

PRUEBA CIENTÍFICA: MITOS Y PARADIGMAS *

Scientific Evidence: Myths and Paradigms

Marina GASCÓN ABELLÁN

Universidad de Castilla-La Mancha (España)

RESUMEN

Este trabajo se propone desactivar algunos de los mitos y paradigmas que informan en la actualidad la valoración de la prueba científica. Para llevarlo a cabo, la autora parte de la premisa de que tales mitos y paradigmas, conformados con el desarrollo de la ciencia forense, están contribuyendo a minimizar el papel de los jueces en los procesos judiciales, en beneficio de los peritos. Estas situaciones pueden llevar a la violación de garantías procesales y al olvido de que los procesos judiciales se dirigen a resolver conflictos. De manera concreta, la autora se refiere a los mitos de la infalibilidad y al conocido como el *paradigma de la individualización*, según el cual las pruebas científicas permiten identificar plenamente a un individuo u objeto a través de vestigios. Mediante el análisis de la reciente doctrina sobre la valoración de la prueba científica, la autora concluye en la necesidad de adoptar el *paradigma de la verosimilitud* y de aplicar el teorema de Bayes, que combina información estadística con información no estadística, ya que así se garantiza mejor el protagonismo del juez en la valoración de la prueba, y el respeto a los fines y garantías del proceso.

Palabras clave: prueba científica, validez de la prueba científica, valoración de la prueba, garantías procesales.

ABSTRACT

The basic purpose of this work consists of eliminating some of the myths and paradigms that currently inform the assessment of scientific evidence. In order to accomplish this objective, the author starts from the premise that these myths and paradigms, produced with the development of forensic science, are minimizing the role of judges in judicial proceedings for the benefit of experts. This can lead to a violation of due process and neglecting that judicial processes are aimed at resolving conflicts. The author refers specifically to the myth of infallibility and the *paradigm of identification*, according to which scientific evidence can positively identify an individual or object through physical traces. By analyzing the recent doctrine on the assessment of scientific evidence, the author concludes on the need to adopt the *paradigm of likelihood*. Thus, the judge still holds the role in the evaluation of evidence, and guarantees that the goals of the process will be respected.

Key words: scientific evidence, validity of evidence, evaluation of evidence, guarantees of legal process.

* Este trabajo se ha realizado en el marco del proyecto PIIIII09-0173-2296, financiado por la Junta de Comunidades de Castilla-La Mancha.

SUMARIO

1. Importancia y caracterización de la prueba científica: el mito. 2. Los paradigmas. 2.1. Sobrevaloración semántica de la prueba. 2.2. Sobrevaloración epistémica. 3. Las consecuencias del paradigma de la verosimilitud: los retos. 3.1. Sobre la admisibilidad de la prueba. 3.2. Sobre la formulación de las conclusiones. 3.3. Sobre la valoración de la prueba.

1. IMPORTANCIA Y CARACTERIZACIÓN DE LA PRUEBA CIENTÍFICA: EL MITO

En los últimos años los constantes avances científicos y técnicos han tenido un profundo impacto en el ámbito de la prueba. La dactiloscopia, la balística, la grafística, etc., son ejemplos de esta proyección de los conocimientos y avances científicos en el campo judicial, donde suelen ser adquiridos a través de informes periciales. Los avances han sido particularmente espectaculares en el campo de la biología molecular. En concreto, lo que se denomina Genética Forense, consistente en el análisis genético de la diversidad humana, ha marcado un antes y un después en la resolución de ciertos problemas judiciales. El rendimiento de la prueba de ADN (o huella genética) radica en que los miles de pares de bases que se reparten de forma secuencial y determinada para cada persona permiten seleccionar a un único individuo entre todos los de su especie si se conoce esa secuencia. El potencial de la huella genética es de tal magnitud que su uso en los tribunales se ha convertido ya en moneda corriente¹. En los últimos años, además, se han ido desarrollando nuevas tecnologías de identificación por ADN que han permitido extender la prueba a muestras que hasta ahora no podían ser objeto de la misma. Y, por otra parte, las intervenciones precisas para obtener muestras para la práctica de un análisis de ADN son cada vez menos invasivas, lo que desde el punto de vista jurídico facilita la práctica de la prueba. Por otra parte, no es sólo la prueba de ADN la protagonista de este boom de la prueba científica. La dactiloscopia y la balística también juegan un papel protagonista en muchos procesos. Y la prueba de locutores de voz, por ejemplo, ha permitido fundar condenas en causas de terrorismo en las que la prueba principal frente al imputado es una llamada de teléfono avisando de la colocación de un artefacto explosivo.

Pero la importancia de las pruebas científicas en la práctica procesal no ha ido acompañada de un proceso paralelo de cautelas y controles en relación con las

1. Son muchas las posibles aplicaciones forenses de la prueba, pero los tipos de pericias más comunes son la investigación biológica de la paternidad, la resolución de problemas de identificación y la investigación de indicios en criminalística biológica, es decir, el análisis de muestras biológicas de interés criminal, como manchas de sangre, saliva, esperma o pelos. Sobre éstas y otras aplicaciones de la prueba de polimorfismos ADN a la Genética Forense, vid. *La prueba del ADN en Medicina Forense. La Genética al servicio de la ley en el análisis de indicios criminales y en la investigación biológica de la paternidad* (Begoña Martínez Jarreta dir.), Barcelona, Masson, 1999. Especialmente, pp. 129 y ss.

mismas. Más bien ha sucedido lo contrario. Sólo por el hecho de presentarse como “científicas”, la validez y valor probatorio de este tipo de pruebas se han asumido como dogmas de fe. Algo que llama particularmente la atención si se considera que nada es menos “científico” que asumir como válido un conocimiento sin un previo control de sus postulados ajustado a una metodología científica.

En efecto, todos los esfuerzos que en los últimos tiempos se están realizando por introducir racionalidad en el ámbito de la prueba, y especialmente en la valoración, que constituye su núcleo esencial, se dirigen fundamental y casi exclusivamente a las pruebas no científicas. Ello es así porque se parte de que, a diferencia de la prueba no científica, estructurada a través de un razonamiento *inductivo*, la prueba científica se articula mediante un razonamiento *deductivo*. Lo que se sostiene, en otras palabras, es que la prueba *no científica* está basada en leyes probabilísticas de débil fundamento epistémico, por lo común máximas de experiencia y leyes del actuar humano habitual, que además son aplicadas dentro de una metodología no científica, de modo que sus resultados han de medirse siempre en términos de simple probabilidad. Son, en definitiva, falibles. En cambio, la prueba *científica* —se argumenta— por estar basada en leyes universales (o en todo caso en leyes probabilísticas que gozan de un fuerte fundamento epistémico) que además son aplicadas dentro de una metodología científica, aparece rodeada de un aura de infalibilidad que proporciona a la decisión probatoria un carácter concluyente o casi concluyente, por lo que no parece necesario buscar estándares adicionales que garanticen su objetividad².

Naturalmente esta convicción es *errónea*, pues las pruebas científicas no constituyen por lo general un razonamiento de tipo deductivo, sino que están basadas prevalentemente en leyes estadísticas cuyos resultados han de ser aún interpretados a la luz de otros datos, y por consiguiente difícilmente puede hablarse de “objetividad” en relación con las conclusiones obtenidas a raíz de las mismas. Pero sobre todo esta convicción entraña un *peligro*, pues propicia la difusa creencia de que las decisiones probatorias apoyadas en pruebas científicas se asuman como incuestionables o irrefutables y, de paso, descarga al juez de hacer un especial esfuerzo por fundar racionalmente la decisión: basta con alegar que hubo prueba científica y que ésta apuntaba justamente en la dirección de la decisión probatoria final.

Seguramente la consecuencia más evidente y adversa de la falta de reflexión crítica sobre el estatuto epistemológico de las pruebas científicas sea la falta de control sobre su validez o fiabilidad, lo que sin duda permite la entrada en el

2. Por lo demás, esa desbordante confianza en la infalibilidad de la prueba científica se ha visto alimentada por el tremendo impacto que en el imaginario popular han tenido series televisivas como CSI, que han generado una especie de beatificación de estas pruebas. De modo que hoy estamos gozosamente dispuestos a creer en los informes prevenientes de los laboratorios de la policía científica como si se tratase de una verdad revelada. (Advierte contra esta actitud J. Igartua, “Prueba científica y decisión judicial (unas anotaciones propedéuticas”, *Diario La Ley*. Núm. 6812, Viernes, 2 de noviembre de 2007.

proceso de auténtica *junk science*, basura sin fundamento científico alguno que es usada a veces por peritos y laboratorios como un jugoso negocio. Piénsese en la grafología, el polígrafo e incluso las huellas dactilares. Pero, sobre todo, el mito de la infalibilidad de la prueba científica³ nos enfrenta al peligro del desconocimiento de los errores judiciales que pueden cometerse con base en ella. El denominado *Innocent Project*, puesto en marcha por los abogados Barry Scheck y Peter Neufeld en la Cardozo Law School para demostrar, mediante pruebas de ADN, la inocencia de un buen número de condenados, ha puesto de relieve no sólo la fragilidad de los medios de prueba tradicionales, como los testimonios y las confesiones, sino también de las pruebas científicas sobre las que se basaban algunas de estas condenas⁴. Por último, el mito de la infalibilidad de la prueba científica entraña un riesgo adicional, un efecto adverso desde el punto de vista jurídico: el de terminar convirtiendo a los peritos en decisores de la causa y, por consiguiente, instaurando un nuevo sistema de prueba fundado en la autoridad de los expertos.

