

CARMEN VÁZQUEZ
(ED.)

**ESTÁNDARES DE PRUEBA
Y PRUEBA CIENTÍFICA**
Ensayos de epistemología jurídica

Marcial Pons

MADRID | BARCELONA | BUENOS AIRES | SÃO PAULO
2013

- 2010: «On Method: The Ascent of Comparative Criminal Justice», London: Sage.
- 2011: «Does Article 6 of the European Convention on Human Rights Require Reasoned Verdicts in Criminal Trials?», *Human Rights Law Review*, 11.
- 2012: «Renegotiating Forensic Cultures: Between Law, Science and Criminal Justice», *Studies in History and Philosophy of Science*, 43, en prensa.
- ROBERTS, P., y WILLMORE, Ch., 1993: *The Role of Forensic Science Evidence in Criminal Proceedings*, RCCJ Research Study No. 11, HMSO.
- ROBERTS, P., y ZUCKERMAN, A., 2010: *Criminal Evidence*, 2.ª ed., Oxford: Oxford University Press.
- ROZENBERG, J., 2003: «Pathologist misled appeal judges in Sally Clark case», *Daily Telegraph*, 12 de abril.
- SANDERS, J., 2007: «Expert Witness Ethics», *Fordham Law Review*, 76.
- SHUMAN, D. W.; CHAMPAGNE, A., y WITHAKER, E., 1996: «Juror Assessments of the Believability of Expert Witness: A Literature Review», *Jurimetrics Journal*, 36.
- SMITH, J. C., 1998: «Is Ignorance Bliss? Could Jury Trial Survive Investigation?», *Medicine, Science and the Law*, 38.
- STEIN, A., 2005: *Foundations of Evidence Law*, Oxford: Oxford University Press.
- TANTON, R. L., 1979: «Jury Preconceptions and their Effect on Expert Scientific Testimony», *Journal of Forensic Sciences*, 24.
- THOMPSON, W. C., 1993: «Evaluating the Admissibility of New Genetic Identification Tests: Lessons From the "DNA War"», *Journal of Criminal Law and Criminology*, 84.
- THORNTON, P., 2004: «Trial by Jury: 50 Years of Change», *Criminal Law Review*.
- TWINING, W., 1982: «Taking Facts Seriously», en GOLD, N. (ed.), *Essays on Legal Education*, Toronto: Butterworths, reimpresso como *Rethinking Evidence*, 2.ª ed., Cambridge: Cambridge University Press, 2006.
- 2007: «Taking Facts Seriously - Again», en ROBERTS, P., y REDMAYNE, M. (eds.), *Innovations in Evidence and Proof*, London: Hart.
- VAN KAMPEN, P., y NIJBOER, H., 1997: «Daubert in the Lowlands», *UC Davis Law Review*, 30.
- VOGLER, R., 2005: *A World View of Criminal Justice*, London: Ashgate.
- VOGLER, R., y HUBER, B. (eds.), 2008: *Criminal Procedure in Europe*, Berlin: Max Planck Institut/Duncker and Humblot.
- WALKER, C., y STOCKDALE, R., 1999: «Forensic Evidence», en WALKER, C., y STARMER, K. (eds.), *Miscarriages of Justice: A Review of Justice in Error*, Oxford: Blackstone.
- WARD, T., 2004: «Experts, Juries and Withc-hunts: From Fitzjames Stephen to Angela Cannings», *Journal of Law and Society*, 31.
- WISTRICH, A. J.; GUTHRIE, Ch., y RACHLINSKI, J. J., 2005: «Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding», *University of Pennsylvania Law Review*, 153.
- WONDER, A. K. Y., 1989: «Science and Law, A Marriage of Opposites», *Journal of the Forensic Science Society*, 29.
- ZUCKERMAN, A. S., 1986: «Law, Fact or Justice?», *Boston University Law Review*, 66.

VII

PRUEBA CIENTÍFICA. UN MAPA DE RETOS

Marina GASCÓN ABELLÁN
Universidad de Castilla-La Mancha.

1. PRUEBA CIENTÍFICA. SOBREVALORACIONES Y PARADIGMAS

1.1. El mito

En los últimos años los constantes avances científicos y técnicos han tenido un profundo impacto en el ámbito de la prueba y juegan un papel cada vez más importante en todos los procesos. Los avances han sido particularmente espectaculares en el campo de la genética forense, que ha marcado un antes y un después en la resolución de numerosos problemas judiciales, como la investigación biológica de la paternidad, la identificación de personas o la investigación de indicios, es decir, el análisis de muestras biológicas de interés criminal, como manchas de sangre, saliva, espermatozoides o pelos¹. El potencial de la huella genética es de tal magnitud que su uso en los tribunales se ha convertido ya en moneda corriente. Pero la prueba de ADN no es la única protagonista de este *boom*. Hay también otras pruebas científicas que juegan un papel decisivo en muchas causas judiciales, como la dactiloscopia, la balística, el análisis de marcas o la prueba de locutores de voz. Esta última, por ejemplo, ha permitido fundar condenas en causas de terrorismo donde la prueba principal frente al

¹ Sobre éstas y otras aplicaciones de la prueba de polimorfismos ADN a la genética forense, véase MARTÍNEZ JARRETA, 1999: 129 y ss.

imputado es una llamada de teléfono avisando de la colocación de un artefacto explosivo. Hoy puede decirse, en suma, que las pruebas científicas se han convertido en la clave para probar muchos hechos que de otro modo resultarían difíciles de probar.

Pero la importancia de las pruebas científicas en la práctica procesal no ha ido acompañada de un proceso paralelo de cautelas y controles en relación con las mismas. Más bien ha sucedido lo contrario. Precisamente por el hecho de presentarse como «científicas» (y porque la mayoría de las veces —al menos en Europa— provienen de los laboratorios oficiales de la policía científica) estas pruebas han ido acompañadas de un aura de infalibilidad que ha frenado (cuando no claramente impedido) cualquier intento de revisión o reflexión crítica sobre las mismas, con el resultado de que su validez y valor probatorio se suelen asumir como dogmas de fe². Esta creencia ciega en la validez y valor de las pruebas científicas es algo que llama particularmente la atención si se considera que nada es menos «científico» que asumir como válido un conocimiento sin un previo control de sus postulados ajustado a una metodología científica. Pero sobre todo entraña un peligro, pues propicia que las decisiones probatorias apoyadas en este tipo de pruebas se asuman como incuestionables o irrefutables y, de paso, descarga al juez de hacer un especial esfuerzo por fundar racionalmente la decisión: basta con alegar que hubo prueba científica y que ésta apuntaba justamente en la dirección de la decisión probatoria final.

La beatificación de este universo probatorio se asienta en dos sobrevaloraciones. Por un lado sus resultados se aceptan como infalibles. Por otro, y sobre todo, se considera que esos resultados dicen cosas distintas de las que en realidad dicen³. La primera es una *sobrevaloración epistémica*. La segunda es, por así decirlo, una *sobrevaloración semántica*⁴.

² Por lo demás, esa desbordante confianza en la infalibilidad de la prueba científica se ha visto alimentada por el tremendo impacto que en el imaginario popular han tenido algunos *booms* televisivos como la norteamericana *CSI (Crime Scene Investigation)*, que han generado una especie de beatificación de estas pruebas. De modo que hoy estamos gozosamente dispuestos a creer en los informes provenientes de los laboratorios de la policía científica como si se tratase de una verdad revelada (advierte contra esta actitud IGARTÚA, 2007).

