

CONOCIMIENTOS EXPERTOS Y DEFERENCIA DEL JUEZ*

(Apunte para la superación de un problema)

Marina Gascón Abellán

Universidad de Castilla-La Mancha

RESUMEN. Este trabajo analiza el problema de la deferencia judicial en el ámbito de las pruebas científicas y apunta algunos elementos para su superación. Se sostiene en primer lugar que, especialmente en el ámbito de las disciplinas forenses identificativas, la deferencia viene alimentada por el uso de expresiones individualizadoras para comunicar las conclusiones de las pericias, y que la crítica a estas prácticas ha impulsado un nuevo paradigma forense que exige recuperar para el juez su papel como decisor. A la luz de este paradigma se examinan dos exigencias para superar la deferencia: una relativa a la formulación del informe pericial y otra sobre el control judicial de fiabilidad de la prueba. Finalmente se sostiene que ninguna medida alcanzará ese objetivo si no se refuerza la base cognoscitiva del juez, y se ofrecen algunas sugerencias para ello.

Palabras clave: prueba pericial, prueba científica, deferencia judicial, control de fiabilidad de la prueba, formación de los jueces, valoración racional de la prueba.

Expert Evidence and Judicial Deference. (Some remarks for overcoming deference)

ABSTRACT. This paper addresses the problem of judicial deference to forensic experts and puts forward some ideas about how to overcome it. It is first argued that in the field of forensic identification deference is prompted by the widespread use of individualization statements when reporting forensic findings to the courts, and it is also held that the new forensic paradigm resulting from the criticism to this practice calls for restoring the courts to their role as decision-makers. On the basis of the new paradigm, two types of measures for overcoming deference are examined: one relating to the content of expert reports, the other to judicial scrutiny for their reliability. Finally, the claim is made that no measure will meet that objective without previously enhancing the cognitive basis of fact-finders, and suggestions in this respect are offered.

Keywords: expert testimony, forensic evidence, judicial deference, reliability of expert evidence, admissibility of scientific evidence, education of the judiciary, rational assessment of evidence.

* Fecha de recepción: 28 de julio de 2016. Fecha de aceptación: 20 de septiembre de 2016.

1. CONOCIMIENTOS EXPERTOS Y PRUEBA. SOBREALORACIÓN DE LAS PRUEBAS «CIENTÍFICAS» Y DEFERENCIA DEL JUEZ

En un mundo donde la especialización y complejidad del conocimiento es cada vez mayor el recurso a los expertos de las distintas ramas del saber juega un importante papel en la vida del derecho. La presencia de la experticia está particularmente presente en la resolución judicial de las controversias, donde expertos en los más variados asuntos auxilian diariamente a jueces y abogados realizando las pericias que precisan para el esclarecimiento de hechos controvertidos. La variedad de tipos de pericia es enorme. Hay expertos en ingeniería, en reconstrucción de accidentes, en contabilidad, en seguros, en paleontología, en medicina (en todas las especialidades médicas que seamos capaces de identificar), en bioquímica, en caligrafía, en marcas de mordeduras, en violencia doméstica, en arquitectura, en sociología, en informática, en diseño industrial, en balística, en análisis de la conducta, en ADN, en autenticidad de obras de arte, en religiones, en epidemiología, en síndrome post-traumático, en toxicología, en abuso sexual infantil, en huellas digitales, etc. Hay expertos para casi todo. Pero entre la inmensa variedad de pericias que juegan un papel en los tribunales destacan, por obvias razones, las que aparecen como «científicas». Sobre ellas centraré mi análisis¹.

Los constantes avances científicos y técnicos han tenido un profundo impacto en el ámbito de la prueba y, a través de las diversas técnicas forenses, juegan un papel cada vez más importante en todos los procesos. Los desarrollos han sido 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 resolución de problemas de identificación y el análisis de muestras biológicas de interés criminal, como manchas de sangre, saliva, esperma o pelos. El potencial de la huella genética es de tal magnitud que su uso en los tribunales se ha convertido ya en moneda corriente. Pero no es solo la prueba de ADN la protagonista de este *boom*. También hay otras pruebas científicas, como las previstas por la balística, la odontología forense, la acústica forense, el análisis de pintura, de drogas y otras sustancias, de fibras y materiales, de fluidos, de restos de fuego o de disparo, el análisis microscópico de cabellos, el examen de huellas dactilares, de escritura manuscrita, de marcas de herramientas o de mordeduras, de impresiones de calzado o de neumáticos, que juegan un papel protagonista en muchos procesos. Es más, las pruebas científicas se han convertido en la clave para probar algunos hechos que de otro modo difícilmente podrían haberse probado. Precisamente por su utilidad y rendimiento en el proceso, el recurso a estas pruebas es cada vez más demandado.

