

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 IMPRECIACIÓN 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.

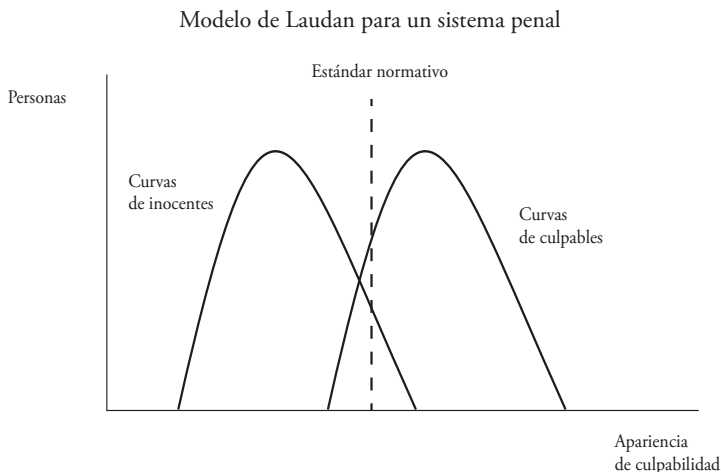
3. EL ENFOQUE DE LARRY LAUDAN SOBRE LA FUNCIÓN DEL ESTÁNDAR DE PRUEBA

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.

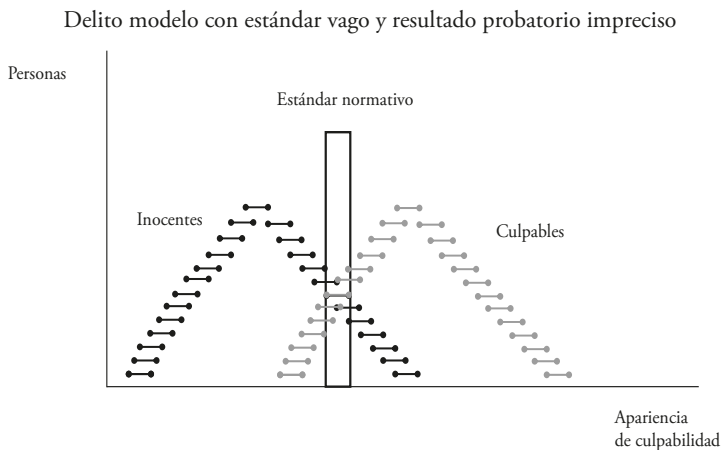
Así, el límite que separa la revocación de una decisión y su reemplazo por una de signo contrario sin otras consecuencias, de la que implica una revocación por aplicación de la doctrina de la arbitrariedad y que en algunos casos podría llegar incluso al inicio de un procedimiento disciplinario o penal 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.

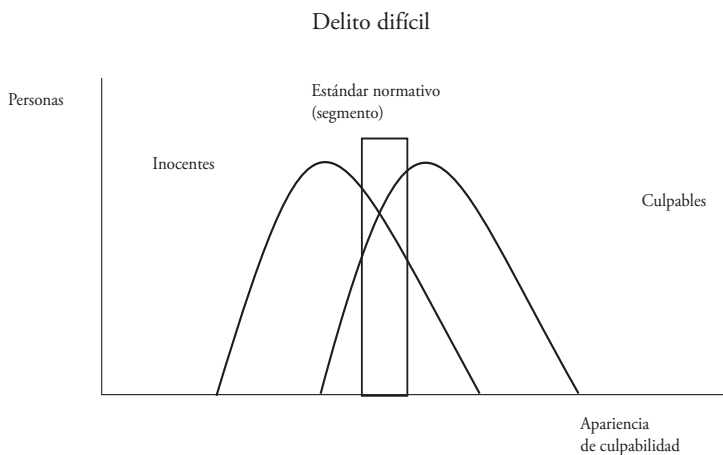
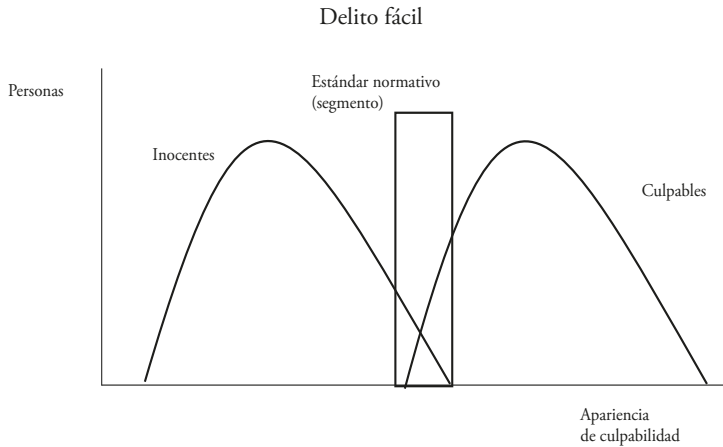
Como es fácil de advertir, para cada delito los conjuntos probatorios de referencia y los contextos de desdicha pueden ser muy diferentes.

