: 338).

Otro indicador del odio, tomado muy en cuenta por querrela, Fiscal y el magistrado Báez, es el lugar del cuerpo en el que fueron realizadas las heridas. Estas serían relevantes en cuanto están dirigidas a lugares representativos de la identidad de género. La magistrada sostiene que esos hechos también han sido tergiversados para hacerlos encuadrar en la calificación de odio de género.

Para decirlo de manera contundente, la magistrada suscribe que la violencia extrema es indicio del odio de género pero niega que en el caso pueda afirmarse que haya la violencia asociada a los delitos de odio. Es decir, niega el indicio, ya sea porque el voto mayoritario no hace explícito los hechos o porque los tergiversa.

Así, afirma la magistrada, que la autopsia indica que «no se han detectado lesiones violentas en zonas genitales» además su cuerpo tampoco fue expuesto de manera particular. Y, en relación con las heridas en las mamas, la magistrada recurre nuevamente a la autopsia para mostrar que posee en la mama izquierda una excoriación de 1 cm producido por un golpe y en la mama derecha hay una herida punzo cortante de 1 cm. El tamaño de las heridas no parece revelar que se hubiera querido negar rasgos característicos de la identidad de género. Y concluye:

Repito, una pequeña herida cortante en el sector superior de la mama y realizada por sobre la remera en un contexto de crimen por apuñalamiento con una víctima que pretendió defenderse, no puede ser tomado seriamente como uno de los dos indicadores —el otro es el de la violencia extrema— de un delito de «odio» (STO NÚM. 4,2018:341).

⁴² Para un resumen del caso véase el párrafo 110 del informe de CIDH. Zamudio fue atacado y torturado durante varias horas por un grupo de cuatro hombres. Le saltaron en la cabeza, orinaron encima, golpearon en los genitales, le hacen tres esvásticas con el cuello de una botella, le quiebran la pierna y se ríen al respecto del ruido que hizo, etc. También en la STO NÚM. 4,2018: 336-337.

En conclusión, la magistrada Bloch niega que tanto el nivel de violencia ni el lugar donde fueron realizadas las heridas pueden utilizarse para probar que el crimen fue cometido con odio. Estos fueron los indicadores utilizados por el tribunal y la acusación, sin embargo, la magistrada cita y analiza otros síntomas para después descartarlos. Entre ellos, pueden mencionarse los siguientes: dibujos o marcas dejados en la escena del crimen, coincidencias con un día significativo para la comunidad LGBTI, ejecución en la misma zona o a la misma hora donde otros hechos similares tuvieron lugar, previo involucramiento del acusado en un delito similar motivado por prejuicio, insultos o comentarios realizados por el imputado que hacen referencia a la orientación sexual, la naturaleza o significado del lugar donde se desarrolló la violencia o, el estatus de la víctima como activista de temas LGBTI (STO NÚM. 4,2018:345)⁴³.

Ahora bien, ¿cuál es el argumento de la magistrada Bloch para descartar que la pertenencia al grupo haya sido relevante en la comisión del delito? Ella afirma que si bien es cierto que hay que prestar atención a los casos donde la víctima es una activista, ello no significa que el conocimiento de este dato por parte de Marino implique la motivación de su delito. De nuevo, el tipo de casos al que se vinculan esos ataques no es el de un contexto similar al del evento donde murió Diana Sacayán. En síntesis, sostiene que no puede vincularse el hecho de que Diana haya sido una referente en la defensa de los derechos LGBTI y el homicidio.

Diana Sacayán trabajaba de modo análogos a estos referentes y también había sufrido violencia policial. Ese también podría haber sido su destino. Es cierto que encontró la muerte, y no niego la relación, entre esa muerte y un contexto hostil, pero ello no alcanza para inferir una responsabilidad personal del imputado más allá de lo que merece su aberrante acto (STO NÚM. 4,2018:354).

Creo que, a diferencia de lo que sostiene la magistrada, el desacuerdo con el tribunal no solo es por el modo en que estos tergiversaron o manipularon los indicadores, con cuya tesis estoy completamente de acuerdo, sino que el desacuerdo también es más profundo e interpretativo.

Creo que la magistrada también desacordaría, al igual que yo, acerca de la manera en que los magistrados del voto mayoritario interpretan la relación entre indicador y prueba de los hechos. Los magistrados del voto mayoritario interpretan los indicadores como si la mera existencia de algunos ellos fuera suficiente para atribuir el odio. Esto es, reduciendo la cuestión de la prueba del odio a una cuestión conceptual. Como señala GONZÁLEZ LAGIER (2005: 61): «cuando la unión entre los hechos probatorios y el hecho a probar viene dado por una teoría o una definición (esto es, cuando el vínculo es conceptual), no nos encontramos en sentido estricto ante un caso de prueba, sino de interpretación o calificación de los hechos».

Lo que quiero afirmar es que incluso si hubiera sido correcto afirmar que las heridas mostraban el grado de violencia típico de los homicidios con odio y hubiera sido

⁴³ . La lista es una combinación de algunos indicadores extraídos del informe contra la violencia de LGBTI de la CIDH y un informe del FBI. Para el análisis de los indicadores véase pp. 345-352.

verdad que las heridas se produjeran en zonas significativas de la sexualidad de Diana Sacayán ello todavía es compatible con que no hubiera habido un homicidio con odio, aunque sí se hubiera podido empezar a suponer que era plausible entender que hubo odio. Estos indicadores, si bien son muy útiles, el tribunal parece interpretarlos como rasgos definicionales de un homicidio con odio. Para el voto mayoritario las heridas con determinado de violencia y en zonas específicas del cuerpo *son* odio. Ha transformado una inferencia probatoria epistémica en una normativa⁴⁴.

Una interpretación alternativa de lo que ha hecho el tribunal, aunque con consecuencias semejantes consiste en haber transformado la cuestión en una regla de *prueba tasada*. Podría ser el caso que alguien distingue el odio de género de la violencia específica, pero que considere que, probada la violencia específica, debe tenerse por probado el odio. Esta regla conecta a las heridas extremadamente violentas, suponiendo que hubieran exhibido el grado de violencia efectivamente requerido, y las partes del cuerpo en que fueron hechas, con la prueba suficiente del odio. Sin embargo, esta regla no estaría empíricamente fundada. La presencia de esos indicios (las heridas y su ubicación en el cuerpo de la víctima) son explicables no solo por la hipótesis del odio, sino por explicaciones alternativas plausibles, como las que enumera la magistrada, que no fueron descartadas y que son menos gravosas para el imputado, violando el *in dubio pro reo* dado que imputa el hecho más gravoso sin descartar sin razones de peso las alternativas menos graves.