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 ellas. Más

¹ Por pruebas científicas entenderé aquí las vinculadas a los conocimientos de las llamadas a veces «ciencias duras», como la física, la química, las matemáticas, la biología, la ingeniería, etc. Quedan en cambio fuera de mi consideración aquí las conectadas a la psicología, la sociología u otros campos similares del saber que juegan también un importante papel en el proceso. Con ello no se pretende decir que esos otros campos del conocimiento no merezcan el nombre de «ciencias». Esa es otra cuestión. Pero me ceñiré aquí a las primeras porque representan el paradigma de lo que se entiende por «ciencia» y porque sus afirmaciones sobre el mundo gozan de un altísimo prestigio.

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 objetividad y prestigio que ha frenado la reflexión crítica sobre las mismas, con el resultado de que su fiabilidad y su valor probatorio se asumen con frecuencia como dogmas de fe². El conocido como «efecto CSI» levanta acta de esta situación. El fenómeno hace referencia al tremendo impacto que en el imaginario popular han tenido algunos *booms* televisivos como la serie norteamericana CSI (*Crime Scene investigation*) y otras similares que, al exhibir el exitoso uso policial de diversas técnicas forenses (especialmente las relacionadas con la criminalística identificativa), han aumentado de manera espectacular las expectativas depositadas en ellas. Con un efecto en la praxis judicial: sobre todo en los procesos penales se ha generado una demanda cada vez mayor de pruebas científicas y una reticencia a condenar en ausencia de las mismas. Diversos estudios cuestionan que estas series televisivas hayan potenciado la sobrevaloración de las pericias forenses, y algunos de ellos sugieren incluso que los espectadores habituales de estas series suelen ser más críticos con este tipo de pruebas³. En todo caso, y sean cuáles sean las causas principales del fenómeno, de lo que no cabe duda es de la gran confianza que en general se deposita en los informes que provienen de los «expertos de bata blanca».

La sobrevaloración de las pruebas científicas alimenta una *actitud deferencial* del juez hacia las declaraciones de los expertos. El fenómeno no puede sorprender. La deferencia es una actitud generalizada ante cualquier pericia (y no solo la científica), pues careciendo el juez de los conocimientos expertos que le permitirían hacer una valoración crítica del informe pericial es lógico que tienda a vincularse él. Pero cuando además se magnifica la confianza que ese tipo de pericia merece, la tendencia a la deferencia resulta mucho más acusada. Y esto es lo que sucede con las pruebas científicas. El alto prestigio del que gozan hace que el informe pericial se asuma sin mayor cuestionamiento y, de paso, el juez queda descargado de hacer un especial esfuerzo por fundar racionalmente la decisión: basta con alegar que hubo prueba científica y que esta apuntaba justamente en la dirección de la decisión probatoria final.

Pero la deferencia plantea problemas. El más evidente desde un punto de vista meramente institucional seguramente sea su complicado encaje en un sistema que confiere a los jueces (independientes e imparciales) la tarea de decidir sobre los hechos controvertidos y por consiguiente la de valorar las pruebas presentadas, incluidas las forenses. Porque, en efecto, la «cientificidad» de la prueba, por sí misma, no cierra la cuestión de su fiabilidad, y menos aún la de su valor probatorio, que debe ser resuelto por el tribunal en cada caso en virtud del principio de libre valoración⁴. Sin embargo cuando se actúa deferencialmente y las conclusiones del informe pericial se asumen sin

² De esta actitud beatificadora de los informes provenientes de los laboratorios de la policía científica advertía hace ya algunos años J. IGARTUA, «Prueba científica y decisión judicial: unas anotaciones propedéuticas», *Diario La Ley*, núm. 6812, 2 de noviembre de 2007.