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.

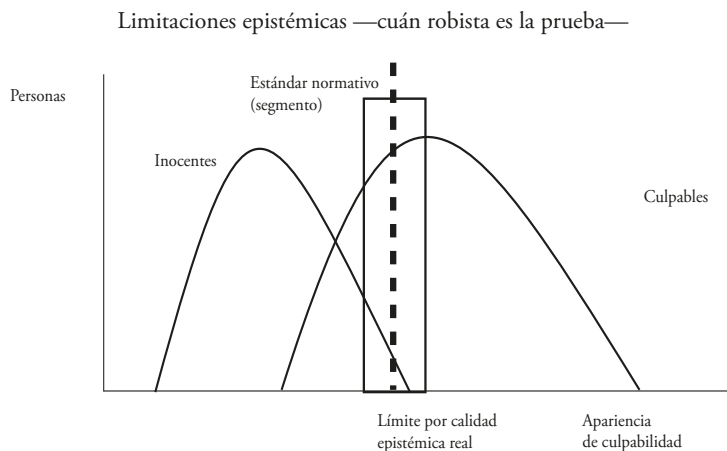
Corresponde observar entonces desde esta perspectiva práctica, cómo impacta la mala calidad en función de sus dos dimensiones («fuerza de la prueba» e «inferencias válidas»).

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.

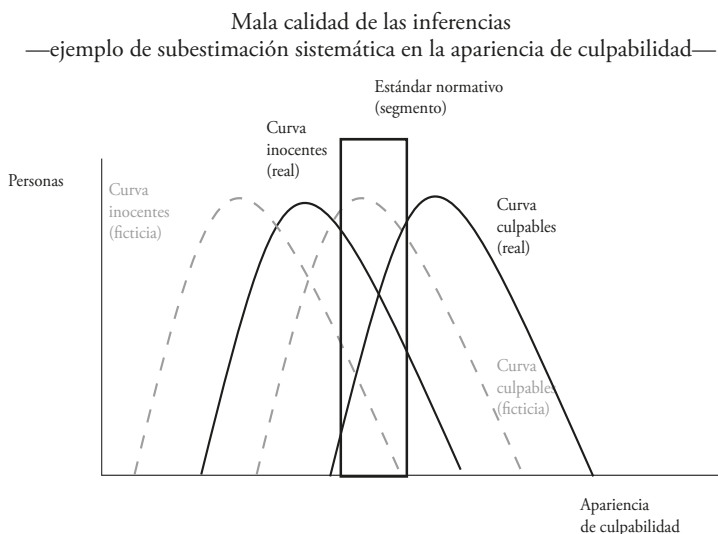
Esa clase de generalizaciones (no científicas, ni tan arraigadas en la experiencia general), son muy abundantes dada la limitada evidencia que suele haber en las causas y juegan un papel crucial en la suerte de los casos. Es en la formulación y discusión de estos fundamentos para nuestras inferencias 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).

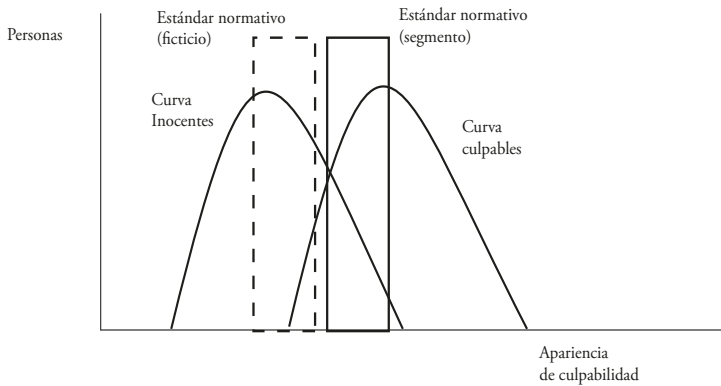
Cabe concluir entonces que la mala calidad de las inferencias en el sistema, en cuanto a su potencial para incidir sobre la distribución de errores y su número total, es muy relevante. Ello se advierte con claridad, por ejemplo, en todo el tramo de la curva de culpables ficticia (verde) que se encuentra 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.