Las anteriores consideraciones apuntan hacia la necesidad urgente de prestar atención a este ámbito del universo probatorio y de articular una reflexión rigurosa sobre las condiciones de validez y el valor probatorio de un tipo de pruebas cuya presencia en la praxis judicial está siendo decisiva en la resolución de muchas causas. Esta reflexión reviste además una importancia crucial, pues no está guiada por un simple prurito epistemológico sino que incide directamente sobre el sistema de garantías y de cautelas que han de rodear esa parte trascendental de la aplicación del derecho. Sólo cuando se toma conciencia de la fragilidad de algunas pruebas científicas, o sea cuando se desconfia del carácter incontrovertible de sus resultados, pueden ponerse en marcha las medidas y controles necesarios para minimizar, y al máximo evitar, los errores judiciales. Cuando, por el contrario, se confía absolutamente (o se actúa “como si” se confiase absolutamente) en la infalibilidad de los resultados de este tipo de pruebas, los controles sobre la calidad de las mismas se relajan en exceso. Lo que desde el punto de vista de la justicia y de los bienes esenciales que están en juego en el proceso puede tener efectos perversos.

2. SOBREALORACIÓN DE LA PRUEBA CIENTÍFICA: LOS PARADIGMAS

La sobrevaloración de la prueba científica se da en una doble vía. Por un lado, sus resultados se aceptan como infalibles. Pero por otro, y sobre todo, se

3. “La llamada *prueba científica* ocupa un lugar mitológico en el saber común, y justamente por la ausencia de conocimiento sobre el tema específico el juez y los abogados —no asesorados— tienden a aceptarla a pie de letra de los dictámenes” (R. Rivera, La valoración de la prueba científica en el proceso judicial, consultado en <http://www.iprocesalcolombovenezolano.org/doctrina/doc2.doc.>, el 14/08/2010).

4. Puede consultarse la página web del proyecto en www.inocenceproject.org

considera que esos resultados dicen cosas distintas de las que en realidad dicen. La primera es, pues, una *sobrevaloración epistemológica*. La segunda es —por así decirlo— una *sobrevaloración semántica*.

2.1. Sobrevaloración semántica de la prueba científica: lo que dicen las pruebas

Si la prueba científica ha sido entronizada en el proceso es porque se da por descontado no sólo el altísimo valor probatorio o incluso infalibilidad de sus resultados, sino también que éstos hablan directamente de aquello que se pretende probar. Se piensa, en concreto, que el resultado de una prueba de ADN señala directamente la pertenencia o no del vestigio analizado a la persona de la que procede la otra muestra de ADN con el que aquél se contrasta (el acusado en una causa penal, por ejemplo); que el resultado de una prueba de balística dice directamente si el casquillo evaluado salió o no de la pistola del acusado; si la impresión de calzado analizada procede de la pisada de una cierta persona; si la escritura manuscrita examinada procede del demandado; si la voz analizada que realizó la llamada telefónica avisando de la colocación del artefacto explosivo pertenece o no al acusado, etc. Se piensa, en definitiva, que el resultado de una prueba científica habla en los términos en que el juez necesita pronunciarse. Esta creencia expresa lo que en el ámbito de la ciencia forense se denomina el *paradigma de la individualización*, que se asienta sobre la supuesta capacidad de llegar a identificar plenamente a un individuo o a un objeto a partir de vestigios. Dicho paradigma se presenta como modelo a seguir para todas las técnicas identificativas en criminalística, y en términos generales consiste en sostener que el resultado de la prueba científica identifica (y además, categóricamente) un vestigio con una fuente, con exclusión de todas las demás⁵.

Las cosas, sin embargo, no son así⁶. Y de hecho en los últimos años el paradigma de la individualización ha recibido fuertes críticas por parte de la co-

5. Aludiendo al paradigma de la individualización, el NRC Report, *Strengthening Forensic Science in the United States*, Washington (USA), February 2009, dice que los analistas creen que algunas marcas son únicas “(típicamente las impresiones de calzado y de neumáticos, las impresiones de crestas procedentes de la dermis, marcas de herramientas y de armas de fuego, y exámenes de escritura manuscrita)” (...) Y creen que “esa unicidad es transmitida fielmente desde la fuente a la evidencia que es objeto de examen (o en el caso de los exámenes de escritura manuscrita, que los individuos adquieren hábitos que dan lugar a una escritura individualizada). Cuando la evidencia y la fuente de donde pueda proceder se comparan, una conclusión de individualización implica que la evidencia se originó desde esa fuente, con exclusión de todas las demás fuentes posibles” (Capítulo 1, Introducción, Sección *Presiones sobre el Sistema de Ciencia Forense*, Párrafo *Ciencia Cuestionable o Cuestionada*).

6. Uno de los ensayos más estimulantes de los últimos tiempos sobre el uso de la estadística en la ciencia forense, la obra de R. Royall *Statistical Evidence*, denuncia explícitamente que el uso de métodos estadísticos estándar conduce muchas veces a tergiversar los resultados de las pruebas: unas veces dándoles un peso mayor o menor del que realmente tienen; otras, considerando que *los datos*

munidad científica, que considera que dicho paradigma se asienta sobre la asimilación entre infrecuencia y unicidad e implica —en palabras de D. H. Kaye— un injustificado “salto de fe” que la ciencia no puede permitirse⁷. Dicho brevemente: no hay fundamento científico alguno para que un perito sostenga que ha sido capaz de identificar a una persona o un objeto a partir de los análisis llevados a cabo en el laboratorio. Pero la crítica va más allá: no hay tampoco fundamento científico alguno para que en el informe pericial se realice siquiera una valoración probabilística sobre la posibilidad de atribuir a una persona o a un objeto el vestigio analizado⁸. El resultado de un análisis de voz, por ejemplo, o el de una comparación de perfiles de ADN, no dice (ni tajante, ni probablemente) que la voz o el ADN del vestigio analizado pertenezca a tal o a cual persona, sino que sólo aporta *datos* que, una vez *interpretados* con las adecuadas herramientas estadísticas, dicen cosas del siguiente tipo: «es X veces más probable que la voz analizada pertenezca al acusado dándose tal rasgo que no dándose»; o «es X veces más probable que el vestigio analizado pertenezca al acusado dándose la coincidencia de perfiles que sin darse».

El llamado por R. Royall *paradigma de la verosimilitud*⁹ levanta acta de esta situación. Dicho paradigma se erige sobre la distinción entre las tres preguntas básicas que cabe formular cuando el perito ha realizado ya los análisis pertinentes en una prueba científica: *¿qué nos dicen los datos* u observaciones resultantes de esos análisis sobre las hipótesis en juego?, *¿qué debemos creer* a partir de esos datos? y *¿qué debemos hacer*? Como enseguida se verá, esta distinción se presenta como una importante herramienta a la hora de reconstruir cómo interacciona con el proceso judicial una prueba que se desarrolla fuera del mismo. Pero la distinción entre estas tres preguntas tiene también una gran trascendencia para un proceso de análisis científico de evidencias tan altamente institucionalizado como es el de la prueba judicial. De hecho —como también se señalará a continuación— permite realizar un deslinde neto de tareas entre el perito y el juez.

estadísticos apoyan un resultado cuando en realidad sucede justo lo contrario (R. Royall: *Statistical Evidence: A Likelihood Paradigm*, Monographs on Statistics and Applied Probability, Chapman & Hall/CRC, London, 1997, Prefacio, p. xi). Y esto último sucede, no porque los expertos usen equivocadamente la estadística, sino justamente porque domina el paradigma de la individualización.

7. D. H. Kaye, *Probability, Individualization and Uniqueness in Forensic Science: Listening to the Academies*, 30th June 2009, Social Science Research Network (SSRN), nota n.º 39, que puede consultarse en el siguiente sitio: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1261970. Como D. H. Kaye comenta (nota n.º 39), existe un riesgo nunca igual a cero al aceptar cualquier inferencia sobre un parámetro poblacional. “La distancia entre la muestra y la población requerirá siempre un salto de fe. Lo único que vale la pena debatir es la longitud del salto”. Cfr. también Saks, M. J., Koehler, J. J., *The Individualization Fallacy in Forensic Science Evidence*, *Vanderbilt Law Review*, Volume 61, Number 1 (January 2008), pp. 199-219.