Por lo dicho anteriormente y porque existen otras hipótesis o explicaciones del evento que poseen al menos la misma plausibilidad que la de la motivación de odio es que la magistrada niega que se pueda aplicar la agravante. Sostiene, que una explicación alternativa plausible es la de un conflicto de carácter económico. Entre otros datos para sostener esta hipótesis tiene en cuenta los siguientes hechos (*i*) tanto víctima como victimario eran consumidores habituales de drogas y habían estado consumiendo cocaína durante la noche (*ii*) el homicidio no se produjo ni bien entró Marino en el departamento; (*iii*) poco tiempo antes de que se desencadenara el hecho, Marino llamó por teléfono a una persona de su confianza al que siempre acudía o bien cuando estaba sobrepasado o bien para pedirle dinero. Sería extraño que una persona que está por cometer un homicidio por odio llame por teléfono a otra, que no está involucrada en el delito, en el medio de la situación de odio (*iv*) el arma homicida estaba en el hogar y no la llevaba para cometer el delito; (*v*) la ira que Marino manifestaba cuando no conseguía lo que esperaba de las personas que le rodeaban, también surgió en ocasiones que Diana poseía una cantidad de dinero considerable en su casa y este no fue encontrado.⁴⁵ Así, concluye la magistrada Bloch, que la explicación del hecho

⁴⁴ Para los diferentes tipos de inferencias probatorias y los problemas que acarrea dicho modo de interpretar los elementos subjetivos, véanse: GONZÁLEZ LAGIER, 2005: 53-105; GONZÁLEZ LAGIER, 2015: 6-10 y 21-26. <http://hdl.handle.net/10045/46907>

⁴⁵ Para un análisis de los motivos alternativos, *e.g.* económico, véase STO NÚM. 4,2018:357-362.

...pudo ser la facilitación de un robo o el procurar su impunidad luego de concretarlo, pudo ser la alteración por un estado de intoxicación, pudo ser el odio de género o todas esas hipótesis juntas, pero ninguna de ellas derriba a la hipótesis rival (STO Núm. 4,2018: 362).

En resumen, podría entenderse el voto en disidencia de este sentido. La magistrada afirma que: (a) no es un caso donde el indicio de violencia específica exista; (b) aun cuando el indicio se da, el tribunal lo convierte en una cuestión o bien definitoria, o bien de prueba tasada. Y, tal como mostró en el voto, la hipótesis de odio de género es compatible con otras que se encuentran al mismo nivel de justificación, entonces no se puede sin violar el principio *in dubio pro reo* elegirse sin más la más gravosa. Esto podría afirmarse, tal como lo hago cuando reconstruyo el voto de la mayoría, incluso si asumimos que había el tipo de violencia extrema compatible con el odio al género. Sin embargo, la violencia extrema todavía puede ser compatible con otras hipótesis. Por ejemplo, que Marino odiaba a Sacayán por celos. En todo caso, suponiendo por hipótesis que la violencia extrema existió, el tribunal debería haber establecido características específicas adicionales de manera que sea un indicador (y solo un indicador) de la expresión de odio al género⁴⁶.

El voto de la magistrada Bloch muestra claramente que hay buenas razones para dudar acerca de la inevitable aplicación de la agravante por odio de género. Además, es rico en detalles dogmáticos y sugerencias de naturaleza político-criminal, pero no abundaré en esas consideraciones. Más bien, mi propósito es avanzar sobre algunas cuestiones de naturaleza conceptual, que son imprescindibles para comprender mejor el desafío de los delitos cometidos con odio hacia la identidad de género.

6. DELITOS DE ODIOS Y DELITOS ODIOSOS

Aunque ya he adelantado algunas ideas acerca de los delitos de odio al comentar la sentencia, creo que vale la pena mostrar de manera expresa cuál es mi posicionamiento en general y los compromisos teóricos que asumo en el análisis de este tema.

Con respecto a la posición que asumió el legislador al momento de establecer la agravante⁴⁷, aunque la extrema vulnerabilidad del colectivo puede haber sido una razón para incorporar la protección especial de la identidad de género en nuestro código penal, la norma penal (art. 80 inc.4) castiga con mayor gravedad a quien comete *el homicidio por odio*. De este modo, se hace hincapié en el estado mental o razón por la cual actúa el sujeto y no por si el colectivo al que pertenece la víctima

⁴⁶ Un desarrollo de las cuestiones teóricas vinculadas a la inferencia a la mejor explicación, véanse: HARMAN, 1965: 88-95; LIPTON, 1991; AMAYA: <http://ssrn.com/abstract=2044136>; LAUDAN: [HTTP://SSRN.COM/ABSTRACT=1153062](http://SSRN.COM/ABSTRACT=1153062), TUZET, 2013: Cap XII.

⁴⁷ Para una discusión sobre los distintos modelos legislativos acerca de los delitos de odio, véase, por ejemplo, DÍAZ LÓPEZ, 20013: 27.

tiene o no determinadas características.⁴⁸ Además, si asumiésemos por un momento que en la incorporación legal del inc. 4 hace hincapié en la protección de grupos vulnerables —y no en los motivos del agente— entonces se generarían dos inconvenientes importantes⁴⁹. En primer lugar, no hay espacio para que se hubiera cometido un homicidio contra una persona travesti y no fuera calificado. Y, en segundo lugar, se debe exigir a la víctima probar su pertenencia al grupo vulnerable (la homosexualidad, la minoría racial o religiosa, etc.). Así, por ejemplo, el fiscal en este caso debería esforzarse en probar la identidad travesti de Diana. Creo que ello, además de ser una interpretación equivocada de nuestro código penal, es una consecuencia desagradable e inasumible para un diseño institucional preocupado por resguardar las acciones privadas que no dañan a terceros (DÍAZ LÓPEZ, 2013: 113)⁵⁰.