³ La sobrevaloración de la prueba científica ha sido denunciada en el National Research Council Report, 2009, *Strengthening Forensic Science in the United States. A Path Forward*, donde se lee expresamente: «Durante décadas las disciplinas de la ciencia forense han producido pruebas valiosas que han contribuido a la imputación y condena de delinquentes y a la exoneración de personas inocentes [...]. Sin embargo estos avances también han revelado que, en algunos casos, información sustantiva y testimonios basados en análisis defectuosos de la ciencia forense pueden haber contribuido a la condena errónea de personas inocentes. Este hecho ha demostrado el potencial peligro de atribuir un peso indebido a la información y testimonios derivados de pruebas y análisis defectuosos. Es más, los informes periciales imprecisos o exagerados han contribuido algunas veces a la admisión de pruebas erróneas o engañosas» (Summary, Introduction, Paragraph: *Challenges Facing the Forensic Science Community*, p. 4).

⁴ Sobre estas dos sobrevaloraciones, véase más extensamente GASCÓN, 2010: 81 y ss.

1.2. Sobrevaloración epistémica

Pero cabe preguntar, ¿por qué se asume con este fervor casi dogmático el mito de la infalibilidad de las pruebas científicas? ¿Por qué todos los esfuerzos que en los últimos tiempos se están realizando por introducir racionalidad en el ámbito de la prueba se dirigen fundamental y casi exclusivamente a las pruebas no científicas? ¿Por qué la prueba científica no se sitúa también bajo esa mirada crítica?

No voy a desarrollar aquí una respuesta detallada a estas cuestiones pero sí puede señalarse someramente su causa principal. No se ha reflexionado porque se asume que una prueba científica transcurre por derroteros epistemológicos distintos al resto de las pruebas; que la arquitectura inferencial de ambos tipos de prueba es distinta; que mientras una prueba no científica se estructura a través de un razonamiento *inductivo*, la prueba científica se articula mediante un razonamiento *deductivo*. Lo que se sostiene, en otras palabras, es que la prueba *no científica* está basada en leyes probabilísticas de débil fundamento epistémico, por lo común máximas de experiencia y leyes del actuar humano habitual, que además son aplicadas dentro de una metodología no científica, de modo que sus resultados son falibles y han de medirse siempre en términos de simple probabilidad. En cambio la prueba *científica* —se argumenta— está basada en leyes universales (o en todo caso en leyes probabilísticas que gozan de un fuerte fundamento epistémico), leyes que además son aplicadas dentro de una rigurosa metodología científica, por lo que sus resultados pueden tenerse por incuestionables o fuera de toda duda. Se piensa, en definitiva, que el conocimiento que se obtiene en las salas de los tribunales es frágil, en cambio lo que sucede en los laboratorios de la policía científica es otra cosa.

Naturalmente esta convicción es *errónea*, pues las pruebas científicas no constituyen por lo general un razonamiento de tipo deductivo, sino que están basadas prevalentemente en leyes estadísticas y sus resultados han de ser aún interpretados a la luz de otros datos, y por consiguiente difícilmente puede hablarse de «objetividad»⁵ y mucho menos de infalibilidad en relación con las conclusiones obtenidas a raíz de las mismas. Por supuesto es muy probable que el mito de la infalibilidad de las pruebas científicas tenga mucho que ver

⁵ «Cualquier juicio sobre probabilidad en un caso particular, incluso aunque el juicio esté basado en una frecuencia relativa, tiene una componente basada en conocimiento personal». Esto es lo mismo que decir que cualquier juicio sobre probabilidad es esencialmente personal y, por tanto, subjetivo. Los que entienden los resultados estadísticos como resultados objetivos en el sentido de interpretarlos como incontrovertibles y universalmente alcanzables tienen una concepción irreal de lo que la ciencia puede llegar a hacer en la práctica. Puede hablarse de objetividad entendida como acuerdo intersubjetivo. En este sentido es más fácil que los científicos acepten unos resultados si están basados en frecuencias relativas que si lo están en valoraciones subjetivas de probabilidad, pero ese acuerdo no implica, en ningún caso, que los científicos crean que los resultados son incontrovertibles, véase TARONI, AITKEN, GARBOLINO y BIEDERMANN, 2006: 21.

con la prueba del ADN, cuya metodología está ya tan perfeccionada y el grado de probabilidad que arroja es tan alto que en la práctica puede «actuarse como si» fuese infalible. No obstante, aún dando esto por sentado, conviene adoptar una actitud un poco más crítica, pues la calidad epistémica —y por tanto el valor probatorio— de los resultados de una prueba científica depende de varios factores que es necesario ponderar⁶.

En primer lugar depende de la *validez científica y/o metodológica* de la misma. Muchas de estas pruebas pueden realizarse por métodos científicos diferentes y no todos ellos gozan del mismo crédito en la comunidad científica correspondiente, de manera que la validez científica del método usado, y con ello la calidad de los resultados alcanzados, pudiera ser objeto de discusión. En segundo lugar, la fiabilidad atribuible a una prueba científica depende también de su *calidad técnica*. Cabe hablar a este respecto de *corrección técnico-procedimental*, en referencia a todo el proceso que conduce desde el descubrimiento o registro del vestigio o de la muestra hasta su análisis en el laboratorio (por ejemplo, a efectos de atribuir fiabilidad al resultado del análisis de una huella dactilar, el problema no es tanto, o no sólo, la validez científica de la prueba, sino, en primer lugar, saber quién tomó la huella, por orden de quién, en qué objeto estaba depositada, en qué punto concreto, cómo fue la cadena de custodia, etc.; y lo mismo con respecto al análisis de una mancha de sangre, orina, saliva)⁷. Pero cabe hablar también de *corrección técnico-científica*, en referencia a su correcta realización en laboratorio: por personal cualificado y siguiendo los protocolos apropiados. La regla aquí debería ser: «cuanto mayor es la expectativa de valor probatorio depositada en una prueba, más rigurosos deben ser los controles de realización de la misma». Y por último, en tercer lugar, no hay que olvidar que las pruebas las realizan personas de carne y hueso, y que por tanto pueden cometer errores. No son pocos los estudios que en los últimos tiempos llaman la atención sobre los *riesgos cognitivos* de algunas pruebas científicas⁸, sobre todo de aquellas tradicionales que, como la dactiloscopia y la grafística, tienen un fuerte componente comparativo que las deja enteramente bajo la supervisión del perito⁹.

⁶ En su informe, CHAMPOD y VUILLE (2010) realizan un exhaustivo repaso por los factores que pueden influir negativamente en el valor probatorio atribuible a los resultados de las pruebas científicas. A continuación destacamos algunos, pero para una visión completa remitimos a ese excelente estudio (que puede consultarse en www.coe.int/t/e/legal_affairs, pp. 10 y ss.).

⁷ Es evidente, por ejemplo, que, pese al potencial de la prueba de ADN, el estado en el que llegan los vestigios biológicos al laboratorio es crucial: si los vestigios no han sido bien recogidos o conservados (por ejemplo, porque han sido contaminados por un ADN extraño) la posibilidad y el rendimiento del análisis se reduce. Por eso la *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.

⁸ Con respecto a los riesgos cognitivos cfr. DROR, 2009: 93-110; DROR y COLE, 2010: 161-167.

⁹ CHAMPOD y VUILLE, 2010: 10-11.