8. Para un análisis detallado de las críticas recibidas por el paradigma de la individualización, remito al trabajo conjunto Gascón, M., Lucena, J. J. y González, J., “Razones científico-jurídicas para valorar la prueba científica: una argumentación multidisciplinar”, en *Revista La Ley* (en prensa).

9. R. Royall, *Statistical Evidence: A Likelihood Paradigm*, cit.

Responder a la primera pregunta (*qué dicen los datos*) es obviamente tarea del perito, quien, por ejemplo en una prueba de cotejo de voces, deberá *interpretar* el resultado de la misma en el laboratorio (supongamos que positivo) dándole el alcance en relación con las hipótesis enfrentadas (supongamos que *A*: la voz es del acusado; o *B*: la voz no es del acusado). Pero, una vez interpretados esos datos y expresados adecuadamente en el informe pericial, hay que responder a la segunda pregunta (*qué debe creerse a partir de esos datos*), es decir hay que *evaluar* la veracidad de las hipótesis disponibles (*A* y *B*) a partir de esos datos. Y esta es —también obviamente— tarea que debe hacer el juez, y no el perito, a la vista no sólo de lo que dicen esos datos sino también de lo que dicen el resto de las pruebas e informaciones disponibles. Y es que, en efecto, es el juez quien tiene *institucionalmente* atribuida la función de proveer una solución jurídica para un caso controvertido o litigioso y por lo tanto la de determinar previamente cuáles han sido los hechos que han dado origen al conflicto. En otras palabras, es a él a quien corresponde institucionalmente la tarea de creer o no creer la hipótesis que se pretende probar. Pero no sólo eso. También el principio de *libre convicción*, en la medida en que consagra que el juez no deba dar por probado un hecho entretanto no se haya convencido de su verdad, proscribe cualquier valoración predeterminada de la prueba y por consiguiente sólo es acorde con la atribución al juez de la tarea de determinar lo que hay que creer a partir de la prueba. Por lo demás, lo que acaba de decirse puede parecer claro y poco objetable, incluso puede parecer una obviedad, pero merece ser destacado, pues la confusión entre las tareas señaladas no sólo no es infrecuente sino que está en la base del paradigma de la individualización, tan extendido en laboratorios de todo el mundo, sobre todo en relación con las ciencias forenses tradicionales.

En suma, la neta distinción entre la tarea del perito (expresar lo que dicen los datos) y la del juez (valorarlos a la luz de los demás datos y pruebas disponibles) es el núcleo —no importa repetirlo— del paradigma de la verosimilitud, y es la única que se muestra respetuosa con la atribución exclusiva al juez de la tarea de valorar la prueba y con el principio de libre convicción. El paradigma de la individualización, por el contrario, al no distinguir entre la tarea del perito y la del juez, al propiciar que el perito declare en términos de «el vestigio *v* procede de la fuente *f* (por ejemplo, el acusado)», que son los términos en que debe expresarse el juez en ejercicio de la tarea de valoración de la prueba que sólo a él le corresponde, se muestra contraria a derecho. Y no sólo eso. También comporta un riesgo importante de comisión de falacias. Como la conocida falacia del fiscal (*prosecutor fallacy*) a la que más adelante me referiré.

Por último, y aunque sea una cuestión de menor importancia, conviene señalar que el paradigma de la verosimilitud también traza la distinción entre *lo que debe creerse* y *lo que debe hacerse* a partir de la realización de una prueba científica. Dicha distinción resulta sumamente apropiada para reconstruir otra distinción esencial dentro del universo de la prueba: la existente entre la valoración de la prueba y los estándares de prueba. Y es que mientras que la primera tarea (lo que debe *creerse*) puede identificarse —ya se ha dicho— con la *valoración* de la

prueba, la segunda (lo que debe *hacerse*) expresa la cuestión de los *estándares* de prueba; es decir, la determinación de los grados de probabilidad que debe haber alcanzado una determinada hipótesis para que podamos considerarla probada y actuar en consecuencia. “Y así como la valoración de la prueba es (o debería ser) una cuestión completamente librada a la *racionalidad epistémica*, ajena por tanto a condicionamientos jurídicos o valorativos de cualquier tipo, la fijación de estándares de pruebas es una cuestión de *policy*, enteramente dependiente de lo tolerante que el sistema esté dispuesto a ser con cada uno de los dos errores que pueden cometerse al adoptar una decisión: declarar probado lo falso y declarar no probado lo verdadero”¹⁰.

2.2. La sobrevaloración epistemológica: la infalibilidad

Pero la falta de mirada crítica sobre la prueba científica no sólo está relacionada con la errónea interpretación del significado de sus resultados (que hablarían —según el dominante paradigma de la individualización— en términos de identificación del vestigio analizado con una fuente) sino que, además, hay una cierta tendencia a sobrevalorar también el alcance epistémico de los mismos: los resultados de una prueba científica suelen considerarse infalibles. En los planteamientos más reflexivos no se niega, obviamente, que dichos resultados se expresen en términos de probabilidad, pero se actúa acríticamente “como si” esta probabilidad —precisamente porque estamos ante una prueba “científica”— fuese tan extremadamente elevada que pudieran considerarse absolutamente fuera de duda.

Desde luego es muy probable que el mito de la infalibilidad de las pruebas científicas tenga mucho que ver con la prueba del ADN, cuya metodología científica está ya tan perfeccionada y el grado de probabilidad que arroja es tan alto que en la práctica puede “actuar como si” fuese infalible. No obstante, aún dando esto por sentado, conviene adoptar una actitud un poco más crítica, pues la calidad epistémica de los resultados de una prueba científica —y por lo tanto su valor probatorio— es un factor que depende de más variables¹¹.

En primer lugar, la calidad epistémica de una prueba científica depende de la *validez científica y/o metodológica* de la misma. Muchas de estas pruebas, en efecto, pueden realizarse por métodos científicos diferentes, y no todos ellos gozan

10. Cfr. Gascón, M., Lucena, J. J. y González, J., “Razones científico-jurídicas para valorar la prueba científica”, cit., adonde remito de nuevo para una exposición más detenida de las tres preguntas sobre las que se asienta el *paradigma de la verosimilitud*.

11. En su informe *Scientific Evidence in Europe. Admissibility, Appraisal and Equality of Arms*, CDPC (2010), 10, Christophe Champod y Joëlle Vuille realizan un exhaustivo repaso por los factores que pueden influir negativamente en el valor probatorio atribuible a los resultados de las pruebas científicas. A continuación destacamos algunos, pero para una visión completa remitimos a ese excelente estudio [que puede consultarse en www.coe.int/t/e/legal_affairs, pp. 10 ss.].

del mismo crédito en la comunidad científica, de manera que la validez científica del método usado, y con ello la calidad de los resultados alcanzados, pudiera ser objeto de discusión. Incluso en relación con una prueba tan extendida y consolidada como la prueba de polimorfismos ADN, por ejemplo, no son indiferentes las *técnicas o métodos usados* en laboratorio para el análisis de la muestra. Existen muchas técnicas de extracción y tipificación del ADN y no todas ellas ofrecen el mismo rendimiento para los distintos tipos de muestras (semen, sangre, saliva, pelos...). Por eso, dependiendo del tipo de muestra e incluso del soporte en que se encuentre (piedra, madera, cristal, una alfombra) puede ser más adecuado usar una técnica u otra.

En segundo lugar, la fiabilidad y el valor probatorio atribuible a una prueba científica depende también de su *calidad técnica*, es decir de su correcta realización en el laboratorio. Es importante que los laboratorios cuenten con las infraestructuras apropiadas y con personal adecuadamente formado, así como que se sigan rigurosos protocolos en la realización de los análisis y estudios. Además, no sólo cabe hablar de calidad *técnico-científica* sino también de calidad, por así decir, *técnico-procedimental*, en referencia a todo el proceso que conduce desde el descubrimiento o registro del vestigio o de la muestra hasta su análisis en el laboratorio. Por ejemplo, a efectos de atribuir fiabilidad al resultado de una prueba dactilar, el problema no es tanto o no sólo la validez científica de la prueba, sino, en primer lugar, saber quién tomó la huella, por orden de quién, en qué objeto estaba depositada, en qué punto concreto, cómo fue la cadena de custodia, etc. (corrección técnico-procedimental). Y lo mismo con respecto al análisis de una mancha de sangre, orina, saliva¹². Por otra parte, en segundo lugar, es evidente que la validez de la prueba dependerá también de su correcta realización en el laboratorio (corrección técnico-científica): por personal cualificado, siguiendo los protocolos apropiados y —en relación con algunos tipos de pruebas, como la de ADN— prestando atención escrupulosa a las técnicas de asepsia para evitar las contaminaciones durante el trabajo, etc. La regla aquí debería ser: “cuanto mayor es la expectativa de valor probatorio depositado en una prueba, más rigurosos deben ser los controles de realización de la misma”. A este respecto es muy importante la *estandarización*, tanto de los laboratorios (superación de controles de calidad internos y externos) como de la realización misma de la prueba. Y por último, en tercer lugar, no hay que olvidar tampoco que las pruebas las realizan personas

12. Es evidente, por ejemplo, que, pese al potencial de la prueba de ADN, el estado en el que llegan los vestigios biológicos al laboratorio es crucial: si los vestigios no han sido bien recogidos o conservados (por ejemplo, porque han sido contaminados por un ADN extraño) la posibilidad y el rendimiento del análisis se reduce. Por eso la *recogida de indicios* ha de hacerse con sumo cuidado, y el *mantenimiento de la cadena de custodia* es fundamental para que los indicios no pierdan su valor probatorio. Llama la atención a este respecto el entusiasmo con el que muchos países se han abierto a las pruebas científicas (particularmente a las de ADN) sin un marco normativo previo que regule los procedimientos de obtención y conservación de los datos y garantice en consecuencia la fiabilidad de los resultados.

de carne y hueso, y que por lo tanto pueden cometer errores. No son pocos los estudios que en los últimos tiempos llaman la atención sobre los *riesgos cognitivos* de algunas pruebas científicas¹³, sobre todo de aquellas tradicionales que, como la dactiloscopia y la grafística, tienen un fuerte componente comparativo que las deja bajo la supervisión de un perito¹⁴.