Como mencioné al inicio de este trabajo, creo que el homicidio de Diana Sacayán fue, por el impacto que generó en la comunidad, un delito odioso antes que un delito cometido por odio. Para determinar si un crimen fue odioso es irrelevante que el sujeto activo haya experimentado odio al cometer su conducta o haya decidido realizarla por el odio que le tenía a determinada (clase de) persona. El delito puede ser identificado como odioso sin recurrir a las motivaciones del autor (DÍAZ LÓPEZ, 2013: 73). El criterio que sirve para identificar estos tipos de delito es el asco, la repugnancia, la indignación, el odio o el desprecio que la conducta del agente provoca en la comunidad o en nosotros. Así Frederick Reamer en su libro *Henious Crimes*, refiriéndose a esa clase de delitos, afirma: «Son esa clase de delitos tan horribles que noquea (shocks) nuestra conciencia colectiva» (REAMER, 2005: 4). El autor da un gran número de ejemplos de delitos odiosos que abarcan una amplia gama de tipo de delitos. Entre ellos, se puede mencionar la condena a Andrea Yates en 2001 que confesó asesinar a sus cinco hijos (de edad entre 6 meses y 7 años) en una bañera (REAMER, 2005: 1-3). Pero también podrían identificarse casos como de abusos de menores, estafas masivas a adultos mayores, etc. También podríamos pensar como un delito odioso situaciones donde a primera vista parecen imprudentes pero lo suficientemente desaprensivas como para ocasionarnos un desagrado visceral. Por ejemplo, el hallazgo de los 39 cadáveres dentro de un camión en Inglaterra donde se suponía que se debía a una mafia encargada de traficar con migrantes⁵¹. Por su-

⁴⁸ Por supuesto, esto no significa que en la dogmática penal se hayan realizado esfuerzos para mostrar que esta es una posición errada. Esta es la actividad que generalmente realiza la dogmática. 'Interpreta' aquello que dice el legislador de un modo que sea coincidente con las razones sustantivas que el teórico quiere que se sigan. Para ello véase MANRIQUE, NAVARRO y PERALTA, 2017.

⁴⁹ Sobre los pros y contra de ciertos modos de legislar, véase VÁZQUEZ, 2019: 193-219.

⁵⁰ Asumir un enfoque u otro también genera cuestiones dogmáticas importantes, en particular para el caso de tentativa y error. Por ejemplo, si una persona mata a otra porque cree que es homosexual y este no lo es, en el sistema que hace énfasis en la protección de clases vulnerables el homicidio no será calificado, incluso aunque el agente hubiera odiado a esa persona por lo que creía que era. Con este párrafo no quiero fundamentar definitivamente la opción que no tomó el legislador sino que pongo de manifiesto algunos problemas.

⁵¹ https://elpais.com/internacional/2019/10/23/actualidad/1571820020_804808.html

puesto, las comunidades van modificando su sensibilidad y lo que en un momento puede considerarse odioso en otro momento puede que no lo sea y viceversa. Así, las violaciones en masa en la época de la colonización en América Latina no generaban comparativamente un impacto como el que podría generar hoy en día.

La caracterización de un delito como odioso genera algunos interrogantes que no trataremos aquí. Por ejemplo, ¿qué ocasiona que ciertos delitos nos desagraden o repugnen tanto como para calificarlos como odiosos?, o ¿está justificado el desagrado ante ciertos hechos?, ¿está justificado aumentar una sanción o disminuir garantías constitucionales por la repulsión que el hecho genera en la comunidad?⁵², ¿La categoría de «odioso» se aplica a clase de delitos o a casos individuales que instancian ciertos delitos?

Un interrogante que vale la pena responder aquí es ¿cómo se identifican y diferencian las acciones realizadas con odio y aquellas que generan consecuencias odiosas?, ¿Un delito con odio genera siempre que este sea odioso? En el ámbito de la filosofía de la acción hay dos grandes maneras de identificar aquello que hace el agente. La primera, que podríamos llamar «psicológica», está centrada en los estados mentales del individuo que realiza la acción *i.e.* deseos, creencias, emociones, motivos, etc. Desde este punto de vista la explicación de una acción depende de los estados mentales que el agente *efectivamente* tuvo al momento de realizar la acción. La segunda manera, que podríamos llamar “semántica”, identifica las acciones no por los estados mentales que tenía el agente al momento de realizar una conducta sino por el *significado* que determinado comportamiento tiene en una comunidad. La acción de X de tatuarse una esvástica en el brazo *cuenta como* una manifestación del odio hacia determinado grupo de personas independientemente de que el agente haya experimentado o no esa emoción en particular (VON WRIGHT, 1983: 117).

Además, aquellos delitos que son cometidos por odio hacia una cierta clase normalmente nos parecen odiosos. Esta es una de las razones por las que consideramos que vale la pena identificar a los delitos por odio como una categoría especial dentro del resto de homicidios. Sin embargo, aunque «lo odioso» sea la razón por la que decidimos sancionar de modo especial los delitos de odio esto no significa que los dos criterios colapsen o que el criterio para identificar el delito cometido por odio sea que nos parezca odioso. Que un delito haya sido cometido por odio es una discusión acerca del caso individual, de los hechos del caso. Por el contrario, el calificar de odioso a ciertos tipos de delitos depende de la valoración general que se haga acerca de ciertos casos (clase de casos). A cierto tipo de casos le atribuimos la categoría de odioso (VON WRIGHT, 1983: 117).

⁵² En ciertas situaciones se considera que así como los delitos con odio implican un aumento de la pena los delitos odiosos provocan la eliminación de ciertas garantías constitucionales al identificar a quien comete este delito odioso como un «enemigo» de la sociedad. Sobre el derecho penal del enemigo, véanse por ejemplo: CANCIO MELIÁ Y GÓMEZ JARA, 2006; SILVA SÁNCHEZ, 2008: 73. Para una breve conexión de este discurso con los delitos odiosos, DÍAZ LÓPEZ, 2013: 68-77.

Ello significa que puede haber casos de delitos cometidos con odio que sean a su vez un crimen odioso pero no debe olvidarse que crímenes de odio y crímenes odiosos pueden surgir de manera separada. La muerte de los tres jóvenes afroamericanos en manos del KuKlux Klan en 1964 en Misisipi puede ser un ejemplo de delito cometido con odio que a su vez fue odioso.

Ahora bien, teniendo en mente esta categoría es que pueden entenderse algunas de las afirmaciones del magistrado Báez. Por ejemplo: «Estoy persuadido que el odio a la identidad de género denotado por el encartado podemos hallarlo en su *abominación* sobre el cuerpo transexual de Sacayan» (STO Núm. 4, 2018: 200). Por supuesto, el fin del argumento del magistrado es mostrar que Marino odiaba a Diana pero tanto esfuerzo en adjetivar durante todo su voto y el desentendimiento de los aspectos mentales del imputado podría explicarse por el desagrado que le genera el imputado y el hecho cometido por él. De todas maneras, aunque podría ponerse en duda si incluso si el hecho de Marino fue un delito odioso es importante destacar que nuestro código penal no castiga de mayor manera el homicidio por el impacto o la conmoción que este generó en la comunidad⁵³.