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

Del mito de la infalibilidad (o lo que es lo mismo, de la falta de reflexión seria sobre el estatuto epistémico de las pruebas científicas) derivan algunas consecuencias de importancia. Seguramente la más evidente y adversa sea la falta de control sobre su validez y fiabilidad, lo que sin duda permite la entrada en el proceso de auténtica *ciencia basura*, datos sin fundamento científico alguno con los que los laboratorios y los peritos hacen un auténtico negocio. Piénsese sin ir más lejos en la grafología o en el polígrafo. Por otra parte, y vinculado a lo anterior, el mito de la infalibilidad de la prueba científica nos enfrenta al peligro del desconocimiento de los errores judiciales que pueden cometerse con base en ella. El denominado *Innocent Project*, puesto en marcha por los abogados Barry 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 condenas¹⁰. Por último, este mito entraña un riesgo adicional, un efecto adverso desde el punto de vista jurídico: en la medida en que el informe pericial se asume como infalible podría decirse que *es el propio perito quien indica al juez lo que debe creer sobre la hipótesis en consideración*, sin que el juez, por consiguiente, pueda separarse de tal juicio sin razones poderosas. Pero de este modo se termina convirtiendo a los peritos en decisores de la causa y, por consiguiente, instaurando un nuevo sistema de prueba fundado en la autoridad de los expertos.

1.3. Sobrevaloración semántica

Pero no es sólo el dogma de la infalibilidad de las pruebas científicas lo que debe ser revisado. También debe ser revisada otra cuestión: la que al prin-

¹⁰ Puede consultarse la página web del proyecto en www.innocenceproject.org.

cipio llamé la «sobrevaloración semántica»; más concretamente, el hecho de considerar que los resultados de la prueba (además de infalibles) dicen cosas distintas de las que en realidad dicen.

Si la prueba científica ha sido entronizada en el proceso es porque se da por descontado no sólo el altísimo valor probatorio o incluso infalibilidad de sus resultados, sino también que éstos hablan directamente de aquello que se pretende probar. Se piensa, en concreto, que el resultado de una prueba de ADN señala directamente la pertenencia o no del vestigio analizado a la persona de la que procede la otra muestra de ADN con el que aquél se contrasta (el acusado en una causa penal, por ejemplo); que el resultado de una prueba de balística dice directamente si el casquillo evaluado salió o no de la pistola del acusado; si la impresión de calzado analizada procede de la pisada de una cierta persona; si la escritura manuscrita examinada procede del demandado; si la voz analizada que realizó la llamada telefónica avisando de la colocación del artefacto explosivo pertenece o no al acusado, etc. Se piensa, en definitiva, que el resultado de una prueba científica habla en los términos en que el juez necesita pronunciarse.

Esta creencia expresa lo que en el ámbito de la ciencia forense se denomina el *paradigma de la individualización*, que se asienta sobre la supuesta capacidad de llegar a identificar plenamente a un individuo o a un objeto a partir de vestigios. Dicho paradigma se presenta como modelo a seguir para todas las técnicas identificativas en criminalística, y en términos generales consiste en sostener que el resultado de la prueba científica identifica (y además, categóricamente) un vestigio con una fuente, con exclusión de todas las demás¹¹.

Las cosas, sin embargo, no son así¹². Y de hecho en los últimos años el paradigma de la individualización ha recibido fuertes críticas por parte de la comunidad científica, que considera que, por más que se trate de un esquema

¹¹ Aludiendo al paradigma de la individualización, el *NRC Report, Strengthening Forensic Science in the United States. A Path Forward*, cit., 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» (cap. I, Introduction, Section *Pressures on the Forensic Science System*, Paragraph *Questionable or Questioned Science*: 43-44).

¹² Uno de los ensayos más estimulantes de los últimos tiempos sobre el uso de la estadística en la ciencia forense, es la obra *Statistical Evidence* (ROYALL, 1997: xi) que denuncia explícitamente que el uso de métodos estadísticos estándar conduce muchas veces a tergiversar los resultados de las pruebas: unas veces dándoles un peso mayor o menor del que realmente tienen; otras, considerando que *los datos estadísticos apoyan un resultado cuando en realidad sucede justo lo contrario*. Y esto último sucede, no porque los expertos usen equivocadamente la estadística, sino justamente porque domina el paradigma de la individualización.

de análisis simple e intuitivo, no resulta aceptable: la pretensión no comprobada de que es posible vincular un vestigio desconocido a una única fuente representa una equivocada intuición probabilista que iguala infrecuencia con unicidad¹³. O dicho muy brevemente, se considera imposible obtener prueba concluyente de la unicidad¹⁴.

Como resultado de lo anterior los científicos consideran que hay que interpretar los resultados de las pruebas, no en términos de identificación (de un vestigio con una fuente) sino en términos de *razón de verosimilitud* (*likelihood ratio*). Lo que se sostiene, más concretamente, es que no hay fundamento científico alguno para que un perito sostenga *categóricamente* que ha sido capaz de identificar a una persona o un objeto a partir de los análisis llevados a cabo en el laboratorio. Pero la crítica va más allá: no hay tampoco fundamento científico alguno para que en el informe pericial se realice siquiera una valoración *probabilística* sobre la posibilidad de atribuir a una persona o a un objeto el vestigio analizado¹⁵. El resultado de un análisis de voz, por ejemplo, o el de una comparación de perfiles de ADN, no dice (ni conclusiva ni probablemente) que la voz, o el ADN del vestigio analizado, 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 se observe tal rasgo en la voz analizada si ésta pertenece al acusado que si no pertenece a él»; o «es X veces más probable que coincidan los perfiles genéticos si el vestigio analizado procede del acusado que si procede de una fuente distinta». En otras palabras, las pruebas científicas hablan de la *probabilidad de los datos analíticos y técnicos* resultantes tras el análisis en el laboratorio a la luz de las hipótesis judiciales examinadas, y no al revés; es decir no hablan de la probabilidad de las hipótesis judiciales consideradas a la luz de esos datos. Por eso la extendida costumbre de expresar los resultados de las pruebas científicas en términos de identificación categórica o incluso probabilística (particularmente extendida en las técnicas forenses tradicionales) debe ser abandonada. Por lo demás, como ha subrayado CHAMPOD en su dura crítica al paradigma de la individualización, la ciencia forense, para ser útil, no necesita manejar ese paradigma, por lo que debe ser erradicado de la comunidad forense¹⁶.

¹³ SAKS y KOEHLER, 2008: 199-219; KAYE, 2009: n. 39, véase en línea: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1261970. Como KAYE comenta (n. 39), existe un riesgo nunca igual a cero al aceptar cualquier inferencia sobre un parámetro poblacional. «La distancia entre la muestra y la población requerirá siempre un salto de fe. Lo único que vale la pena debatir es la longitud del salto». Cfr. también STONEY, 1991.

¹⁴ Por eso, «un científico o un juez objetivo que revisara el estado actual de las ciencias forenses tradicionales probablemente calificaría las afirmaciones hechas por éstas como plausibles, carentes de suficiente investigación y sobrevaloradas», SAKS, KOEHLER, 2005: 892.

¹⁵ Para un análisis detallado de las críticas recibidas por el paradigma de la individualización, remito al trabajo conjunto GASCÓN, LUCENA y GOZÁLEZ, 2010.

¹⁶ CHAMPOD, 2009.

2. EL PARADIGMA DE LA VEROSIMILITUD

Pero si el paradigma de la individualización no resulta válido, entonces cabe preguntar, ¿cómo deben interpretarse los datos aportados por una prueba científica?, ¿cuál es el valor que hay que dar a los resultados de este tipo de pruebas? El llamado por ROYALL *paradigma de la verosimilitud*¹⁷ ofrece una respuesta a esta cuestión. Dicho paradigma se erige sobre la distinción entre las tres preguntas básicas que cabe formular una vez que el perito ha realizado ya los análisis pertinentes en una prueba científica: ¿*qué nos dicen los datos* u observaciones resultantes de esos análisis sobre la hipótesis A en relación con la hipótesis B?, ¿*qué debemos creer a partir de esos datos?* y ¿*qué debemos hacer?*¹⁸. Como 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. De hecho permite realizar un deslinde neto entre las tareas del perito y las del juez.