Finalmente, conviene no olvidar algo que ya se indicó al principio: que las leyes científicas sobre las que se levantan estas pruebas son en la mayoría de los casos de *naturaleza probabilística*, y los resultados de las mismas han de ser aún interpretados recurriendo a la estadística. En otras palabras, que el resultado de una prueba científica —no importa reiterarlo— es siempre sólo la simple probabilidad, por más alta que esta pueda ser. Precisamente por eso —como veremos más adelante— el modo de formular las conclusiones es muy importante, pues se corre el riesgo de que los jueces atribuyan a las mismas un valor probatorio desproporcionado o en todo caso inadecuado. En genética forense, por ejemplo, como luego veremos, es realmente fácil —y desgraciadamente frecuente— cometer la llamada *falacia del fiscal* (*prosecutor's fallacy*), que malinterpreta el valor real de los datos resultantes de la prueba.

En definitiva, la validez de una prueba científica (y por consiguiente la fiabilidad de sus resultados) no es algo que haya que dar por descontado, sino que depende de la *validez científica del método usado*, de que se haya utilizado la *tecnología apropiada* y de que se hayan seguido rigurosos *controles de calidad*. Por consiguiente, incluso en relación con la prueba de ADN, que con el tiempo ha conseguido un altísimo grado de solidez científica, no puede haber ninguna duda sobre la necesidad de prestar atención a estas dos últimas cuestiones a la hora de evaluar el crédito que dicha prueba merece. Ahora bien, cumplidas todas estas precauciones, el problema de las pruebas científicas no reside ni mucho menos en la prueba del ADN. El problema reside en que hay otras muchas áreas de la policía científica que están muy lejos de haber gozado del mismo grado de atención y desarrollo científico que la genética forense, pero sobre las que sin embargo se fundan diariamente muchas decisiones judiciales.

13. Con respecto a los riesgos cognitivos cfr. Dror, I., "How can Francis Bacon help forensic science? The four idols of human biases", *Jurimetrics Journal*, 50, 2009, pp. 93-110; y Dror, I. E., & Cole, S. A., "The vision in 'blind' justice: Expert perception, judgment, and visual cognition in forensic pattern recognition", en *Psychonomic Bulletin & Review*, 17, 2010, pp. 161-167. Recuérdese el caso Mayfield, en el que tres expertos del FBI y un experto independiente le asociaron una huella detectada en el transcurso de la investigación de los atentados de Madrid de 2004. La asociación se mostró errónea, y que los cuatro peritos cometieran el mismo error se explica porque los tres peritos oficiales, cuando realizaron su trabajo, conocían perfectamente la conclusión a la que había llegado el perito independiente. Es decir se explica por un sesgo cognitivo.

14. Ch. Campod & J. Vuille, *Scientific Evidence in Europe*, cit., pp. 10-11.

3. LAS CONSECUENCIAS DEL PARADIGMA DE LA VEROSIMILITUD: LOS RETOS

Del paradigma de la verosimilitud derivan algunas consecuencias de importancia que plantean importantes retos de cara a la consecución del objetivo final de racionalización de la función de juzgar cuando se utiliza prueba científica. La primera tiene que ver con el control de la validez y fiabilidad de la prueba. La segunda está relacionada con el contenido y sentido del informe pericial, esto es, con el modo en que han de formularse las conclusiones en el informe. La tercera, por último, hace referencia al modo en que ha de valorarse la prueba científica.

3.1. Sobre la admisibilidad de la prueba científica. ¿El juez como *gatekeeper*?

La admisibilidad es el control de entrada de la prueba en el proceso y es, por eso mismo, una tarea que corresponde al juez. Todos los esquemas de control de validez y fiabilidad (en definitiva, de calidad) de las pruebas científicas tienen como objetivo común dejar a la pseudociencia fuera del ámbito de las decisiones judiciales. Cómo conseguir esto es lo que marca la diferencia entre ellos. En línea de principio cabe adoptar dos grandes posturas sobre la admisibilidad de la prueba científica. Cabe sostener, por un lado, que *sólo debe haber un control de admisibilidad procesal* de la prueba (esto es, un control de su pertinencia, necesidad y legalidad) y que, por consiguiente, todas las pruebas científicas que pasen ese control deben ser admitidas, dejando para el momento de la valoración la cuestión de su validez científica¹⁵. Pero cabe plantear también la necesidad de incorporar la cuestión de la validez científica de la prueba al control de admisibilidad de la misma¹⁶. Lo que se sostiene, en definitiva, es que *junto al control de admisibilidad procesal* (pertinencia, necesidad y legalidad) *hay que agregar otro de admisibilidad científica* (validez científica).

En Europa en general la postura que se ha seguido para filtrar la prueba científica que no cumple los estándares de validez y fiabilidad es la primera. Las pruebas científicas pertinentes y legalmente válidas entran en el proceso sin traba alguna y son valoradas por el juez, que les atribuye mayor o menor fuerza probatoria dependiendo, entre otros factores, de su validez científica y fiabilidad. En otras palabras, la validez y fiabilidad de una prueba científica no se evalúa en el trámite de admisibilidad de la prueba, sino que se pondera en el momento de la valoración y se refleja en el valor probatorio que se le atribuye a dicha prueba.

15. Lo que se sostiene, en definitiva, es que, si la prueba no es válida, ya se encargará el juez en el momento de la valoración de privarla de valor probatorio.

16. Puede consultarse *in extenso* una confrontación de los modelos de control aludidos en A. M. Hayajneh & S. H. Al-Rawashdeh, "Theoretical Approaches to Admitting Scientific Evidence in the Adversarial Legal System", en *European Journal of Scientific Research*, Vol. 41, N.º 2 (2010), en particular, pp. 187 ss.

Pero —repárese— de este modo la cuestión de la validez científica de la prueba se confunde con la de su fuerza probatoria.

Desde luego no puede desconocerse que tal modelo de control presenta una importante ventaja desde el punto de vista jurídico, pues deja enteramente en manos de los jueces la decisión sobre la admisibilidad de la prueba, lo que en principio se acomoda a la perfección a la atribución exclusiva a los mismos del poder de juzgar en general y de valorar la prueba en particular; y de paso se acomoda también, implícitamente, a la exigencia de que tal control venga realizado por una instancia independiente, como los jueces son. Sin embargo me parece que la opción comentada presenta más problemas que ventajas. En efecto, dejar enteramente en manos de los jueces —a través del trámite de valoración de la prueba— la libre (o mejor, libérrima) apreciación de la validez y fiabilidad de la prueba científica sin establecer al propio tiempo estándares objetivos para determinar dicha validez puede acarrear dos consecuencias igualmente adversas, resultado ambas de la falta de conocimientos expertos por parte de los jueces. Por un lado puede conducir a rechazar pruebas perfectamente válidas y fiables desde el punto de vista científico, que podrían haber ayudado al esclarecimiento de la verdad en el proceso. Por otro puede conducir a dar crédito a pruebas que son pura charlatanería pseudocientífica y que pueden convertirse en fundamento de decisiones judiciales erróneas que a veces son de máxima gravedad, como las decisiones condenatorias. Por eso parece más sensata la opción alternativa, es decir la que apuesta por la instauración de un control de validez científica en el trámite de admisibilidad guiado por parámetros objetivos. Y es que si la prueba no es científicamente válida, si no ha pasado unos mínimos controles objetivos de validez científica y fiabilidad, ni siquiera debería entrar en el proceso, evitando así que despliegue ninguna fuerza probatoria, sea mucha o poca. No puede desconocerse la gran capacidad de persuasión e influencia que una prueba científica puede tener sobre la convicción judicial, con lo que si se admite su entrada en el proceso se corre el riesgo de que, al final, la decisión venga determinada por pura y simple *junk science*¹⁷. Por eso parece lógico que “el sistema judicial deba ser diseñado para prevenir, más que para remediar, los extravíos de la justicia que resultan de admitir pseudo ciencia o mala ciencia”¹⁸.