El uso del término «travesticidio» puede ayudar a pasar por alto esta diferencia, con independencia de otras consecuencias positivas que pueda acarrear el uso de este nuevo vocablo.⁵⁴ Así, la utilización de «travesticidio» no es, al menos, para la interpretación del agravante, útil ni positivo. El concepto hace hincapié en la protección de cierta clase de víctima vinculándose por analogía a los genocidios o el femicidio⁵⁵ donde el foco está puesto en quién es la víctima y no en qué pretendía el agente o qué motivación tenía al momento de realizar la conducta. Esto no sucede en el caso de Diana Sacayán. Podría responderse que el código penal simplemente protege a cierta clase de personas que han sido históricamente maltratadas. Algunos códigos penales

⁵³ En Brasil, por ejemplo, existe una enumeración de ciertos delitos como «crímenes hediondos» por ejemplo la explotación sexual infantil. Ser calificado como tal genera sanciones más graves y la limitación de ciertas condiciones en la ejecución de la pena (fianza, libertad condicional, ser amnistiado, etc.).

Creo que esta idea también podría utilizarse en el caso más reciente de los rugbiers en Villa Gesell. La indignación que generó en la comunidad el hecho de que 10 jóvenes jugadores de rugby, educados y adinerados mataran a golpes a la salida de una discoteca a un joven humilde y trabajador.

⁵⁴ Aquí no estoy haciendo una evaluación sobre los aspectos positivos de utilizar el término «travesticidio» en general. Tiendo a pensar que la utilización del vocablo puede mostrar de manera simple un problema profundo. Tampoco quiero afirmar que sería negativa la opción, como sugería un evaluador anónimo, de incorporar el término en el código penal al igual que sucedió con el femicidio. Mi breve comentario sobre el término está limitado solo a que en este caso, puede haber llevado a confundir dos maneras diferentes de justificar los delitos por odio. Para una breve descripción de los contextos en los que puede ser útil el término, véanse, por ejemplo: RADI Y SARDÁ-CHANDIRAMANI 2016; PIZZI Y SARALEGUI, 2018.

⁵⁵ No estoy afirmando que en este tipo de delitos no sean relevante los estados mentales de los agentes sino que más bien lo que nos impacta es o la cantidad de víctimas de una cierta clase (judíos, tutsis, haitianos, armenios, etc.) o ciertas personas históricamente desprotegidas (e.g. mujeres en el caso del femicidio).

pretenden proteger a un colectivo al agravar las penas por los delitos de odio. Esto es lo que sucede por ejemplo con el femicidio en el código penal argentino. Aunque se exigen elementos añadidos a la muerte de una mujer, sin duda, se pretende proteger a esta clase de personas⁵⁶. Sin embargo, esta no es la decisión que ha tomado el legislador argentino con respecto a los delitos de odio. El código penal argentino fundamenta la agravante por la emoción de odio que tenía el agente y que se manifiesta al momento de cometer el delito y cuya acción es una manifestación de esta. Es decir, que el agravante fundamenta su mayor reproche en ciertas características de la acción del agresor y no de la víctima.⁵⁷

Independientemente de cuál ha sido de hecho la opción del legislador, hay razones sustantivas que justifican un sistema donde el centro de la atención para atribuir responsabilidad penal y el grado de atribución de la misma radica en aquello que hace el agente. La manera de entender qué hizo el agente es preguntarnos qué quería hacer el individuo cuando actuó como lo hizo. Y, si este agente hizo lo que quiso. Preguntarnos sobre los estados mentales del individuo no solo es relevante sino también inescapable. No hay ninguna manera de *explicar* verdaderamente la acción de un individuo sin recurrir a ellos. Por supuesto, podemos preferir un sistema donde las consecuencias perjudiciales provocadas sean el objeto de evaluación, pero habremos abandonado el reproche por las acciones⁵⁸.

Otra estrategia menos radical sería aquella que afirma que la intención (el dolo) conformada por deseos y creencias o solo por creencias es importante para atribuir responsabilidad, pero no son igualmente relevantes otros estados mentales vinculados a la acción pero que no la determinan. En terminología de González Lagier, esta tesis sostendría que hay que ser descriptivista en cuanto a las intenciones, pero atribucionista en sentido fuerte en cuanto a los motivos⁵⁹. Podrían esgrimirse distintas razones (normativas, ideológicas) para sostener esta idea, pudiendo una de ellas ser epistémica: probar los motivos con los que realizamos una conducta, *e.g.* el odio, la envidia, la compasión, la empatía es más difícil que probar la intención *i.e.* el deseo y la creencia. Esta tesis tendría un problema central. Debería mostrar que hay algo específico en cuanto a la caracterización de los estados mentales concretos que permite distinguir entre la intención y un motivo. Es decir, deberían mostrar que el hecho de que algo cuente como motivo lo convierte en un estado mental de naturaleza diferente a la intención.

⁵⁶ Para un estudio de cuáles pueden ser los tipos de fundamentos de la incorporación del agravante en nuestros códigos penales, véase, por ejemplo: DIAZ LOPEZ, 2013; PERALTA, 2012.

⁵⁷ Por supuesto este enfoque no está exento de problemas teóricos. Aquí surgen las dudas sobre si reprochar los motivos es reprochar el carácter o los meros pensamientos de los individuos. Véanse, por ejemplo, MANRIQUE, 2019: 192-220; PERALTA, 2012.

⁵⁸ Para el vínculo entre la explicación de la acción y el reproche penal véase, por ejemplo, SANCI-NETTI, 2005.

⁵⁹ GONZÁLEZ LAGIER, 2003: 635-716.

Creo que a primera vista no hay nada que indique que probar la intención sea categorialmente más sencillo o menos complejo que probar los motivos. Este es un problema filosófico importante y con gran relevancia práctica. No estoy en condiciones de dar una respuesta aquí, sin embargo, creo que esto dependerá del caso en particular que estemos analizando. No hay que descartar de antemano que al Estado en algunos casos le puede resultar difícil probar los motivos porque está pretendiendo mostrar hechos que no existieron. Independientemente de que los magistrados del voto mayoritario no se esfuerzan lo suficiente en probar más allá de toda duda razonable el odio que en teoría poseía Marino hacia las personas que eligieron determinado género y a Diana Sacayán en concreto, tal vez (y solo tal vez) las dificultades de probar que Marino odiaba a Diana Sacayán surgen porque *en realidad* Marino no odiaba a Diana Sacayán por su identidad de género. En síntesis, no hay suficientes datos ni son suficientemente contundentes e independientes de la retórica para mostrar más allá de toda duda razonable que Marino odiaba a Diana por su identidad de género o su expresión⁶⁰.