2.1. Lo que dicen los datos

Responder a la primera pregunta (*qué dicen los datos*) es la principal tarea del perito en el proceso, quien (por ejemplo en una prueba de comparación de voces) deberá *interpretar el resultado* de la prueba en el laboratorio (supongamos que positivo: las voces comparadas coinciden en el rasgo-R) dándole un alcance en relación con las hipótesis enfrentadas (supongamos que hipótesis-A: la voz pertenece al acusado; o hipótesis-B: la voz no pertenece al acusado). Muy simplemente, se trata de interpretar qué nos dicen los datos resultantes de la prueba sobre una hipótesis en relación con la otra (por ejemplo en estos términos: «es X veces más probable que se observe el rasgo-R en la voz analizada si ésta pertenece al acusado que si no pertenece a él»). Insistamos en que la interpretación de esos resultados ha de hacerse en términos de *razón de verosimilitud*, y es importante que se haga bien, porque es lo que se refleja en el informe de conclusiones y por consiguiente lo que se va a comunicar al juez.

2.2. Lo que debe creerse

Una vez que los datos han sido interpretados y expresados adecuadamente en el informe pericial hay que responder a la segunda pregunta (*qué debe creerse a partir de esos datos*), es decir hay que *evaluar* la veracidad de las

¹⁷ ROYALL, 1997.

¹⁸ Para un desarrollo más extenso del alcance de este paradigma véase GASCÓN, LICENA y GÓZALEZ, 2010.

hipótesis enfrentadas (A y B), y ello ha de hacerse a partir de lo que dicen esos datos pero teniendo en cuenta también lo que dice el resto de las pruebas e informaciones disponibles.

Es evidente que esta tarea de valoración de las hipótesis en juego a partir de los datos aportados por la prueba corresponde al juez, al menos por dos razones. Primero porque es el juez quien tiene *instruccionamente* atribuida la función de proveer una solución jurídica para un caso controvertido o litigioso, y por tanto la de determinar previamente cuáles han sido los hechos que han dado origen al conflicto. Y segundo porque también el principio de *libre convicción*, en la medida en que consagra que el juez no debe dar por probado un hecho entretanto no se haya convencido de su verdad, exige atribuir a éste la tarea de determinar lo que hay que creer a partir de la prueba, incluida la científica¹⁹. Resulta claro, por consiguiente, que el paradigma de la individualización no es compatible con el hecho de que sea el juez quien tiene asignada en exclusiva la tarea de valorar la prueba, pues si el informe pericial hablara en términos de «el vestigio *v* proviene (o no proviene) de la fuente *f*» habría sido el perito (y no el juez) quien habría determinado lo que hay que creer sobre las hipótesis en conflicto.

En suma, la neta distinción entre la tarea del perito (interpretar y comunicar lo que dicen los datos) y la del juez (valorarlos a la luz de los demás datos y pruebas disponibles) es el núcleo del paradigma de la verosimilitud. El paradigma de la individualización, por el contrario, no distingue entre la tarea del perito y la del juez, pues propicia que el perito declare en términos de «el vestigio *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 corresponde²⁰. Pero de este modo importantes principios jurídicos, como el de la libre convicción, resultan conculcados.

2.3. Lo que debe hacerse

Por último, el paradigma de la verosimilitud también traza la distinción entre *lo que debe creerse* y *lo que debe hacerse* a partir de la realización de una prueba científica. Dicha distinción resulta sumamente apropiada para reconstruir otra diferenciación esencial dentro del universo de la prueba: la

¹⁹ Por eso la «cientificidad» de la prueba, por sí misma, no cierra la cuestión de su valor probatorio, que debe ser resuelto por el tribunal en cada caso en virtud del principio de libre valoración. En el mismo sentido WRÓBLEWSKI (1969: 185): «La referencia a la ciencia especializada no excluye la valoración de las pruebas».

²⁰ Por lo demás, este paradigma comporta un riesgo importante de comisión de falacias. Como la conocida falacia del fiscal (*prosecutor fallacy*). Sobre la falacia del fiscal y la falacia de la defensa, cfr. THOMPSON Y SCHUMANN, 1987: 167. Sobre la falacia del fiscal puede consultarse más recientemente, LEUNG, 2002: 44-50.

existente entre la valoración de la prueba y la decisión sobre los hechos probados. Mientras que lo que debe *creerse* fundadamente sobre una hipótesis se expresa como la probabilidad de esa hipótesis y puede identificarse con la *valoración* de la prueba, lo que debe *hacerse* (o sea, la toma de decisiones en presencia de esa probabilidad) expresa la cuestión de los *estándares* de prueba; es decir, la determinación del grado de probabilidad que debe haber alcanzado una determinada hipótesis para que podamos considerarla probada y actuar en consecuencia. Y así como la valoración de la prueba es (o debería ser) una cuestión completamente librada a la *racionalidad epistémica*, la fijación de estándares de prueba es también una cuestión de *policy*, es decir política o valorativa, dependiente de lo tolerante que el sistema esté dispuesto a ser con cada uno de los dos errores que pueden cometerse al adoptar una decisión: declarar probado lo falso y declarar no probado lo verdadero. Es evidente, por ejemplo, que el estándar de prueba que suele exigirse en el proceso civil es mucho más bajo que el que se exige en el penal. De manera que mientras que en el primero podríamos aceptar que la prueba de un hecho descansase sobre una probabilidad preponderante (digamos, del 55 por 100), en el segundo sólo estamos dispuestos a condenar con una probabilidad muy cualificada; y ello por la simple razón de que somos menos tolerantes con el error consistente en condenar a un inocente que con el consistente en absolver a un culpable.

En suma, *lo que hay que creer* sobre una cierta hipótesis depende de lo que dicen los datos sobre la misma y del resto de las informaciones disponibles. *Lo que debe hacerse* depende obviamente de lo que hay que creer, pero también del contexto normativo en el que se encuadra la decisión a adoptar. Por eso, ante una misma creencia, no es lo mismo tomar una decisión en el marco de un proceso civil que en el marco de un proceso penal. Y aún en este último supuesto, no es lo mismo que la decisión se adopte en el marco de la instrucción que en el del juicio oral.

3. LOS RETOS

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

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

La admisibilidad es el control de entrada de la prueba en el proceso. Controlar que los informes periciales admitidos al proceso tengan un sólido fundamento científico es de un interés epistémico máximo, pues dicho control aspira a dejar a la pseudociencia fuera del ámbito de las decisiones judiciales. De hecho todos los esquemas de control de validez y fiabilidad (en definitiva, de calidad) de las pruebas científicas comparan ese objetivo. Cómo conseguir esto es lo que marca la diferencia entre ellos.

En línea de principio cabe adoptar dos grandes posturas sobre la admisibilidad de la prueba científica. Cabe sostener, por un lado, que *no es necesario un control específico de admisibilidad científica de la prueba* (validez científica) *junto con el control de admisibilidad procesal* (relevancia y legalidad) y que, por consiguiente, todas las pruebas científicas relevantes y legalmente obtenidas deben ser admitidas, dejando la cuestión de su validez científica para el momento de la valoración²¹. Pero también cabe sostener que *junto al control de admisibilidad procesal debe haber otro de admisibilidad científica* diferente de la fase de valoración de la prueba, es decir de atribución a la misma de valor probatorio²².