Situar el control de validez científica de la prueba en el control de admisibilidad es lo que ha sucedido prevalentemente en Estados Unidos, donde en los últimos tiempos se ha producido además un importante cambio en el test o criterio utilizado para efectuar dicho control. Durante la mayor parte del siglo XX el control de admisibilidad fue el instaurado por la sentencia *Frye*¹⁹, que —por decirlo resumidamente— simplemente exigía que la prueba gozase de una *gene-*

17. La admisión al proceso de prueba científica no fiable “lead the judges [a]stray and thus undermine the factual accuracy of criminal adjudication”, Paul Roberts y Adrian Zuckerman, *Criminal Evidence*, Oxford University Press, Oxford, 2004, p. 69.

18. A. M. Hayajneh & S. H. Al-Rawashdeh, “Theoretical Approaches to Admitting Scientific Evidence in the Adversarial Legal System”, cit., p. 187.

19. *Frye v. United States* 293 F 1013 (D.C. Cir. 1923).

ral acceptance en la comunidad científica. Ahora bien, el *Frye test* no representó precisamente un paradigma de control objetivo realizado además por instancias independientes, pues la *general acceptance* no sólo ha sido un criterio sumamente impreciso y abierto que ha permitido que entraran en el proceso todo tipo de pruebas pseudocientíficas, sino que, además, supuso dejar la admisibilidad, de hecho, en manos de los científicos²⁰.

Los criterios de admisibilidad de la prueba se endurecieron con la sentencia *Daubert*²¹, que, además del criterio de la *general acceptance* por la comunidad científica concernida, incorporó al test de admisibilidad la *validez científica* de la técnica aplicada. La validez científica del *Daubert test* ha de acreditarse evaluando los siguientes factores: (a) si la teoría o la técnica se puede probar y si ha sido efectivamente puesta a prueba (verificabilidad y refutabilidad como criterios mínimos de científicidad de cualquier teoría); (b) si se han realizado publicaciones especializadas en el tema; (c) cuáles son los márgenes de errores potenciales asociados al método, y (d) cuáles son —si es que existen— los procedimientos estandarizados y los controles de calidad que rigen la práctica. No es que el *Dauber test* exija que cualquier técnica científica usada como fuente de prueba cumpla todos los criterios mencionados, pues esto es algo que queda al criterio del juez. Pero sí exige a los jueces —y esto es lo más significativo— escrutar de manera más estricta muchas pruebas científicas que hasta ahora no habían sido cuestionadas y conduce así a que éstos recuperen su papel de *gatekeeper*²². Por

20. Con lo cual, en determinados momentos, y dependiendo de las instancias consultadas, es muy posible que determinadas áreas de la ciencia forense no suficientemente desarrolladas quieran ser promocionadas, y aparezcan entonces revestidas de una *general acceptance* relativamente fácil de acreditar.

21. *Daubert v. Merrell Dow Pharmaceuticals Inc.*, 113 S.Ct.2786 (1993).

22. Prácticamente una réplica del *Dauber test* cabe encontrar en los criterios sugeridos para el Reino Unido por la *Law Commission* (The Law Commission, “The Admissibility of Expert Evidence in Criminal Proceedings in England and Wales — A New Approach to the Determination of Evidentiary Reliability — A Consultation Paper”, http://www.lawcom.gov.uk/expert_evidence.htm Consultation Paper No 190, 2009, pp. 53 ss.).

“In determining whether scientific (or purportedly scientific) expert evidence is sufficiently reliable to be admitted, the court shall consider the following factors and any other factors considered to be relevant:

(a) whether the principles, techniques and assumptions relied on have been properly tested, and, if so, the extent to which the results of those tests demonstrate that they are sound;

(b) the margin of error associated with the application of, and conclusions drawn from, the principles, techniques and assumptions;

(c) whether there is a body of specialised literature relating to the field;

(d) the extent to which the principles, techniques and assumptions have been considered by other scientists — for example in peer-reviewed publications — and, if so, the extent to which they are regarded as sound in the scientific community;

(e) the expert witness’s relevant qualifications, experience and publications and his or her standing in the scientific community;

(f) the scientific validity of opposing views (if any) and the relevant qualifications and experience and professional standing in the scientific community of the scientists who hold those views; and

decirlo muy sumariamente, la sentencia Dauber supone un llamamiento a los jueces para que miren más críticamente las pruebas científicas²³. De hecho, después de esta sentencia, ámbitos tradicionales de la policía científica, como las huellas papilares, la grafística y la balística empiezan a ser puestos seriamente en cuestión.

Pero el *Dauber test* tampoco ha zanjado la polémica sobre el sistema de control de la validez y fiabilidad de las pruebas científicas. Ni mucho menos. Precisamente por dejar también en manos del juez la decisión de si una técnica determinada puede o no ser considerada como científicamente válida, el *test* ha sido objeto de críticas: primero por la inseguridad jurídica que deriva del hecho de que cada juez pueda decidir según su criterio; y segundo, y sobre todo, porque los criterios sentados por *Dauber* presuponen en todo caso que el juez es capaz de comprender y manejar conceptos científicos que la mayoría de las veces son complejos, lo que acaso no siempre sea cierto²⁴.

Con todo, hay que decir que, más allá del sistema por el que se opte, la admisibilidad de las pruebas depende mucho de si se da o no por descontada su fundamentación científica. Hoy en día, por ejemplo, ya no se discute la admisibilidad de las pruebas basadas en la genética forense. Recordemos a este respecto —como se dijo más arriba— que ha sido precisamente la prueba del ADN la que se ha utilizado en el *Innocence Project* para esclarecer numerosos errores judiciales. Ello no significa —conviene subrayar— que en esta prueba no siga habiendo muchas cuestiones debatibles²⁵ y que en todo caso puede haber fallos técnicos en su realización. Significa tan sólo que el nivel de investigación y discusión científica sobre esta prueba es muy alto y el consenso alcanzado sobre la misma también, de modo que hoy es comúnmente aceptado que si en su realización se han seguido rigurosos controles de calidad, no hay motivo alguno para cuestionar su admisibilidad. No en vano este tipo de pruebas se han terminado convirtiendo en el modelo a seguir por las demás²⁶.

(g) whether there is evidence to suggest that the expert witness has failed to act in accordance with his or her overriding duty of impartiality.”

23. Con elocuentes palabras Ch. Champod y J. Vuille expresan esta idea: la sentencia Dauber “insiste de manera implícita sobre el escepticismo que el juez debe mantener hacia el perito, el cual deja de ser considerado como el miembro de una élite con autoridad y pasa a ser un agente social comparable a cualquier otro, eventualmente sometido a presiones de orden político y económico que pueden alterar su dictamen”, *Scientific Evidence in Europe*, cit., p. 26.

24. Precisamente por eso, en su informe de 2009, *The Law Commission* propone para el Reino Unido la posibilidad de que excepcionalmente, en los casos muy delicados, el juez pueda solicitar el auxilio de un experto para evaluar la validez científica de la técnica propuesta. Estimo que esta sugerencia es muy interesante.

25. *Vid.*, por ejemplo Thompson, W. C., “Tarnish on the ‘gold standard’: understanding recent problems in forensic DNA Testing”, en *The Champion*, 2006, pp. 10-16; o Balding, D. J., & Buckleton, J., “Interpreting low template DNA profiles”, en *Forensic Science International: Genetics*, 4, 2009, pp. 1-10.

26. Saks y Koehler, por ejemplo, hacen un llamamiento para que los otros ámbitos de las ciencias forenses se unan al rigor y transparencia atribuidos a la genética forense (Saks, M. J., &

En todo caso, aunque poco a poco se consolidaran científicamente las diferentes áreas de la ciencia forense aproximándose así a los niveles de calidad y consenso de que hoy gozan las pruebas de ADN, la admisibilidad de la prueba científica sigue y seguirá siendo una cuestión problemática. Ello es así, en primer lugar, porque hay y siempre habrá especialidades de la ciencia forense con mayor fragilidad científica que la prueba de ADN²⁷, y, como se dijo un poco más arriba, el *Dauber test* (y en general cualquier otro test que pretendiera que los jueces aplicasen estrictamente una serie de criterios científicos de admisibilidad de las pruebas) da por sentado que los jueces son capaces de comprender y aplicar los conceptos científicos necesarios para evaluar la validez de estas pruebas; pero esto es algo de lo que cabe dudar. No debe extrañar por ello que para muchos cada vez esté más clara la necesidad de formar a los jueces en los distintos ámbitos de especialidad forense, y sobre todo la necesidad de proporcionarles una formación específica en las nuevas áreas que vayan surgiendo. Obviamente no se trata de convertir a los jueces en *amateur scientists*, pero sí de tener claro que, si han de poder aplicar estrictamente los criterios científicos de admisibilidad de la prueba, entonces acaso deban recibir previamente algún adiestramiento en este sentido. Por lo demás, en segundo lugar, estamos ante una cuestión problemática porque, precisamente en contestación a la sugerencia de formación de los jueces en las principales áreas de especialidad de las pruebas científicas, cabe diseñar una tercera respuesta a la pregunta de quién y/o cómo debe controlarse la admisibilidad de una prueba científica en el proceso: ahora ya no 1) los jueces, libremente o sin ajustarse a regla previa alguna, en el momento de valoración; ni tampoco 2) los jueces aplicando estrictamente los criterios técnicos de admisibilidad (validez y fiabilidad) científica de la prueba, para lo que —como se ha dicho— habrían de recibir algún adiestramiento previo; sino 3) *los jueces, pero ayudados por los peritos, o bien cualquier otra instancia independiente*.