Exigir la prueba del «motivo de odio» no solo es una garantía para el imputado de que se lo está responsabilizando por aquellas cosas que hizo, sino que también le da una importancia mayor a la víctima como individuo perteneciente a una comunidad y no meramente como una *clase de persona* que merece protección (judío, mujer, transexual, homosexual, etc.). Asumir que, porque la víctima pertenecía a una determinada clase de personas que históricamente ha sido maltratada, la muerte se ocasionó por ser como era invisibiliza la vida que como persona ha llevado y el esfuerzo que en ocasiones puede haberle hecho superar el estereotipo al cual ha sido sometida. Y esto creo que es lo que ha sucedido en el caso Diana Sacayán, es cierto que la comunidad trans y travesti ha sido vapuleada históricamente y que consta que la misma Diana ha sufrido maltrato, por ejemplo, por parte de la policía, sin embargo, ella superó gran parte de la marginalidad que se asocia a este colectivo, *i.e.* tenía un trabajo, no se dedicaba a la prostitución, tenía un lugar estable donde vivía y militaba y se codeaba con personas de importancia política. Asumir que porque ella pertenecía al colectivo trans/travesti su muerte fue provocada por ello hace que dejemos de ver todo lo que Diana Sacayán era como persona además de ser travesti⁶¹.

En conclusión, mientras más fina sea la explicación de por qué el agente actuó de determinada manera nos encontramos en mejores condiciones de imponer, en su caso, un castigo más adecuado a la conducta efectivamente realizada. Por supuesto, el derecho no le otorga relevancia a todos los motivos sino solo a los que considera importantes (para mal o para bien) y al tenerlos en cuenta nos hace identificar mejor

⁶⁰ Creo que el análisis que hace la magistrada del tipo y cantidad de las heridas en Diana es una muestra de ello.

⁶¹ Obviamente aquí no estoy disminuyendo la relevancia de la identidad de género como construcción de nuestros rasgos identitarios. Solo quiero decir que nuestra identidad se construye junto con otras características *además* de la identidad de género.

cómo castigar⁶². Los motivos como el odio hacia alguna raza o algún género han sido tomados como agravantes de ciertas conductas, en este sentido hacen pensar que ellos solo reflejan a un autor más culpable que realizó la misma conducta que otro, *e.g.* homicidio, que no poseía ese motivo pero ello no tiene que ser necesariamente el caso. En ocasiones los motivos definen aquello que el agente hace *e.g.* tortura, defensa propia, estado de necesidad, etc., y no hay necesidad en entender que ellos cumplen una sola función en la descripción y la explicación de la acción⁶³.

7. BALANCE

Los temas centrales de este trabajo pueden dividirse en dos grandes partes. En primer lugar, puede destacarse la crítica a los fundamentos de la decisión de la mayoría en el caso Sacayán. Es ampliamente aceptado que para distinguir entre la imposición de una pena y el maltrato estatal, es fundamental —entre otras cosas— que sea posible controlar racionalmente la prueba incorporada y la valoración de los hechos que realiza el tribunal. En este sentido, la exigencia es que la decisión esté basada en una descripción verdadera de los hechos y que esa descripción conecte el comportamiento del agente responsable, el resultado de sus acciones y las consecuencias del evento (ilícito). Sin embargo, como he destacado en la primera parte del trabajo, los esfuerzos del voto mayoritario no alcanzan a mostrar más allá de toda duda razonable cómo y por qué sucedieron los hechos. Más bien, el uso de lenguaje emotivo y las simplificaciones sobre el curso de los acontecimientos procuran *persuadir* del punto de vista del tribunal y no constituyen una razón suficiente para calificar el homicidio de Diana Sacayán como un caso de odio de género⁶⁴.

En segundo lugar, aunque la sentencia carezca de justificación podría construirse una explicación tentativa de la convicción a la que arriba la mayoría del tribunal.

⁶² La discusión sobre si el Estado eligió bien los motivos que considera relevantes es una discusión política que no daré aquí. Sin embargo, creo que esta es una discusión importante que debe darse en el ámbito del reproche de los motivos de odio. Para un estudio de las distintas maneras en que el odio es relevante en derecho penal, véase FUENTES OSORIO, 2017.

⁶³ En el mismo sentido MATHIS, 2018:10. En términos dogmáticos estoy afirmando que los motivos no muestran solo cuán culpable es el agente, sino que en ocasiones los motivos impactan en el injusto (el hecho). Ellos definen qué hizo el agente. El hecho que la teoría del delito no encuentra una categoría clara donde colocar los motivos no es un problema para evaluar los motivos, sino que es algo vinculado a las categorías que hemos establecido para analizar el delito.

⁶⁴ Obviamente no quiero negar que el maltrato sistemático a travestis y trans sea un problema que deba formar parte del discurso de toda la comunidad y ello también afecta a los jueces. Sin embargo, aunque los jueces también pueden tomar en consideración estas cuestiones lo deben hacer solo a través de decisiones (sentencias) correctas. En este trabajo he intentado criticar ese aspecto de la decisión y no la pretensión de proteger a un colectivo altamente vulnerable. No debe olvidarse que a través de esta decisión se impacta sobre la libertad de otro individuo, incluso aunque el hecho cometido por este individuo nos repugne. Para entender el modo en que travestis y trans han sido históricamente maltratados, véase, por ejemplo: MAFFIA, 2016. <https://www.youtube.com/watch?v=b936aFQpWvg>

Esta explicación giraría en torno del profundo desagrado o repugnancia que produce el homicidio de Diana Sacayán. Tal vez porque el crimen parece odioso también es tentador asumir que el odio *es* el único motivo para ejecutarlo. Sin embargo, como mostré más arriba, este es un argumento falaz. Los delitos cometidos por odio son, al menos parcialmente, independientes de los delitos odiosos. Mientras que la prueba de los primeros exige indagar acerca del marco epistémico-volitivo que determinó la conducta al agente, los segundos dependen del marco social de referencia en el que se producen. De este modo, la prueba de los delitos cometidos por odio requiere de un análisis de las circunstancias de hecho que hacen posible un cierto delito, la prueba de los segundos exige una compenetración con las convenciones sociales vigentes en una cierta comunidad. Por ello, aunque un hecho nos cause profundo desagrado e indignación, ello todavía no es una razón, al menos en el sistema jurídico argentino, para sostener que el hecho del agente debe considerarse agravado en los términos del artículo 80 inc.4 del código penal argentino. Tal vez, puede ser teóricamente interesante discutir si todo delito odioso es, por sí mismo, un delito cometido por odio, pero este debate sería, en el mejor de los casos, una consideración adicional (*obiter dicta*) y no aporta nada decisivo al momento de fundamentar las sentencias⁶⁵.