³ Cfr. N. J. SCHWEITZER y J. SAKS, «The CSI Effect: Popular Fiction About Forensic Science Affects Public Expectations About Real Forensic Science», *Jurimetrics*, vol. 47, Spring 2007.

⁴ En el mismo sentido, J. WRÓBLEWSKI: «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, 1969, 185.

mayor cuestionamiento podría decirse que es el propio perito quien indica al juez lo que debe creer sobre el *factum probandum*. Pero de este modo los peritos se convierten en decisores de la causa y, por consiguiente, se termina instaurando un nuevo sistema de prueba fundado en la autoridad de los expertos. Naturalmente podría alegarse que la deferencia no solo es comprensible sino que además está justificada precisamente porque el juez carece de los conocimientos necesarios para cuestionar el informe pericial⁵. En definitiva, porque no puede ser *peritus peritorum*, por más que esa expresión esté muy extendida en la comunidad jurídica. Sin embargo eso no alivia en lo esencial el problema señalado, que sigue siendo la alteración que la deferencia produce en un sistema de prueba que confiere a los jueces (y no a los expertos) la decisión sobre los hechos.

Pero no es este el único (ni desde luego el más grave) problema que la deferencia plantea. El problema más serio es que el proceder deferencial implica una falta de control judicial sobre la fiabilidad de lo que el experto declara (y en particular sobre la fiabilidad de los conocimientos y técnicas usadas), lo que muchas veces permite la entrada en el proceso de auténtica basura pericial (*junk science*, en la ya famosa expresión popularizada por Peter HUBER⁶), datos con poco o nulo fundamento científico que son aceptados por los jueces y que pueden conducirles a cometer serios errores. Este peligro ha sido denunciado en el *NRC Report*, donde expresamente se dice que en algunos casos testimonios e informes periciales basados en análisis forenses defectuosos o en metodología con poco o insuficiente fundamento científico han podido contribuir a la condena de gente inocente, y se advierte, en consecuencia, del potencial peligro que comporta atribuir un peso indebido o desmedido al testimonio de los expertos⁷. Es más, 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 solo la fragilidad de los medios de prueba tradicionales, como los testimonios y las confesiones, sino también de muchas pruebas forenses sobre las que se basan algunas de estas condenas⁸. Muy brevemente, cuando el derecho y la pseudociencia se combinan los errores judiciales no son raros. Por eso la deferencia (al comportar una falta de control sobre la fiabilidad del dictamen pericial) no es la mejor actitud para el interés epistémico del proceso: realizar afirmaciones *veraces* sobre los hechos controvertidos⁹.

⁵ O podría alegarse que por lo menos algún tipo de «deferencial débil» puede estar justificada, como defiende por ejemplo T. WARD, «Expert Testimony, Law and Epistemic Authority», *Journal of Applied Philosophy*, 2016.

⁶ P. HUBER, *Galileo's Revenge: Junk Science in the Courtroom*, New York, New York Basic Books, 1991.

⁷ National Research Council Report, *Strengthening Scientific Evidence in the United States: A Path Forward*, Washington D.C., The National Academies Press, 2009 (en adelante NRC Report 2009), donde se lee expresamente: «For decades, the forensic science disciplines have produced valuable evidence that has contributed to the successful prosecution and conviction of criminals as well as to the exoneration of innocent people... Those advances, however, also have revealed that, in some cases, substantive information and testimony based on faulty forensic science analyses may have contributed to wrongful convictions of innocent people. This fact has demonstrated the potential danger of giving undue weight to evidence and testimony derived from imperfect testing and analysis. Moreover, imprecise or exaggerated expert testimony has sometimes contributed to the admission of erroneous or misleading evidence» (p. 4).

⁸ Puede consultarse la página web del proyecto en www.innocenceproject.org.

⁹ Cfr. en este punto R. J. ALLEN («The Conceptual Challenge of Expert Evidence», *Discusiones filosóficas*, vol. 14, núm. 23, 2013), quien subraya cómo el proceder deferencial ante los conocimientos expertos está en contradicción con la aspiración epistémica del proceso.