Por último, es importante observar también que la discusión sobre cuál ha de ser el modelo de control de la validez de la prueba científica se conecta con otra cuestión teórica de más calado que en mi opinión debe ser elucidada: la de la distinción entre la *validez o fiabilidad científica* de la prueba y su *valor probatorio*. Recuérdese a este respecto que la opción por ubicar el control de validez científica en la valoración de la prueba (la opción europea, por así decirlo) se asienta en una asimilación de ambos conceptos. Esclarecer esta distinción es, pues, otra importante tarea pendiente.

Koehler, J. J., "The coming paradigm shift in forensic identification science", en *Science*, 309, 2005, pp. 892-895). Y esta misma necesidad es apuntada en el último informe NRC (2009).

27. En cualquier caso, es importante resaltar que en los últimos tiempos se han ido elaborando tanto a escala nacional, como transnacional, importantes protocolos de realización de las pruebas científicas que pretenden maximizar la fiabilidad de estos análisis. Y hay que decir además que la mayoría de los laboratorios —al menos los oficiales— se ajustan a dichos protocolos.

3.2. Sobre la formulación de las conclusiones

Aparte del central asunto de la fiabilidad de la prueba científica, hay otras dos importantes cuestiones que afectan a la calidad de los resultados obtenidos mediante este tipo de prueba y a los que desgraciadamente no se ha prestado demasiada atención: cómo (y quién) ha de valorar esos resultados y cómo han de expresarse los mismos por los peritos²⁸. La primera cuestión se conecta con la más amplia de la valoración de la prueba, y será abordada en el siguiente epígrafe. La segunda tiene que ver simplemente con el modo de formular los informes periciales para que pueda hacerse una correcta valoración de los mismos.

Recordemos que seguramente la característica más significativa del paradigma de la verosimilitud (que, recordémoslo también, es el que —frente a la práctica más extendida en laboratorios de todo el mundo— es recomendado por los científicos) es la distinción entre *lo que dicen los datos* resultantes de los análisis que constituyen las pruebas y *lo que hay que creer* a partir de esos datos. La segunda operación es tarea del juez. La primera tarea corresponde en cambio incuestionablemente al perito. Se trata, es evidente, de tareas distintas, pero íntimamente relacionadas. Más concretamente, la interpretación de los datos y la correcta comunicación de la misma al juez es *conditio sine qua non* para que el juez pueda realizar posteriormente una correcta valoración de la prueba científica.

Pero ¿qué quiere decir que una vez realizados los análisis en los laboratorios y obtenidos unos datos éstos han de ser aún interpretados? *Interpretar* los datos analíticos resultantes de la prueba quiere decir determinar su significado en el concreto contexto del caso. Pero para verlo más claramente tomemos el ejemplo de una prueba de ADN y de otra de locutores de voz.

La prueba de ADN consiste en la comparación de perfiles genéticos: uno procedente del vestigio analizado y otro de la persona con la que se quiere comparar el primero, por ejemplo el acusado. El resultado de la comparación puede ser la *exclusión* (no coincidencia de los perfiles comparados) o la *no-exclusión* o coincidencia (*match*). Cuando el resultado de la comparación es la exclusión, dicho resultado puede aceptarse en principio como infalible, aunque en la práctica se aconseja repetir la prueba para verificar que no ha habido fallos o errores. Pero si el resultado fuera la no-exclusión o coincidencia, habrá que *valorar aún esa coincidencia*; más exactamente, habrá que valorar la probabilidad de que el vestigio analizado provenga de ese individuo, y esto es algo que dependerá del porcentaje de individuos de la población general que presentan ese perfil genético. Ello explica la intensa labor de análisis de genética poblacional que se realiza en los laboratorios de genética forense. Si ese porcentaje fuera del 50% (la mitad de los individuos presentan ese rasgo del fenotipo) habría que concluir que el valor

28. De ello se lamentaba hace poco, entre otros, Willis, S. M., “Forensic science, ethics and criminal justice”, en J. Fraser & R. Williams (Eds.), *Handbook of forensic science*, Cullompton, UK: Willan Publishing, 2009, pp. 523-545.

probatorio de la coincidencia es bajo; pero si el porcentaje fuera del 1% (uno de cada cien individuos presentan ese rasgo fenotípico) entonces habría que concluir que el valor probatorio de la coincidencia es muy alto²⁹. Estos son los datos interpretados que el perito debe expresar en su informe.

La prueba de locutores de voz, por su parte, consiste también en cotejar en el laboratorio dos voces, una dudosa (procedente, por ejemplo, de una llamada anónima amenazante) y otra indubitada (normalmente la de un sospechoso). Una vez comparadas las voces, el resultado puede ser igualmente la exclusión o la no exclusión. Si el resultado es la no exclusión, ésta deberá ser aún interpretada o valorada por el perito teniendo en cuenta también las características de funcionamiento del sistema, en concreto las tasas de falsos positivos y de falsos negativos que produce. Estos son los datos interpretados que el perito deberá expresar en su informe.

Puesto que el juez debe conocer los datos de la prueba una vez interpretados para poder hacer una correcta valoración de los mismos, una cuestión que adquiere una importancia crucial es el modo en que el perito transmite al juez esos datos; es decir, *cómo se formulan o redactan las conclusiones del informe pericial*. Pues bien, a efectos de redactar apropiadamente el informe para que el juez conozca exactamente cuál es el alcance de esos datos y evitar malinterpretaciones de los mismos, hay un elemento importantísimo a tomar en consideración. Es el siguiente: las conclusiones de la prueba que el perito formula en su informe se orientan a determinar *el significado (o el grado de probabilidad) de los datos analíticos o técnicos resultantes a la luz de todas las hipótesis en consideración, y no el significado (o el grado de probabilidad) de las hipótesis en consideración a la luz de los datos analíticos o técnicos resultantes*. Entender que las conclusiones de la prueba hablan en estos últimos términos, esto es, que las pruebas le dicen al juez justamente lo que él debe determinar para fundar su decisión, sin necesidad de tomar en cuenta otras informaciones de las que sólo él dispone, es propio del paradigma de la individualización, que no distingue entre lo que dicen los datos y lo que debe creerse a partir de los datos, y da lugar a múltiples malinterpretaciones y en ocasiones a serios errores judiciales. Esto, por ejemplo, es lo que explica que se produzcan las conocidas como falacia del fiscal y falacia de la defensa.

Para ilustrar en qué consisten estas falacias imaginemos el siguiente caso. En Madrid, una mujer es agredida violentamente en la oscuridad de la noche mientras caminaba hacia su casa por un individuo al que no puede reconocer. Como consecuencia del forcejeo mantenido queda una mancha de sangre del agresor en la ropa de la víctima. A raíz de otras agresiones posteriores que tuvieron lugar en la misma zona se detiene a un individuo, al que se acusa también de esta agresión. La prueba del ADN muestra que hay una coincidencia (*match*) entre la sangre

29. Vid. sobre el particular Carracedo A., Barros F., *Problemas bioestadísticos en Genética Forense*. Santiago de Compostela. Universidad de Santiago de C., 1996; y Aitken GGG. *Statistics and the Evaluation of Evidence for Forensic Scientists*. New York. Wiley, 1995.

analizada y la del acusado, y según declaran los expertos la frecuencia del perfil genético (del rasgo fenotípico) de la mancha en la población de referencia es del 1%.

El *fiscal* podría presentar el caso así: “Sólo un 1% de la población presenta ese rasgo, de modo que sólo hay un 1% de probabilidades de que la sangre proceda de otro que no sea el acusado. Ello significa que la probabilidad de que el acusado sea inocente es del 1%. Luego el acusado tiene un 99% de probabilidad de ser culpable”. La prueba tiene mucho valor probatorio. El *defensor* por su parte podría presentar el caso así: “El 1% de la población posee ese grupo de ADN. Lo que significa que en una ciudad como Madrid ese rasgo lo poseerían 5000 personas. Por tanto, el acusado es una más entre las 5000 personas que han podido cometer el crimen. Una probabilidad bajísima”.