8. BIBLIOGRAFÍA

- ALCHOURRÓN, C. y BULYGIN E., 1991: *Análisis lógico y derecho*, Madrid: Centro de Estudios Constitucionales.
- ÁLVAREZ, J.T., 2019: «Homicidios vinculado a la identidad de género o su expresión. Comentario acerca de la sentencia por el travesticidio de Amacay D. Sacayán», *Temas de derecho penal y procesal penal*, Buenos Aires: Erreius, 201365-383.
- ARENA, F., 2016: «Los estereotipos normativos en la decisión judicial. Una exploración conceptual», *Revista de Derecho de la Universidad Austral de Chile*, (Vol. XXIX, N.1): 51-75.
- 2018 «Estadísticas, estereotipos y grupos desfavorecidos. Algunos límites del apoyo estadístico de los estereotipos», *Anuario de la Facultad de derecho*, UNC: 1-18.
- 2019 «Algunos criterios metodológicos para evaluar la relevancia jurídica de los estereotipos», *Derecho y Control* (2), Córdoba: Ferreyra Editor: 11-44.
- AROCENA, G., 2014, «La incorporación de los discursos de género en la tipificación del homicidio en el derecho argentino», *Diritto Penale Contemporaneo*. <https://archiviopdc.dirittopenaleuomo.org/d/3311-la-incorporacin-de-los-discursos-de-genero-en-la-tipificacin-legal-del-homicidio-en-el-derecho-arge>
- BAYÓN MOHINO, J.C ,2009: «Epistemología, moral y prueba de los hechos: hacia un enfoque no benthamiano», *Analisi e diritto. Ricerche di giurisprudenza analitica*, 2009-2010, 1.: 15-34.
- BOUZAT, A y CANTARO, A., 2003: «Verdad y prueba en el proceso acusatorio», *Discusiones*, 3: 67-79.

⁶⁵ Ya Lord Devlin intentaba mostrar que el impacto en la comunidad es un factor para determinar que una conducta es incorrecta. Hart fue un gran crítico de esta idea. Para un análisis crítico sobre la manera en que emociones como la repugnancia no deben utilizarse como criterio de corrección, véase NUSSBAUM, 2006. Tiendo a pensar, como Hart y como Nussbaum que fundamentar el reproche en este impacto es completamente dañino. De todas maneras, es un tema que merece mayor análisis.

- BOUVIER, H. y PÉREZ BARBERÁ, G., 2004: «Casación, lógica y valoración de la prueba: Un análisis de la argumetación sobre los hechos en las sentencias de los tribunales casatorios», *Pensamiento Penal y Criminológico: revista de Derecho Penal integrado*, 9: 171-195.
- CANCIO MELIÁ y GÓMEZ JARA (coord.), 2006: *Derecho penal del enemigo. El discurso penal de la exclusión*, España: Edisofer.
- CARACCILO, R., 2000: «El problema de los hechos en la justificación de la sentencia», III Seminario Internacional de Teoría del Derecho, Universidad Nacional del Sur, septiembre 2000.
- COHEN, J. L., 1989: *An Introduction to the Philosophy of Induction and Probability*, Oxford: Clarendon Press.
- COHEN, L. J., 1977: *The probable and the provable*, Oxford: Clarendon Press.
- CARRIÓ, G., 1983: *Notas sobre derecho y lenguaje*, Buenos Aires: Abeledo Perrot.
- COPI, I. y COHEN C., 2013: *Introducción a la lógica*, México: Limusa.
- DÍAZ LÓPEZ, J., 2013: *El odio discriminatorio como agravante penal*, Pamplona: Civitas.
- FERRER BELTRÁN, J., 2001: «Está probado que p'», en Triolo, L., *Prassi giuridica e controllo di razionalità*, Torino: G. Giappichelli: 73-96.
- 2006: «Legal proof and fact finders' beliefs», *Legal Theory*, 12, 4: 293-314.
- 2007: *La valoración racional de la prueba*, Madrid: Marcial Pons.
- 2010: «Una concepción minimalista y garantista de la presunción de inocencia».
- 2011: «Apuntes sobre el concepto de motivación de las decisiones judiciales», *Isonomía. Revista de Teoría y Filosofía Del Derecho*, 34: 87-108.
- 2011: «La valoración de la prueba: verdad de los enunciados probatorios y justificación de la decisión», en Ferrer Beltrán, J., Gascón, M., González Lagier, D. y Taruffo, M., *Estudios sobre la prueba*, México: Fontamara, pp. 15-59.
- 2013: «La prueba es libertad, pero no tanto: una teoría de la prueba cuasibenthamiana», en Vázquez, C., *Estándares de prueba y prueba científica*, Madrid - Barcelona - Buenos Aires - Sao Paulo: Marcial Pons, pp. 21-39.
- 2018: «Prolegómenos para una teoría sobre los estándares de prueba. El test case de la responsabilidad del Estado por prisión preventiva errónea», en Papayannis, D. M. & Pereira Fredes, E., *Filosofía del derecho privado (en prensa)*, Madrid-Barcelona-Sao Paolo-Buenos Aires: Marcial Pons, pp. 401-430.
- FUENTES OSORIO, J.L., 2017, «El odio como delictio», *Revista Electrónica de Ciencias Penales y Criminología*, vol 19-27.
- GASCÓN ABELLÁN, M., 2004: *Los hechos en el derecho: Bases argumentales de la prueba*, Madrid- Barcelona: Marcial Pons.
- GARCÍA CARPINTERO, M., 1996: *Las palabras, las ideas y las cosas*, Madrid, Ariel.
- GIMBERNAT ORDEIG, E., 1999: *Concepto y método de la ciencia del derecho penal*, Madrid: Tecnos.
- GONZÁLEZ LAGIER, D., 2003: «Buenas razones, malas intenciones: (sobre la atribución de intenciones)», *Doxa*, (26): 635-716.
- 2005: *Quaestio facti. Ensayos sobre prueba, causalidad y acción*, México D.F.: Fontamara.
- 2015: «Presunción de inocencia, verdad y objetividad» <http://hdl.handle.net/10045/46907>
- HEMPEL, C., 1996: *La explicación científica*, Barcelona-Buenos Aires: Paidós.
- KENNY, A., 2004: *Actions, Emotions and Will*, London and New York: Routledge.
- LAUDAN, L. 2003: «Is reasonable doubt reasonable?», *Legal Theory*, 9, 4: 295-331.
- 2005: «Por qué un estándar de prueba subjetivo y ambiguo no es un estándar», *DOXA: Cuadernos de Filosofía del Derecho*, 28: 95-113.
- 2011: «The Rules of Trial, Political Morality, and the Costs of Error: Or, Is proof Beyond a Reasonable Doubt Doing More Harm than Good?», en Green, L. Leiter, B., *Oxford studies in philosophy of law*, Oxford: Oxford University Press, pp. 195-227.
- 2013a: «La elemental aritmética epistémica del derecho II: los inapropiados recursos de la teoría moral para abordar el derecho penal», en Vázquez, C., *Estándares de prueba y prueba científica*, Madrid - Barcelona - Buenos Aires - Sao Paulo: Marcial Pons, pp. 119-134.
- 2013b: *Verdad, error y proceso penal: un ensayo sobre epistemología jurídica*, Madrid: Marcial Pons.