En Europa en general se ha seguido la primera postura. Las pruebas científicas relevantes y legalmente válidas entran en el proceso sin traba alguna y son valoradas por el juez, que les atribuye mayor o menor fuerza probatoria dependiendo, entre otros factores, de la validez científica y fiabilidad de la metodología usada para obtenerlas. En otras palabras, la validez y fiabilidad de una prueba científica no se evalúa en el trámite de admisibilidad de la prueba sino que se pondera en el momento de la valoración y se refleja en el valor probatorio que se le atribuye a dicha prueba. Pero esta opción presenta algunos inconvenientes.

En primer lugar una importante cuestión conceptual que, no obstante, tiene trascendencia práctica. Este modelo no distingue entre la fase de admisibilidad científica de la prueba y la fase de valoración de la misma, de modo que la *fiabilidad científica* de la prueba se confunde a menudo con su *valor probatorio*. Debe observarse, sin embargo, que admitir una prueba científica no implica pronunciarse sobre su valor probatorio sino meramente sobre la posibilidad de tomarla en consideración para la decisión del caso. Ambos conceptos (validez científica y valor probatorio) deben ser claramente distinguidos.

²¹ Lo que se sostiene, en definitiva, es que, si la prueba no es válida, ya se encargará el juez en el momento de la valoración de privarla de valor probatorio.

²² Puede consultarse in extenso una confrontación de los modelos de control aludidos en HAVAN-NEH y AL-RAWASHDEH, 2010: 187 y ss.

En segundo término está también el problema de que los jueces, por lo general, carecen de los conocimientos científicos necesarios para evaluar la validez y fiabilidad de estas pruebas. Por eso dejar enteramente en sus manos la libre apreciación de este dato, sin establecer un estándar objetivo para guiarlos, puede conducir a rechazar pruebas científicamente válidas que podrían haber ayudado al esclarecimiento de la verdad en el proceso, o —lo que es más probable— puede conducir a dar crédito a pruebas que son pura charlatanería pseudocientífica y que pueden convertirse en fundamento de decisiones judiciales erróneas. No puede desconocerse la gran capacidad de persuasión e influencia que una prueba científica puede tener sobre la convicción judicial, con lo que si se admite su entrada en el proceso sin ningún control se corre el riesgo de que, al final, la decisión venga determinada por pura y simple *junk science*²³. Por eso parece más apropiada la opción alternativa, esto es, la que aboga por establecer un estándar objetivo de admisibilidad científica de las pruebas junto al examen de su admisibilidad procesal.

Ésta es la opción que se ha seguido en Estados Unidos. Allí, el que se conoce como test Daubert²⁴ ha establecido como estándar de admisibilidad, junto con la *aceptación general* de la teoría y técnica usadas por la comunidad científica concernida, la efectiva *validez científica* de la técnica aplicada, que ha de acreditarse evaluando los siguientes factores: *a)* si la teoría o la técnica se puede probar y si ha sido efectivamente *puesta a prueba* (verificabilidad) y refutabilidad como criterios mínimos de científicidad de cualquier teoría; *b)* si se han realizado publicaciones especializadas en el tema sometidas a una *revisión de pares*, y *c)* cuáles son los *márgenes de error* potencial asociados al método. No es que el test Daubert exija que cualquier técnica científica usada como fuente de prueba cumpla todos los criterios mencionados, pues esto es algo que queda al criterio del juez. Pero sí exige a los jueces —y esto es lo más significativo— 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*²⁵. Por decirlo muy sumariamente, la sentencia *Daubert*

²³ La admisión al proceso de prueba científica no fiable «confunde a los jueces, socavando la corrección fáctica de las decisiones en materia penal» (ROBERTS y ZUCKERMAN, 2004: 69).

²⁴ El estándar ha sido establecido en el caso *Daubert v. Merrill Dow Pharmaceutical Inc.*, 113 S. Ct. 2786 (1993). El caso que allí se abordaba era la admisibilidad de nueva o inusual prueba científica, y la cuestión legal que se debatía era si el tradicional *Frye test* vigente desde 1923, que cifraba la admisibilidad de la prueba en la «aceptación general» del método usado por parte de la comunidad científica, había sido sustituido por la *Federal Rule of Evidence 702*, de 1975, que establecía que el testimonio de los expertos resultará admisible cuando, aparte de relevante, este basado en datos científicos que sean el resultado de principios y métodos científicos válidos que hayan sido aplicados correctamente por el experto. La sentencia *Daubert* respondió afirmativamente a esa cuestión y estableció además —y éste es el corazón del test Daubert— que *son los tribunales* los que deben escrutar si el testimonio de los expertos cumple esos parámetros de relevancia y de efectiva «validez científica» de método usado (agradezco a Susan HAACK las atinadas precisiones que me hizo sobre el tema).

²⁵ Prácticamente una réplica del test Daubert cabe encontrar en los criterios sugeridos para el Reino Unido por la LAW COMMISSION, 2009: 53 y ss. (LAW COMMISSION, *The Admissibility of Expert Evidence in Criminal Proceedings in England and Wales. A New Approach to the Determination of*

supone un llamamiento a los jueces para que miren más críticamente las pruebas científicas²⁶. De hecho, después de esta sentencia, ámbitos tradicionales de la policía científica, como las huellas papilares, la grafística y la balística empiezan a ser puestos seriamente en cuestión.

Pero el test Daubert tampoco ha zanjado la polémica porque su aplicación presupone que los jueces son capaces de comprender y aplicar los conceptos científicos necesarios para evaluar la validez de estas pruebas: lo que acaso no siempre sea cierto. Por supuesto cuando la prueba ha sido ya objeto de un intenso debate y aceptación por parte de la comunidad científica, como sucede con la prueba del ADN, no se plantea ninguna cuestión sobre su admisibilidad. Pero siempre habrá otras especialidades de la ciencia forense con mayor fragilidad científica que la prueba de ADN. No debe extrañar por ello que para muchos cada vez esté más clara la necesidad de formar a los jueces en los distintos ámbitos de especialidad forense, y sobre todo la necesidad de proporcionarles una formación específica en las nuevas áreas que vayan surgiendo. En todo caso muchas cuestiones permanecen abiertas: ¿es realista instar a los jueces a adquirir conocimientos científicos para poder evaluar la fiabilidad de la prueba científica? ¿no acentuaría esto la disparidad de las actuaciones en práctica? ¿Pero no es aún más peligroso dejar esta tarea enteramente en manos de los

Evidentiary Reliability. A Consultation Paper, www.lawcom.gov.uk/keeper_evidence.htm, Paper 2009, 190, pp. 53 y ss.)

«Para determinar si la prueba pericial científica (o supuestamente científica) es lo suficientemente fiable como para ser admitida, el tribunal debe considerar los siguientes factores y cualesquiera otros que se consideren relevantes:

- a)* Si los principios, técnicas y presupuestos en los que se basa han sido adecuadamente sometidos a prueba y, de ser así, el grado en que los resultados de tales pruebas demuestran su solidez.
- b)* El margen de error asociado a la aplicación de esos principios, técnicas y presupuestos, así como a las conclusiones obtenidas a partir de los mismos.
- c)* Si hay un cuerpo de literatura especializada en ese campo de conocimiento.
- d)* En qué medida los principios, técnicas y presupuestos han sido examinados por otros científicos —por ejemplo en publicaciones sometidas a revisión por pares— y, de ser así, en qué medida se consideran sólidos en la comunidad científica.
- e)* La cualificación, experiencia y publicaciones relevantes del testigo experto y su prestigio en la comunidad científica.

f) La validez científica de los puntos de vista opuestos (si es que los hay) y la cualificación, experiencia y prestigio profesional en la comunidad científica de los científicos que los mantienen.

g) Si hay pruebas que sugieran que el testigo experto no ha actuado de conformidad con su primordial deber de imparcialidad.