2. ALGUNAS CAUSAS DE LA DEFERENCIA. LA INDIVIDUALIZACIÓN EN LAS DISCIPLINAS FORENSES IDENTIFICATIVAS

La sobrevaloración de las pruebas científicas que subyace a la actitud deferencial hacia los informes y declaraciones de los expertos es en cierto modo sorprendente. Primero porque nada es menos «científico» que asumir como válido un conocimiento sin un previo control de sus postulados ajustado a una metodología científica. Pero además (y sobre todo) porque la mayor parte de estas pruebas están basadas en leyes estadísticas y sus resultados han de ser aún interpretados a la luz de otros datos, y por tanto difícilmente puede hablarse de «objetividad» y mucho menos de infalibilidad en relación con las conclusiones obtenidas a raíz de las mismas. No obstante hay varios elementos que explican esa sobrevaloración.

El primero de ellos tiene que ver con una deficiente cultura científica que conduce a depositar unas expectativas exageradas en los análisis y ensayos que provienen de la ciencia. En efecto, si las pruebas científicas no se han situado bajo una mirada crítica (como sí ha sucedido en cambio para el resto de pruebas) es porque se tiene una concepción irreal de lo que la ciencia puede hacer en la práctica. Porque se ignora que la mayoría de las pruebas científicas se expresan en un juicio de probabilidad en el caso particular que tiene un componente personal o subjetivo¹⁰. Porque se asume —aunque sea de forma vaga e intuitiva— que mientras las pruebas no científicas están basadas en generalizaciones empíricas de débil fundamento (por lo común máximas de experiencia), las científicas se fundan en conocimientos y leyes concluyentes que además son aplicadas con una rigurosa metodología, por lo que sus resultados, si la prueba se ha realizado bien, pueden tenerse por incuestionables o fuera de toda duda. En definitiva, porque se piensa que el conocimiento que se obtiene en las salas de los tribunales es frágil, y en cambio lo que sucede en los laboratorios de la policía científica es otra cosa.

Por otra parte, tampoco puede desconocerse la gran influencia que sobre esa visión acrítica de la ciencia forense ha tenido la prueba del ADN. El gran desarrollo científico de sus métodos y técnicas de análisis permite hoy sostener fundadamente su capacidad para vincular, con un alto nivel de fiabilidad, una muestra examinada con una persona concreta o un objeto. Ello no significa que en esta prueba no siga habiendo muchas cuestiones debatibles y que en todo caso puedan existir fallos de todo tipo en su realización¹¹. Significa tan solo que el nivel de investigación y discusión científica

¹⁰ «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 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 (F. TARONI, C. AITKEN, P. GARBOLINO y A. BIEDERMANN, *Bayesian Networks and Probabilistic Inference in Forensic Science*, Chichester, John Wiley and Sons Inc., 2006, cap. 1: «The Logic of Uncertainty», p. 21).

¹¹ Me remito al interesante y esclarecedor libro de E. E. MURPHY, *Inside the Cell: The Dark Side of Forensic DNA*, 2015. La autora muestra en su estudio la cara oscura de la prueba de ADN: laboratorios poco o nada supervisados que producen resultados inconsistentes; fiscales que fuerzan el análisis de muestras biológicas

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 gran motivo para desconfiar de sus resultados. Por ello no debe extrañar que el enorme crédito alcanzado por esta prueba genética (el *gold standard* de la ciencia forense) se haya proyectado también, como una suerte de «efecto irradiación», sobre otras muchas disciplinas forenses que, sin embargo, no gozan en absoluto del grado de investigación y desarrollo del análisis de ADN.

Pero la magnificación de las pruebas científicas no se explica solo por el aura de objetividad y prestigio que rodea la ciencia. Obedece también (y diría que sobre todo) al modo en que los expertos de las áreas de criminalística identificativa (es decir, las que tratan de identificar personas u objetos a partir de vestigios recogidos en la inspección ocular y relacionados con el hecho criminal) suelen comunicar a los tribunales las conclusiones de los análisis que han realizado. Poner el foco en lo que sucede en estas disciplinas me parece importante, porque constituyen una parte esencial de la aportación de la ciencia forense al proceso.