- 2016: *The Law's Flaws. Rethinking Trials and Errors?*, Milton Keynes, Lightning Source.
- LAUDAN, L. Y SAUNDERS, H. 2009: «Re-Thinking the Criminal Standard of Proof: Seeking Consensus About the Utilities of Trial Outcomes», Disponible en SSRN: <https://ssrn.com/abstract=1369996> o <http://dx.doi.org/10.2139/ssrn.1369996>.
- MANRIQUE, M.L., NAVARRO P.E., y PERALTA JOSÉ M., 2017: «Criminal Law and Legal dogmatics», *Revus*, (31).
- MANRIQUE, M.L., 2019: «Delitos de odio y motivos emocionales», *Análisis Filosófico*, Buenos Aires, SADAFA, (Vol XXXIX, Núm. 2): 192-220.
- MAFFÍA, D., 2016: «Jornada de capacitación sobre “Crímenes de género: del femicidio al travesticidio / transfemicidio”» <https://www.youtube.com/watch?v=b936aFQpWvg>
- MATHIS, S., 2018: «Motive, Action, and Confusions in the Debate over Hate Crime Legislation», *Criminal Justice Ethics*, (Vol 37,1): 1-20.
- NUSSBAUM, M., 2006: *El ocultamiento de lo humano —Repugnancia, vergüenza y ley—*, Buenos Aires: Katz.
- PERALTA, J.M., 2013: «Homicidios por odio como delitos de sometimiento», *Indret*, (4).
- PIZZI, L. Y SARALEGUI, N., 2018: «El continuum de violencias contra el colectivo travesti y trans a la luz del fallo de Diana Sacayán», *Estudios sobre jurisprudencia*, Ministerio Público de la Defensa <https://jurisprudencia.mpd.gov.ar/Estudios/Forms/DispForm.aspx?ID=32&source=/Estudios/forms/AllItems.aspx>
- RADI, B. y SARDÁ-CHANDIRAMANI A., 2016: «“Travesticidio / transfemicidio” Coordinadas para pensar los crímenes de travestis y mujeres trans en Argentina» Boletín Núm. 9, *Observatorio de Género en la Justicia*, Consejo de la Magistratura de la Ciudad de Buenos Aires.
- REAMER, F., 2005: *Heinous Crime: cases, causes and consequences*, Columbia University Press.
- ROSS, A., 1963: *Sobre el derecho y la Justicia*, Buenos Aires: Eudeba.
- SÁNCHEZ BRÍGIDO, R. Y SELEME, H. 2002: «Justificación sin verdad», *Analisi e diritto*, 2001: 285-294
- SANCINETTI, M., 2005: *Teoría del delito y disvalor de acción*, Buenos Aires: Hamurabi.
- SHAFER, G. 1988: «The Construction of Probability Arguments», en Tillers, P. y Green, E. D., *Probability and inference in the law of evidence : the uses and limits of bayesianism*, Dordrecht [etc.]: Kluwer Academic Publishers, pp. 185-204.
- SILVA SÁNCHEZ, J., 2008: *La expansión del derecho penal. Aspectos de la política criminal en las sociedades post industriales*, Madrid: Cívitas.
- STEVENSON, C., 1944: *Ethics and Language*, USA: Yale University Press.
- TARUFFO, M. 1992: *La prova dei fatti giuridici. Nozioni generali*, Milano: Giuffrè.
- 2002: *Sui confini. Scritti sulla giustizia civile*, Bologna: Il Mulino.
- 2003: «Algunas consideraciones sobre la relación entre prueba y verdad», *Discusiones*, 3: 15-41.
- TILLERS, P Y GREEN, E., 1988: *Probability and Inference in the Law of Evidence: the Uses and Limits of Bayesianism*, Dordrecht [etc.]: Kluwer Academic Publishers.
- TUZET, G., 2013: *Filosofía della prova giuridica*, Torino: Giappichelli.
- VÁZQUEZ C., 2019: «Técnicas legislativas del femicidio y sus problemas probatorios», *DOXA*, (42): 193-219.
- VON WRIGHT, G.H., 1984: *Practical Reason: Philosophical Papers I*, Oxford: Blackwell.

CÓMO PUBLICAR EN LA REVISTA

1. REMISIÓN DE ARTÍCULOS A EFECTOS DE SER EVALUADOS

Todo artículo debe ser enviado a *revistaquaestiofacti@gmail.com* en formato .doc o .docx.

Revisión: una vez recibidos, todos los textos se someterán a doble revisión ciega por pares.

Idiomas: se aceptan textos escritos en español, inglés, italiano y portugués.

Extensión: se sugiere que los textos no superen las 15.000 palabras. Las *refutaciones* en la sección *Conjeturas y refutaciones* no podrán superar las 5.000 palabras.

2. REGLAS DE EDICIÓN

2.1. Presentación del trabajo

Nombre y apellido del autor, afiliación académica y correo electrónico de contacto.