²⁶ Con elocuentes palabras CHAMPD y VUILLE (2010: 26), la sentencia *Daubert* «insiste de manera implícita sobre el escepticismo que el juez debe mantener hacia el perito, el cual debía 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».

²⁷ Esta es la razón por la que, en su informe de 2009, la Law Commission propone para el Reino Unido la posibilidad de que excepcionalmente, en los casos muy delicados, el juez pueda solicitar la ayuda de los expertos para evaluar la validez científica de la técnica propuesta (me refiero al informe *The Admissibility of Expert Evidence in Criminal Proceedings in England and Wales. A New Approach to the Determination of Evidentiary Reliability*, 2009). Estimo esta sugerencia sumamente interesante.

expertos, que podrían tener intereses ocultos en la promoción de ciertas técnicas? ¿No sería posible un sistema donde los jueces decidieran con la ayuda de un órgano independiente y científicamente competente?²⁸

3.2. Sobre el modo de transmitir los resultados de la prueba científica

Aparte del central asunto de la fiabilidad de la prueba científica hay otra importante cuestión relativa al informe pericial y a la que desgraciadamente no se le ha prestado demasiada atención²⁹. La cuestión concierne, más precisamente, al modo en que los peritos deberían interpretar y expresar los resultados de los test en su informe para evitar que sean malentendidos o sobrevalorados por los jueces. Muy brevemente, cómo deben formularse las conclusiones del informe pericial para comunicar correctamente los resultados de las pruebas a los tribunales.

De acuerdo con el paradigma de la verosimilitud al perito no le corresponde emitir una opinión sobre la hipótesis en consideración sino sólo dar cuenta de los datos de un modo científicamente riguroso, permitiendo así que el juez entienda exactamente su significado y pueda valorarlos junto con el resto de las pruebas disponibles. En relación con esto conviene insistir en una cuestión central: las conclusiones de la prueba que el perito formula en su informe se orientan a determinar el significado (o el grado de probabilidad) de los datos analíticos o técnicos resultantes a la luz de todas las hipótesis en consideración, y no el significado (o el grado de probabilidad) de las hipótesis en consideración a la luz de los datos analíticos o técnicos resultantes. Entender que las conclusiones de la prueba hablan en estos últimos términos, esto es, que las pruebas le dicen al juez justamente lo que él debe determinar para fundar su decisión, es propio del paradigma de la individualización, que no distingue entre lo que dicen los datos (una tarea pericial) y lo que debe creerse a partir de los datos (una tarea judicial), y da lugar a múltiples malinterpretaciones y, en ocasiones, a serios errores judiciales. Esto, por ejemplo, es lo que explica que se produzcan las conocidas como falacia del fiscal y falacia de la defensa.³⁰

²⁸ «Una posibilidad podría ser establecer un órgano europeo, al estilo del British Forensic Science Advisory Council and Forensic Science Regulator, que actuara como el principal consultor de las autoridades políticas y jurídicas en relación con la fiabilidad de las técnicas científicas que se usen [...]». Ese órgano no tendría poderes vinculantes, pero podría emitir recomendaciones para los tribunales nacionales que auxiliarían a los jueces cuando tuviesen que pronunciarse sobre la admisibilidad de una nueva técnica forense o sobre la fiabilidad de una nueva forma de prueba, o cuando pareciera adecuado abandonar una forma de prueba que hubiera quedado obsoleta». CHAMPDÓN Y VUILLE, 2010: 109.

²⁹ De ello se lamentaba no hace mucho, entre otros, WILLIS, 2009: 523-545. El *NRC Report* también presta atención a este asunto, al establecer que es absolutamente necesario sentar las bases para la correcta valoración y comunicación de los hallazgos forenses («Reporting Results Section», p. 186).

³⁰ Sobre la falacia del fiscal y la falacia de la defensa, cfr. THOMPSON Y SCHUMANN, 1987: 167 y ss. Y, entre nosotros, CARACEDO, 1999: 303. Sobre la falacia del fiscal puede consultarse más recientemente LEUNG, 2002: 44-50.

Probablemente no hay fórmulas mágicas para expresar las conclusiones en el informe, y probablemente no hay una sola fórmula. Pero lo que sí está claro es que debe huirse de las expresiones (categóricas o incluso sólo probabilísticas) sobre las hipótesis planteadas por las partes con respecto a un delito criminalístico, y debe en cambio optarse por pronunciamientos del tipo «es R veces más probable que coincidan los perfiles genéticos si el vestigio x procede de la fuente f que si no procede de f »; o «es R veces más probable que coincidan las huellas dactilares si la huella es del acusado que si no lo es»³¹. Por eso las llamadas escalas verbales de probabilidad, ampliamente usadas en el análisis de huellas digitales, en el examen de escritura, en la balística, en la comparación de caras, en la identificación de voz mediante análisis lingüístico o en la comparación de huellas de calzado, de neumáticos y de herramientas son erróneas. De hecho, estas escalas reproducen el paradigma de la identificación estableciendo, no lo que dicen los datos, sino lo que debe creerse sobre la hipótesis en consideración a partir de los mismos. En este sentido un importante reto que la ciencia forense tiene ante sí es el establecimiento de unos estándares de conclusiones que guíen el trabajo de los laboratorios y que al propio tiempo aseguren el uso de criterios uniformes en este punto³².

Y una última observación al respecto. La estadística juega un papel crucial en la interpretación de los datos resultantes de las pruebas científicas. De hecho, los datos plasmados en las conclusiones del informe pericial que después han de ser valorados por los jueces son, en la mayoría de los casos, datos estadísticos. Ahora bien, está comprobado el peso desproporcionado (*overweighting thesis*) que el público en general y los jueces en particular atribuyen a la estadística³³. Precisamente por eso no parece descabellado plantear la necesidad de que los jueces reciban alguna formación en este sentido. Obviamente no se trata de convertir a los jueces en matemáticos, o en *amateur scientists*, pero sí de proporcionarles los conocimientos necesarios para conjurar el riesgo de malinterpretación o de sobrevaloración de los datos estadísticos en los que se expresa el resultado de una prueba científica.

De todos modos, y aunque no cabe dudar de los beneficios que reportaría un adiestramiento como el comentado, ello no anularía el importante papel auxiliar de los peritos en la tarea de interpretación de los datos del informe. Muy simplemente, no siempre resultará fácil para el juez, como para cual-

³¹ Me permito remitir aquí, de nuevo, al trabajo de GASCÓN, LUCENA y GÓZALEZ, 2010.

³² En relación con este punto cabe mencionar que el Laboratorio Sueco de Ciencia Forense (SKL) anunció en 2009, tras años de debate interno, su decisión de unificar el modo de redactar las conclusiones para así proporcionar la misma información a los tribunales independientemente del tipo de prueba de que se tratase.

³³ Por citar solo a Europa, son conocidos, entre otros, el caso de Sally Clark, en Inglaterra, o el de Lucia de Berk, en los Países Bajos, absuelta en abril de 2010. Ambos, y algunos más, son citados como ejemplo de lo dicho por CHAMPDÓN Y VUILLE, 2010: 10.

quiera que no sea un experto, aislar en dicho informe *qué cosa* dicen exactamente los datos y *qué valor* hay que atribuir a eso que dicen los datos. Por más claro y riguroso que sea el informe, puede ser inevitablemente complejo y objeto de malinterpretaciones. Por eso es necesario que, como regla general, se requiera la presencia del experto en el acto de juicio para ayudar a interpretar los datos del informe y evitar así que éstos sean malinterpretados por el juez. Esta presencia representa, en este sentido, una *garantía epistemológica*³⁴.