Los expertos de estas disciplinas comunican con frecuencia sus hallazgos usando expresiones individualizadoras: afirmaciones de que un vestigio hallado en la escena del crimen que presenta una coincidencia relevante con una fuente específica (porque ambos comparten un número suficiente de características) *proviene de esa fuente*. Por ejemplo, que el casquillo examinado mediante una prueba de balística salió de una cierta pistola; o que la impresión de calzado analizada pertenece a la víctima; o que la huella dactilar latente encontrada en el arma del crimen pertenece al acusado; o que la escritura manuscrita examinada fue hecha por el demandado; o que la voz analizada que realizó la llamada telefónica avisando de la colocación del artefacto explosivo pertenece al acusado, etc. En definitiva, los peritos expresan con frecuencia los resultados de las pruebas justamente en los términos en que el juez debe pronunciarse, y de este modo las pruebas aparecen como una poderosa herramienta que «cierra» la decisión judicial. Es verdad que frecuentemente los expertos comunican sus conclusiones de un modo más «técnico». Diciendo, por ejemplo, que los cabellos encontrados en el cuerpo de la víctima son compatibles con el sospechoso y añadiendo que la probabilidad de que provengan de cualquier otra persona es 1 entre 10.000. Pero esta forma más depurada de testificar no cambia, en mi opinión, mucho las cosas. Sigue sugiriendo que los análisis forenses son capaces de vincular el vestigio analizado con su fuente de procedencia, y en la medida en que los análisis se presentan como «científicos» sus resultados son muy persuasivos, particularmente cuando el experto viste una bata de laboratorio.

Esta manera de interpretar y comunicar los resultados de las pruebas expresa lo que en el ámbito de la ciencia forense se denomina el *paradigma de la individualiza-*

excesivamente pequeñas y de pobre calidad que hacen que sus resultados sean altamente subjetivos y puedan conducir a error; trazas del ADN de una persona aparecidas en un lugar donde esa persona nunca ha estado o en un objeto que nunca ha tocado y que podrían incriminarla en un hecho delictivo; errores que los jueces pueden cometer al interpretar los datos estadísticos que expresan la coincidencia entre los perfiles comparados; por no hablar de la masiva e inquietante compilación de datos que se realiza sin mayor control en las bases de datos de ADN policiales. Sobre los problemas de esta prueba genética puede verse también W. C. THOMPSON, «Tarnish on the “gold standard”: understanding recent problems in forensic DNA Testing», en *The Champion*, 2006, 10-16; o D. J. BALDING y J. BUCKLETON, «Interpreting low template DNA profiles», en *Forensic Science International: Genetics*, 4, 2009, 1-10.

ción, que se asienta sobre la asunción teórica de que la *unicidad* existe. La unicidad significa que un vestigio o marca desconocida m_x solo puede provenir de una fuente f_1 , con exclusión de todas las demás, y se refleja en la creencia de que podemos llegar a identificar plenamente un individuo o un objeto a partir de vestigios. La individualización, en palabras simples, reposa en la asunción teórica de que «la naturaleza nunca se repite». Mediante un análisis científico forense —es la idea rectora de este paradigma— se puede *individualizar* la marca o vestigio analizado, o sea se puede determinar con adecuada fiabilidad si esa marca o vestigio procede de una fuente específica (un individuo o un objeto), con exclusión de cualquier otra en el mundo¹². La mayoría de las técnicas forenses tradicionales de identificación criminal declaran en términos de individualización, asumiendo así no solo que la «unicidad» existe sino que además la individualización puede ser conseguida.

Las cosas, sin embargo, no son así, pues los conocimientos científicos disponibles no permiten sostener una tesis tan fuerte. De hecho en los últimos años el paradigma de la individualización ha recibido fuertes críticas por parte de la comunidad científica.