Título del artículo. Si posee subtítulo debe ir en una línea independiente respecto del título.

Título y subtítulo en inglés.

Abstract de no más de 150 palabras en el idioma en que está redactado el trabajo.

Abstract en inglés, si no fuera este el idioma en que el artículo está escrito.

5 palabras clave en el idioma de redacción del texto.

5 palabras clave en inglés, si ese no fuera el idioma de redacción.

Versión para revisión ciega. Todos los textos deben acompañarse de una versión destinada a ser sometida a revisión sin dato alguno del autor, eliminados los agradecimientos y cualquier otra referencia que pueda indicar su identidad.

2.2. Presentación del contenido

Fuente y tamaño. Los textos deben ser enviados en letra Times New Roman, cuerpo 12, interlineado 1,5. Sangría al inicio de cada párrafo.

Notas. Las notas deben estar editadas en cuerpo 10 e interlineado sencillo, sin sangría. Numeración correlativa iniciando desde 1 en cada artículo.

Subtítulos. El orden de los encabezamientos será el siguiente: 1, 1.1, 1.1.1, 1.1.2, 2, etc.

Citas

Las citas bibliográficas seguirán el modelo anglosajón. Los apellidos en VERSALITAS, el año de edición del original (cuando el autor tenga varias obras publicadas en un mismo año se pondrán letras al lado de la fecha, comenzando por a, b, c...). Como ejemplo:

(YOUNG, 1971a: 25-26)

Si hay varios autores citados del mismo apellido, añadir inicial del nombre, después apellido:

COLOMER, J., 1995: ...

Si se trata de una cita de un artículo de revista del que no consta el año, la manera de citar sería, por ejemplo:

COLOMER, J., «The Spanish ...»: 40-52

Las citas en texto principal o cuerpo de notas irán en redonda (*i. e.* no cursiva) y entre comillas «angulares».

Si la cita en el cuerpo posee una extensión de más de tres renglones no irá entre comillas y deberá aparecer separada del texto principal, en cuerpo 10 y centrada.

Comillas. Las comillas dentro de comillas pasarán a ser comillas “voladitas”.

Bibliografía. La bibliografía irá al final, en orden alfabético, respetando el siguiente formato:

Libros: APELLIDO EN VERSALITAS, inicial del nombre, año: *título del libro en cursivas*. Lugar: Editorial.

Artículos de revista: APELLIDO EN VERSALITAS, inicial del nombre, año: «título del artículo entre comillas angulares», en *nombre de la revista en cursiva*, número (si tiene número de tomo o volumen, entre paréntesis): páginas.

Capítulos de libros: APELLIDO EN VERSALITAS, inicial del nombre, año: «título del capítulo entre comillas angulares», en *título del libro en cursiva*. Lugar: Editorial, páginas.

Por ejemplo:

PUTNAM, R., 1995: *Making democracy work. Civic traditions in modern Italy*. Princeton: Princeton University Press.

COLOMER, J., 1998: «The Spanish “State of Autonomies”: Non-Institutional federalisms», en *West European Politics*, 21 (4): 40-52.

Seguir criterios en cuanto a: derecho (siempre en caja baja); Estado; ordenamiento jurídico; juez; tribunal; fiscal; Tribunal Supremo (organismos en general, en alta); capítulo.

3. COMPROMISO ÉTICO Y DERECHOS DE AUTOR

3.1. Compromiso ético

Quaestio Facti. Revista Internacional sobre Razonamiento Probatorio adhiere a las directrices fijadas por el EASE (European Association of Science Editors) a efectos de la gestión, edición, revisión y publicación de resultados científicos en revistas de diferentes áreas de conocimiento. De conformidad con esas directrices:

Los autores se comprometen a remitir trabajos inéditos, que no hayan sido publicados anteriormente y que no se encuentren sometidos a evaluación por parte de otras revistas, al menos hasta tanto no se complete el proceso de evaluación por parte de esa revista. El envío de trabajos para su evaluación requiere omitir en la copia ciega cualquier dato que pueda ser identificado por los revisores respecto de su autoría y de igual modo implica aceptar las normas de publicación, revisión y evaluación de la revista que se detallan a continuación, así como en cada número de la misma.

Los revisores de los trabajos asumen el compromiso de realizar una revisión crítica, honesta y constructiva sobre la calidad científica del texto dentro el ámbito de sus conocimientos y habilidades. Por ello, solo revisarán un trabajo en caso de ser competentes en la temática a revisar y si no existen conflictos de interés. *Quaestio facti* informará a autores y autoras en tiempo y forma adecuados acerca de la aceptación o rechazo de sus trabajos, asegurando una edición respetuosa respecto del estilo y de las ideas de cada autor, dando un tiempo adecuado a efectos de la corrección de la edición final.

Los Directores, miembros del Consejo de Redacción, miembros del Consejo Editorial y los miembros del Consejo Asesor se comprometen a respetar la imparcialidad y mantener la confidencialidad de los trabajos enviados, sus autores y revisores, de

forma que el anonimato preserve la integridad de todo el proceso de evaluación. A tal fin garantizarán la selección de los revisores más cualificados y especialistas en la materia para emitir una apreciación crítica y experta del trabajo. Asimismo, en la medida de lo posible, evitarán todo tipo de conflictos de intereses y cumplirán estrictamente con los tiempos de evaluación, edición y publicación que exige la periodicidad de la revista.

3.2. Derechos de autor

La remisión de artículos a efectos de ser publicados por *Quaestio facti. Revista Internacional sobre Razonamiento probatorio* implica la aceptación de los siguientes términos:


- El autor o autora se compromete a respetar las pautas éticas antedichas.
- Se concede a *Quaestio facti* el derecho de primera publicación y otorgamiento de licencia para publicar el texto remitido en todos los formatos y medios impresos y electrónicos de la revista.
- A efectos de toda reimpresión, abreviación o traducción de un artículo de cualquiera de sus secciones, *Quaestio facti* se compromete a solicitar autorización sin demora a cada autor y autora del trabajo en cuestión, a efectos de que él o ella decidan si autorizan dicha publicación. Al mismo tiempo, los autores que decidan republicar su trabajo en otros sitios se comprometen a dejar constancia en la nueva publicación de la previa aparición de la contribución en *Quaestio facti*.

Qf

 Marcial
Pons

 Càtedra
de Cultura
Jurídica

Universitat
de Girona

 Diputació de Girona

www.quaestiofacti.com