3.3. Sobre la valoración de la prueba

La valoración de la prueba es el núcleo de la decisión probatoria y consiste en determinar lo que hay que creer sobre la hipótesis en consideración. Valorar la prueba científica, por consiguiente, consiste en atribuirle valor probatorio en relación con la hipótesis que describe los hechos del caso. Son dos las cuestiones relevantes que se plantean en relación con la valoración de la prueba: *quién debe valorar y cómo llevar a cabo la valoración*.

¿Quién debe valorar? La respuesta a esta cuestión puede parecer evidente, porque así es.

Si es el juez quien tiene la tarea de proveer una solución normativa para el caso planteado, entonces es también a él a quien corresponde la tarea previa de determinar autoritadamente cuáles han sido los hechos que han dado origen al conflicto. Sin embargo no es ocioso insistir en ello, pues, precisamente en consonancia con el paradigma de la identificación, muchos tribunales actúan «como si» fuese el perito, y no el juez, quien ha de valorar la prueba. Y es que, en efecto, al expresar las conclusiones de la prueba conforme al paradigma de la identificación (es decir, al declarar en términos de, por ejemplo, «a la vista de los datos resultantes de los test, la hipótesis es probable en un grado x»,

«los datos resultantes de los test hacen la hipótesis probable en un grado x» o «son compatibles con la hipótesis en un grado x», o aún más directamente, «el vestigio v procede de la fuente f »), es el experto (y no el juez) quien habrá establecido lo que hay que creer sobre la hipótesis. En virtud del paradigma de la verosimilitud, sin embargo, es el juez quien ha de valorar la prueba, o sea, quien ha de decir lo que ha de creerse sobre la hipótesis que se pretende probar.

Y una última consideración relacionada con este aspecto. Suele decirse muchas veces que, en virtud del principio de libre convicción, «el juez es *peritus peritiorum*», expresión con la que se pretende destacar que es él quien tiene la última palabra sobre el valor que debe atribuirse a los datos expresados en el informe pericial y que, por consiguiente, *no está estrictamente vinculado* a dicho informe; lo cual —se dice también— resulta un tanto paradójico si se tiene en cuenta que él no es un experto y que por tanto carece de los conocimientos necesarios para formarse una opinión propia sobre la prueba realizada. La cuestión, sin embargo, está mal planteada, pues en realidad carece de sentido afirmar que «el juez es *peritus peritiorum*». Conforme al paradigma de la verosimilitud lo que el juez debe hacer no es, o bien *vincularse* al informe pericial incluso si está convencido de que debe apartarse del mismo (infringiendo entonces el principio de libre convicción) o bien, por el contrario, *apartarse* del mismo en su condición de *peritus peritiorum* (asumiendo entonces que posee conocimientos especializados que muy probablemente no posee). Lo que el juez debe hacer es otra cosa: *determinar lo que hay que creer sobre la hipótesis en consideración a la luz del informe pericial y (también) del resto de las pruebas e informaciones que obran en la causa*. El juez valora la prueba, no es perito de peritos.

¿Cómo valorar? La segunda y central cuestión es la de cómo realizar la valoración (o sea, cómo determinar lo que hay que creer sobre la hipótesis a probar) cuando se ha aportado prueba científica.

Cabe distinguir dos supuestos. Cuando únicamente existe prueba científica en apoyo de la hipótesis, lo que hay que creer sobre la misma viene enteramente determinado por (o es coincidente con) el valor probatorio atribuido a la prueba científica. Cuando, por el contrario, además de la prueba científica existe también prueba no científica, que es la situación más frecuente, determinar lo que hay que creer sobre la hipótesis a probar requiere ponderar el valor probatorio de los datos estadísticos aportados en el informe pericial con el valor probatorio atribuido al resto de las pruebas. Naturalmente el juez realiza normalmente esta valoración sin auxilio alguno de las matemáticas; «subjetivamente», por así decirlo. Sin embargo, dado que el valor probatorio de la prueba científica puede ser expresado en términos estadísticos, se han propuesto también instrumentos matemáticos para llevar a cabo la valoración.

³⁴ 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. Se trataría, por consiguiente, de extender también a las pruebas científicas provenientes de los laboratorios oficiales el mismo tratamiento que ya la reforma del art. 788 de la Ley de Enjuiciamiento Criminal española confirió a los informes forenses emitidos por laboratorios oficiales sobre la naturaleza, cantidad y pureza de las drogas. «Este tipo de propuestas me parecen profundamente desacertadas. Podrán servir —esto no se cuestiona— para agilizar el acto de juicio (puesto que no será necesario que declaren los peritos sino que el juez, por su cuenta, examinará el documento del informe pericial), así como para optimizar los recursos de los laboratorios oficiales y facilitar el trabajo de sus expertos, que no se verán obligados a acudir al tribunal. Pero, al menos en el proceso penal, lesionan valores fundamentales vinculados al derecho de prueba, como la publicidad, la contradicción, la oralidad y la inmediación, como lesionan también el que constituye el principio medular de la decisión probatoria: el de *libre valoración*» (GASCÓN Y LUCENA, 2010: 105-106).

El principal intento se ha realizado mediante la aplicación del teorema de Bayes³⁵ a las inferencias jurídicas basadas en probabilidades subjetivas. En su formulación más simple la fórmula bayesiana afirma que la probabilidad de un evento H, dado un evento E, puede determinarse en función de la frecuencia estadística con la que dado H se verifica E y de la probabilidad atribuida precedentemente al evento H.

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

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

P (E/H): grado de confianza de que ocurra E si H es verdadera.

P (E/no-H): grado de confianza de que ocurra E independientemente de que H sea verdadera.

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

P (H): grado de creencia de que H es verdadera antes de saber si E ocurriría (*prior probability*).

La utilidad de la fórmula bayesiana en el proceso reside en que permite combinar la información estadística aportada por una prueba científica (la procedente del informe pericial y expresada mediante una *likelihood ratio*) con el valor atribuido anteriormente por las otras pruebas a la hipótesis que se pretende probar (la *prior probability*), y el resultado de esa combinación es la *posterior probability*, que expresa el resultado de la valoración de la prueba. En realidad, como puede verse, la fórmula bayesiana mide el impacto que, sobre el hecho que se pretende probar, provoca la inclusión de nuevos datos de naturaleza estadística, como son los resultantes de las pruebas científicas³⁷.

Sin embargo las ventajas de utilizar el teorema de Bayes para mejorar la precisión y controlabilidad (en definitiva, la fiabilidad) de la valoración judicial sobre hechos se ven atenuadas por los problemas que presenta su uso en el proceso. De hecho el bayesianismo, como modelo general de valoración,

³⁵ El teorema de Bayes no es el único instrumento matemático que ha sido propuesto como instrumento de valoración. Las *redes bayesianas* también constituyen una poderosa herramienta al servicio de las inferencias probabilísticas en ciencia forense cuando concurren múltiples pruebas, científicas o no, bajo las hipótesis defendidas por las partes que intervienen en el proceso. Cf. TARONI, ALITKEV, GARROLINO y BIEDERMANN, 2006.

³⁶ La probabilidad subjetiva indica el grado de creencia personal de un *factfinder* racional sobre la ocurrencia de un suceso incierto.

³⁷ FINKELSTEIN y FAIRLEY, 1970. Sobre el bayesianismo pueden verse además los trabajos del simposio *Decision and Inference in Litigation*, *Cardozo Law Review*, 13, 1991.

ha sido objeto de numerosas críticas. Algunas de estas críticas están animadas por consideraciones *prácticas*. Por ejemplo, llaman la atención sobre lo difícil que resulta para el juzgador cuantificar su valoración subjetiva previa (o sea la *prior probability*); o sobre el riesgo que entraña poner en manos de los jueces instrumentos estadísticos incomprensibles para ellos. Además, la fórmula bayesiana supone un cálculo matemático relativamente simple en el caso-base de tener que valorar una *sola prueba* que versa *directamente* sobre la hipótesis a probar, que es además una *hipótesis simple*. Pero la complicación del cálculo aumenta de manera impresionante cuando se usa para resolver situaciones más complejas, como la *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 inocencia³⁸.