Ambas interpretaciones parecen intuitivamente acertadas, y de hecho se ha demostrado que cuando se presenta cualquiera de las dos —por separado— la mayoría de la gente piensa que es correcta³⁰. Además, si los peritos presentan el informe de forma aparentemente aséptica (“el grupo lo posee el 1% de la población”) muchos individuos hacen alguna de las dos interpretaciones. Sin embargo serían incorrectas, pues no toman en cuenta la existencia de otros datos o pruebas que puedan “apuntar” la inocencia o la culpabilidad. El fiscal no tiene en cuenta la existencia de otros datos obrantes en la causa que pudieran apuntar la hipótesis de la inocencia. El defensor, por su parte, tampoco tiene en cuenta que —pese a que existan 5000 personas que pueden presentar esa coincidencia— pueden existir otros datos o pruebas en la causa que apunten la hipótesis de la culpabilidad. En definitiva, para interpretar el valor incriminatorio de la prueba de ADN ha de tenerse *también* en cuenta la probabilidad (subjética) previa (obtenida sobre la base de otras pruebas)³¹ de que el acusado sea culpable³².

Evitar la falacia del fiscal y otras falacias argumentales que pueden cometerse requiere que el perito formule las conclusiones del informe pericial de manera que no se permitan este tipo de interpretaciones. Por eso, si los datos hablan de la verosimilitud de un cierto resultado a la luz de las hipótesis en liza, es en esos

30. Vid. Carracedo, A., Barros, F., Lareu, M. V., Pestoni, C. y Rodríguez-Calvo, M. S., “The Evaluation of the Evidence in DNA Typing”, *Sc Justice*, 1996, 36 (3), pp. 204 ss; y Carracedo, A. “Valoración de la prueba del ADN”, cit., pp. 303 ss.

31. La *probabilidad subjética* indica el grado de creencia personal de un factfinder racional sobre el acaecimiento de un suceso incierto. Vid., sobre este concepto, I. Hacking, *The Emergence of Probability: a Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference*, Cambridge University Press, 3.ª reimp., 1987.

32. Sobre la *falacia del fiscal* y la *falacia de la defensa*, cfr. Thompson, W. y Schumann E., “Interpretation of statistical evidence in criminal trials. The prosecutor’s fallacy and the defense attorney fallacy”, *Law Human Behav.* 1987, 11, pp. 167 ss. Y entre nosotros Carracedo, A. “Valoración de la prueba del ADN”, en *La prueba del ADN en la medicina forense. La Genética al servicio de la ley en el análisis de indicios criminales y en la investigación biológica de la paternidad* (Begoña Martínez Jarreta dir.), Barcelona, Masson, 1999, cit., p. 303. Sobre la falacia del fiscal puede consultarse más recientemente Leung, W. C., “The prosecutor’s fallacy — a pitfall in interpreting probabilities in forensic evidence”, *Medicine Science and the Law*, 42, 2002, 44-50.

términos, y no en términos de individualización, en los que ha de redactarse el informe pericial. Probablemente no hay fórmulas mágicas, y probablemente no hay una sola fórmula. Pero lo que sí está claro es que debe huirse de las expresiones (categóricas o incluso sólo probabilísticas) sobre las hipótesis planteadas por las partes con respecto a un cotejo criminalístico, y debe en cambio optarse por pronunciamientos del tipo «es R veces más probable que coincidan los perfiles genéticos si el vestigio x procede de la fuente f que si no procede de f »; o «es R veces más probable que coincidan las huellas dactilares si la huella es del acusado que si no lo es»³³. En este sentido un importante reto que la ciencia forense tiene ante sí es el establecimiento de unos estándares de conclusiones que guíen el trabajo de los laboratorios y que al propio tiempo aseguren el uso de criterios uniformes en este punto.

Por lo demás, y aunque diéramos por sentado que los estándares de conclusiones existen y que los laboratorios se ajustan a los mismos, el informe puede ser complejo y objeto de malinterpretaciones. Por eso es necesario que, como regla general, se requiera la presencia del experto en el acto de juicio. Si esta insistencia tiene sentido es porque en los últimos tiempos algunas voces sostienen que no es necesario que los peritos de las instituciones oficiales —esto es, aquellos cuya independencia debemos presumir— acudan al proceso, puesto que lo que tenían que decir ya lo han dicho en su informe, y en consecuencia abogan por una reforma legislativa que confiera a los informes periciales de los laboratorios de la policía científica carácter de prueba documental. Este tipo de propuestas me parecen profundamente desacertadas. Podrán servir —esto no se cuestiona— para agilizar el acto del juicio (puesto que no será necesario que declaren los peritos sino que el juez, por su cuenta, examinará el documento del informe pericial), así como para optimizar los recursos de los laboratorios oficiales y facilitar el trabajo de sus expertos (que no se verán obligados a acudir al tribunal). Pero, al menos en el proceso penal, lesionan valores fundamentales vinculados al derecho de prueba, como la contradicción, la intermediación y la oralidad, y desde luego se alejan del objetivo epistemológico que (también) subyace a la organización institucional de la prueba como mecanismo de averiguación de la verdad. Y es que el juez, por lo general, no posee los conocimientos necesarios para interpretar informes periciales complicados y ponderar su alcance exacto, de modo que la presencia del perito en la sala de juicio representa en estos casos una garantía epistemológica.

3.3. Sobre la valoración de la prueba

Son dos las cuestiones relevantes que se plantean en relación con la valoración de la prueba: quién debe valorar y cómo se realiza la valoración.

33. Me permito remitir aquí, de nuevo, a nuestro trabajo Gascón, M., Lucena, J. J. y González, J., “Razones científico-jurídicas para valorar la prueba científica”, cit.

Quién debe valorar, es decir quién debe determinar lo que hay que creer a raíz de la prueba, es algo que está muy claro. Si es responsabilidad de los peritos presentar el resultado de la prueba de forma que pueda ser correctamente interpretado por el juez, es responsabilidad del juez interpretar esos resultados correctamente y atribuirles un determinado peso en la formación de su convicción sobre el hecho principal. Es el juez, en definitiva, quien, en virtud del principio de libre convicción, debe decidir sobre el valor probatorio que ha de atribuirse al informe pericial. Suele decirse a este respecto que el juez es *peritus peritorum*, destacando así que es él quien tiene la última palabra sobre el valor que debe ser atribuido a los datos expresados en el informe pericial. Ahora bien —se preguntan algunos— ¿no resulta paradójico decir esto y reconocer al propio tiempo que el juez no es un experto? Porque si no siendo un experto carece de los conocimientos necesarios para formarse una opinión propia sobre la prueba realizada, ¿cómo va a quedar en sus manos la atribución de valor probatorio a los resultados de la prueba científica?

La pregunta, a mi juicio, está mal formulada; o más exactamente, su formulación se vincula al paradigma de la individualización, en el seno del cual el informe del experto versa justamente sobre lo que se pretende probar: «el ADN encontrado en la colilla pertenece (o no pertenece) al acusado»; «la voz analizada pertenece (o no pertenece) al acusado»; «la huella dactilar encontrada en el arma homicida es (o no es) del acusado», etc. Sin embargo, en el paradigma de la verosimilitud las cosas no son así. El informe pericial —no importa repetirlo una vez más— no valora directamente estos enunciados. No habla de la probabilidad de las hipótesis judiciales en juego a la luz de los datos analíticos resultantes de la prueba, sino de la probabilidad de esos resultados a la luz de esas hipótesis. Quien tiene la tarea de valorar esos enunciados es el juez, que debe hacerlo además tomando en consideración no sólo los datos del informe sino también los demás datos que obren en la causa. Por eso en este paradigma no tiene sentido decir que el juez es *peritus peritorum*. No tiene sentido decir que ha de valorar el informe pericial y que el resultado de esa valoración es, o bien *apartarse* del mismo (para lo cual necesitaría tener unos conocimientos expertos que no pueden suponersele), o bien *plegarse* al mismo (en cuyo caso parece que se vulneraría el principio de libre valoración de la prueba). En el paradigma de la verosimilitud lo que el juez debe hacer cuando valora el informe no es o bien seguirlo, o bien abandonarlo. Es otra cosa: es determinar lo que hay que creer sobre las hipótesis en conflicto, a partir de los datos del informe pericial y del resto de los datos y pruebas que obran en la causa.