No cabe duda de que las críticas aludidas son serias y deben ser tomadas en consideración. Pero en cualquier caso los problemas del bayesianismo como modelo general de valoración de la prueba no anulan la utilidad de las herramientas estadísticas como modo de cuantificar (mediante una *likelihood ratio*) el valor probatorio de una prueba científica. Esto significa que en la valoración de la prueba acaso puedan coexistir dos distintos conceptos de probabilidad. De un lado la *probabilidad matemática o estadística*, que es usada para cuantificar el *valor probatorio* de la prueba científica y que se expresa mediante una *likelihood ratio*. De otro lado la *probabilidad lógica o inductiva*, que es usada para expresar la *convicción sobre la hipótesis* a probar a la luz de todas las pruebas e informaciones disponibles (científicas o no).

BIBLIOGRAFÍA

- CARRACEDO, A., 1999: «Valoración de la prueba del ADN», en MARTÍNEZ JARRETA, B. (dir.), *La prueba del ADN en la medicina forense. La genética al servicio de la ley en el análisis de indicios criminales y en la investigación biológica de la paternidad*, Barcelona: Masson.

³⁸ El blanco de las críticas lo constituye aquí el papel desempeñado por la *probabilidad previa* del hecho que se quiere probar. Si el valor atribuido a la misma fuera bajo, también sería bajo el resultado de la valoración final, con independencia de que la probabilidad indicada por el nuevo material probatorio fuese alta. Pero además, y conectado con ello, si la formulación de la *probabilidad previa* estuviera viada de un error inicial, podría atentar contra principios básicos del ordenamiento, como la presunción de inocencia. En suma, el uso del teorema de Bayes produce una sobrevaloración de la probabilidad inicial y una minusvaloración del peso que tienen las nuevas pruebas en el cálculo de la probabilidad final. Para las deficiencias del bayesianismo, véanse TRIJBE, 1971; TARCIFRO, 2004; COHEN, 1977. Me remito también, en todo caso, a la ahmada crítica de LAUDAN, 2005, a la aplicación del cálculo probabilístico en la construcción de un estándar de prueba.

- CHAMPD, C., 2009: «Interpretation of evidence and reporting in the light the 2009 NRC report». Keynote Speech of the Interpretation and Evaluation Session of the V Conference of the European Academy of Sciences held in Glasgow (Scotland), University of Strathclyde.
- CHAMPD, C., y VUILLE, J., 2010: *Scientific Evidence in Europe. Admissibility, Appraisal and Equality of Arms*, European Committee on Crime Problems, 10.
- COHEN, H. L., 1977: *The Probable and the Provable*, Oxford: Clarendon Press.
- DROR, I., 2009: «How can Francis Bacon help forensic science? The four idols of human biases», *Jurimetrics Journal*, 50.
- DROR, I. E., y COLE, S. A., 2010: «The vision in "blind" justice: Expert perception, judgment, and visual cognition in forensic pattern recognition», *Psychonomic Bulletin & Review*, 17.
- FINKELSTEIN, M. O., y FAIRLEY, W. B., 1970: «A Bayesian Approach to Identification Evidence», *Harvard Law Review*, 83.
- GASCÓN, M., 2010: «Prueba científica: mitos y paradigmas», en *Anales de la Cátedra de Francisco Suárez, número conmemorativo: «Un panorama de filosofía jurídica y política (50 años de Anales de la Cátedra Francisco Suárez)»*, Granada, 44.
- GASCÓN, M., y LUCENA, J. J., 2010: «Pruebas científicas. La necesidad de un cambio de paradigma», *Jueces para la Democracia*, 71 (11).
- GASCÓN, M., LUCENA, J. J., y GOZÁLEZ, J., 2010: «Razones científico-jurídicas para valorar la prueba científica: una argumentación multidisciplinaria», *Revista La Ley*, 7481 (4/10).
- HAYAINEN, A. M., y AL-RAWASHDEH, S. H., 2010: «Theoretical Approaches to Admitting Scientific Evidence in the Adversarial Legal System», *European Journal of Scientific Research*, 2 (41).
- IGARTÚA, J., 2007: «Prueba científica y decisión judicial: unas anotaciones propedéuticas», *Diario La Ley*, 6812 (2/11).
- KAYE, D. H., 2009: *Probability, Individualization and Uniqueness in Forensic Science: Listening to the Academies*, 30th June, Social Science Research Network (SSRN). Disponible en línea: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1261970.
- LAUDAN, L., 2005: «Por qué un estándar de prueba subjetivo y ambiguo no es un estándar», *Doxa*, 28.
- LAW COMMISSION, 2009: *The Admissibility of Expert Evidence in Criminal Proceedings in England and Wales. A New Approach to the Determination of Evidentiary Reliability. A Consultation Paper*, www.lawcom.gov.uk/expert_evidence.htm, Paper, 190.
- LEUNG, W. C., 2002: «The prosecutor's fallacy - a pitfall in interpreting probabilities in forensic evidence», *Medicine Science and the Law*, 42.
- MARTÍNEZ JARRETA, B. (dir.) 1999: *La prueba del ADN en Medicina Forense. La Genética al servicio de la paternidad*, Barcelona: Masson.
- NATIONAL RESEARCH COUNCIL OF THE NATIONAL ACADEMIES (NRC), 2009: *Report, Strengthening Scientific Evidence in the United States: A Path Forward*, Washington: The National Academies Press.
- ROBERTS, P., y ZUCKERMAN, A., 2004: *Criminal Evidence*, Oxford: Oxford University Press.

- ROYALL, R., 1997: *Statistical Evidence: A Likelihood Paradigm*, *Monographs on Statistics and Applied Probability*, London: Chapman-Hall/CRC.
- SAKS, M. J., y KOEHLER, J. J., 2005: «The Coming Paradigm Shift in Forensic Identification Science», *Science*, 5736 (309).
- 2008: «The Individualization Fallacy in Forensic Science Evidence», *Yanberbilt Law Review*, 1 (61).
- STONE, D. A., 1991: «What Made Us Ever Think We Could Individualize Using Statistics?», *Forensic Science Society*, 31 (2).
- TARONI, F.; AITKEN, C.; GARBOLINO, P., y BIEDERMANN, A., 2006: *Bayesian Networks and Probabilistic Inference in Forensic Science*, Chichester: John Wiley and Sons Ltd.
- TARUFFO, M., 2004: *La prueba de los hechos*, Madrid: Trotta.
- THOMPSON, W., y SCHUMANN, E., 1987: «Interpretation of statistical evidence in criminal trials. The prosecutor's fallacy and the defense attorney fallacy», *Law Human Behaviour Journal*, 11 (3).
- TRIBE, L. H., 1971: «Trial by Mathematics», *Harvard Law Review*, 6 (84).
- WILLIS, S. M., 2009: «Forensic science, ethics and criminal justice», en FRASER, J., y WILLIAMS, R. (eds.), *Handbook of forensic science*, Cullompton, UK: Willan Publishing.
- WRÓBLEWSKI, J., 1969: citado por la traducción de EZQUÍAGA GANUZAS, F. J. (1989): «La referencia a la ciencia especializada no excluye la valoración de las pruebas», en *La prueba jurídica: axiología, lógica y argumentación*, en *id.*, «Sentido» y «hecho» en el derecho, San Sebastián: UPV-EHU.