En este sentido una cuestión que parece relevante es la relativa a la formación matemático-estadística de los jueces y, en general, de los principales actores del proceso³⁴. La estadística, en efecto, juega un papel crucial en la interpretación de

34. Víctor Fairén Guillén recuerda que si la idea de tener jueces-psiquiatras, jueces-ingenieros, etc. (avanzada por Jiménez de Asúa) parece utópica, se abre camino en cambio la de tener *jueces-científicos* o, en su defecto, la del *juez vinculado al dictamen de los peritos*. Como por regla general, sostiene el autor, no se podrá pedir al juez “que posea una ciencia igual o superior a la del científico,

los datos de la mayoría de las pruebas científicas; y precisamente por eso los peritos tienen que redactar las conclusiones del informe en términos de verosimilitud. Es decir, que los datos expresados en las conclusiones del informe pericial que después habrán de ser valorados por el juez son en la mayoría de los casos datos estadísticos. Ahora bien, está comprobado el peso desproporcionado (*overweighting thesis*) que el público en general y los jueces en particular atribuyen a la estadística³⁵. Precisamente por eso no parece descabellado plantear la necesidad de que los jueces reciban alguna formación en este sentido. No se trata —no importa repetirlo una vez más— de convertir a los jueces en matemáticos, o en *amateur scientists*, sino de proporcionarles los conocimientos necesarios para conjurar el riesgo de sobrevaloración de los datos estadísticos en los que se expresa el resultado de una prueba científica.

Pero pasemos a la cuestión central: *¿en qué consiste y cómo se ha de realizar esa valoración?*

Valorar la prueba científica consiste en determinar lo que hay que creer a partir de las conclusiones del informe pericial. Consiste, en otras palabras, en atribuir a las conclusiones del informe un peso en la formación de su convicción sobre las hipótesis de las partes. Esta valoración debe hacerse tomando en consideración los datos estadísticos aportados en el informe pericial y ponderándolos con el resto de las pruebas e informaciones obrantes en la causa. Por supuesto el juez realiza normalmente esta valoración sin auxilio alguno de las matemáticas; “*subjetivamente*”, por así decirlo. Sin embargo la valoración de la prueba científica también puede realizarse recurriendo a instrumentos matemáticos.

La fórmula que permite poner en relación el valor estadístico de una prueba científica con el valor atribuido anteriormente por las otras pruebas e informaciones obrantes en la causa al enunciado fáctico que se pretende probar es el Teorema de Bayes, que en su formulación más simple afirma que la probabilidad de un evento H, dado un evento E, puede determinarse en función de la frecuencia estadística con la que dado H se verifica E y de la probabilidad atribuida precedentemente al evento H.

$$P(H/E) = \frac{P(E/H)}{P(E/\text{no-H})} \cdot P(H)$$

P(H/E): probabilidad de H, dado un evento E (*posterior probability*).

P(E/H): frecuencia estadística con la que, dado H, se da E.

que sería lo ideal para que pudiera rehacer por su cuenta la cabal valoración efectuada por aquél”, deberemos contentarnos con que controle el grado de *aceptabilidad* de los nuevos métodos científicos (*Doctrina General del Derecho Procesal*, Barcelona, Bosch, 1990, p. 463).

35. Por ceñirnos sólo a Europa, son conocidos, entre otros, el caso de Sally Clark, en Inglaterra (descrito por Johnson, 2004), o el de Lucia de Berk, en los Países Bajos, absuelta en abril de 2010. Ambos, y algunos más, son citados como ejemplo de lo dicho por Ch. Champod y J. Vuille, *Scientific Evidence in Europe*, cit., p. 10.

P (E/no-H): frecuencia estadística con la que, sin darse H, se da E.

P (E/H) / P (E/no-H) (*Likelihood ratio* de E respecto de H, que mide la mayor o menor probabilidad con que E será verdadera siendo H verdadera que sin serlo).

P (H): probabilidad atribuida a H antes de saber si E se daba (*prior probability assessments*).

La utilidad del Teorema en el proceso reside en que permite combinar información estadística sobre un cierto suceso (la procedente del informe pericial y expresada mediante la *likelihood ratio*) con información no estadística (la *prior probability*)³⁶. Y el resultado de esa combinación es la *posterior probability*, que expresa el resultado de la valoración de esa prueba. En realidad, como puede verse, el Teorema de Bayes (o más exactamente la *posterior probability*) mide el impacto que, sobre el hecho que se pretende probar, provoca la inclusión de nuevos datos de naturaleza estadística, como son los resultantes de las pruebas científicas.

Supongamos, por ejemplo, que se ha realizado una prueba de locutores de voz para averiguar si un sospechoso *sl* ha sido el autor de la llamada con la que se avisaba de la colocación de un artefacto explosivo, y supongamos también que el resultado del cotejo de voces ha sido positivo. La valoración de esta prueba recurriendo al Teorema de Bayes ha de hacerse tomando en consideración, de un lado, ese resultado positivo una vez interpretado por el perito a la vista de la tasa de falsos positivos y falsos negativos del sistema, y expresado en el informe mediante datos estadísticos; y de otro lado, las probabilidades *a priori* atribuidas precedentemente a ese resultado positivo, y que también han de ser cuantificadas (la probabilidad de que el sospechoso pudiera ser autor de la llamada dudosa antes de que el perito realizase el cotejo criminalístico y su probabilidad complementaria). Si aparte de la prueba científica no hay ningún otro dato o prueba que atribuya alguna probabilidad a la hipótesis de que el sospechoso puede ser el autor de la llamada, entonces la *prior probability* de esa hipótesis será muy baja y el resultado de la valoración mediante la aplicación del teorema de Bayes también. Si por el contrario disponemos de muchas pruebas o informaciones que permiten pensar que el sospechoso probablemente fue el autor de la llamada, entonces la *prior probability* será seguramente bastante alta y el resultado de la valoración mediante el teorema de Bayes también.

Lo anterior no supone desconocer los problemas que aún hoy presenta la introducción de las matemáticas en el proceso de valoración. De hecho, el bayesianismo como modelo de valoración ha sido objeto de numerosas críticas. Algunas de estas críticas están animadas por consideraciones *prácticas*. Por ejemplo, llaman la atención sobre lo difícil que resulta para el juzgador cuantificar su valoración

36. M. O. Finkelstein y W. B. Fairley, "A Bayesian Approach to Identification Evidence", *Harvard Law Rev.*, 83, 1970. Sobre el bayesianismo pueden verse además los trabajos del simposio "Decision and Inference in Litigation", *Cardozo Law Rev.*, 13, 1991.

subjetiva previa (o sea la *prior probability*); o sobre el riesgo que entraña poner en manos de los jueces instrumentos estadísticos incomprensibles para ellos. Además, la fórmula bayesiana supone un cálculo matemático relativamente simple en el caso-base de tener que valorar *una sola prueba* que versa *directamente* sobre la hipótesis a probar, que es además una *hipótesis simple*. Pero la complicación del cálculo aumenta de manera impresionante cuando se usa para resolver situaciones más complejas, como la *pluralidad* de elementos de prueba relativos a una hipótesis, la *cascaed inference* (o prueba mediata) o la prueba de una hipótesis referente a un hecho complejo. Otras críticas, en cambio, *ponen en cuestión la validez epistemológica del modelo en el campo judicial*, en la medida en que conduciría a resultados contraintuitivos y resultaría dudosamente compatible con principios fundamentales del ordenamiento, como la presunción de inocencia³⁷. No cabe duda de que las críticas aludidas deben ser tomadas en consideración. Pero tampoco cabe duda de que dichas críticas están a veces movidas por la tradicional y difusa aversión de los juristas en general y de los jueces en particular a abrirse a conocimientos extrajurídicos. El bayesianismo, al menos en el ámbito de la valoración de las pruebas científicas, es sin duda un camino que debe ser explorado³⁸.

37. El blanco de las críticas lo constituye aquí el papel desempeñado por la *probabilidad previa* del hecho que se quiere probar. Si el valor atribuido a la misma fuera bajo, también sería bajo el resultado de la valoración final, con independencia de que la probabilidad indicada por el nuevo material probatorio fuese alta. Pero además, y conectado con ello, si la formulación de la *probabilidad previa* estuviera viciada de un error inicial, podría atentar contra principios básicos del ordenamiento, como la presunción de inocencia. En suma, el uso del teorema de Bayes produce una sobrevaloración de la probabilidad inicial y una minusvaloración del peso que tienen las nuevas pruebas en el cálculo de la probabilidad final. Para las deficiencias del bayesianismo vid. L. H. Tribe, "Trial by Mathematics", *Harvard Law Rev.*, 84, 1971; M. Taruffo, *La prueba de los hechos*, Madrid, Trotta, 2004; y H. L. Cohen, *The Probable and the Provable*, Oxford, Clarendon Press, 1977. Me remito también, en todo caso, a la atinada y detenida crítica de L. Laudan al bayesianismo, en su trabajo de este mismo volumen.

38. Pero no es el Teorema de Bayes el único instrumento matemático-estadístico que se ha ensayado como modelo de valoración. Las *redes bayesianas* son también otra poderosa herramienta al servicio de la inferencia probabilística en la ciencia forense en la que concurren múltiples evidencias, científicas o no, bajo las hipótesis defendidas por las partes que intervienen en el proceso. Sobre las redes bayesianas cfr. F. Taroni, C. Aitken, P. Garbolino, A. Biedermann, *Bayesian Networks and Probabilistic Inference in Forensic Science*, Wiley, 2006.