3. LA AUTOCRÍTICA EN LA CIENCIA FORENSE: LA NECESIDAD DE UN CAMBIO DE PARADIGMA

En los últimos años, en efecto, la comunidad científica forense, revisando el estado de sus diversas ramas, viene realizando una profunda autocrítica que se concentra particularmente sobre aquellas disciplinas que siguen comunicando sus conclusiones (sea en el informe pericial, sea en la eventual declaración del experto en el juicio) mediante el uso de expresiones individualizadoras. La expresión más acabada de estas críticas la constituye el ya mencionado *NRC Report 2009*, que considera que, por más que se trate de una forma de declarar simple e intuitiva, no resulta aceptable: la tesis no comprobada de que es posible vincular un vestigio de origen desconocido a una *única* fuente responde a una equivocada intuición probabilística que iguala infrecuencia con unicidad¹³. La unicidad no ha sido establecida en ninguna de las disciplinas forenses tradicionales, como el examen de escritura manuscrita, impresiones de neumáticos y de calzado, marcas de herramientas y armas de fuego o huellas dactilares. Pero es más, incluso si se considerase que la unicidad existe, tales disciplinas padecen fuertes

¹² Aludiendo al paradigma de la individualización, el *NRC Report 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» (p. 43).

¹³ M. J. SAKS y J. J. KOEHLER, «The Individualization Fallacy in Forensic Science Evidence», *Vanderbilt Law Review*, vol. 61, núm. 1 (2008), 199-219. D. H. KAYE, «Probability, Individualization and Uniqueness in Forensic Science: Listening to the Academies», *Brooklyn Law Review*, vol. 75, núm. 4 (2010). Como D. H. KAYE comenta (nota núm. 44), «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 D. A. STONEY, «What Made Us Ever Think We Could Individualize Using Statistics?», *J. Forensic Science Society*, vol. 31 (1991).

restricciones epistemológicas que impiden interpretar la *coincidencia* entre una marca desconocida m_x y una fuente f_1 en términos absolutos: la coincidencia entre m_x y f_1 se basa en que ambas comparten una serie de rasgos ($r_a...r_n$) que los expertos consideran suficientes para afirmar que m_x procede de f_1 ; pero lo cierto es que —salvo en casos excepcionales— no se ha comparado m_x (porque no se puede comparar) con todas las demás fuentes ($f_2...f_n$) de las que podrían provenir esos rasgos y, por consiguiente, no se han podido excluir esas fuentes. Por eso afirma el *NRC Report* que, con excepción del análisis de ADN nuclear, no existe base científica para expresar las conclusiones de los informes periciales en términos individualizadores¹⁴. Y también por eso, si un juez desapasionado examinara el estado actual de las ciencias forenses tradicionales muy probablemente consideraría sus afirmaciones como «plausibles, insuficientemente investigadas y sobrevendidas»¹⁵.

Los científicos y los estadísticos sostienen que el único modo riguroso de interpretar los resultados de estas pruebas es hacerlo, no en términos de individualización (o de identificación de un vestigio con una fuente) sino en términos de *verosimilitud* (*likelihood*). Las verosimilitudes expresan la probabilidad de observar un dato encontrado si se considera que es cierta una hipótesis: por ejemplo, qué probabilidad hay de que se observen los datos encontrados en una pericia si se considera cierta la tesis defendida por la acusación, o qué probabilidad hay de que se observen esos datos si se considera cierta la tesis defendida por la defensa. Aunque el juez está interesado en conocer las probabilidades de las hipótesis defendidas por las partes a la luz de los datos resultantes de las pericias, en criminalística identificativa todo lo que el perito puede hacer es calcular verosimilitudes. Por dejarlo más claro: lo que el perito puede (o debe) hacer cuando interpreta y comunica los resultados de las pruebas realizadas es expresar una ratio de verosimilitud (una *likelihood ratio*), o sea la *probabilidad de observar los datos analíticos y técnicos obtenidos en los test* si se considera cierta una hipótesis frente a la probabilidad de observar esos mismos datos si se considera cierta la hipótesis contraria; y no al revés, es decir *no puede hablar de la probabilidad de las hipótesis judiciales* a la luz de esos datos¹⁶. O aún más claramente: la *probabilidad* de una hipótesis a la luz de los datos obtenidos en las pruebas practicadas (que es lo que el juez desea saber) mide la *incertidumbre* de esa hipótesis; la *ratio de verosimilitud* (que es lo que el experto puede aportarle al juez) mide la *fuerza con que los datos apoyan* una hipótesis frente a la otra.

La consecuencia de lo anterior es clara: no existe fundamento racional para que el perito sostenga categóricamente en su informe que ha sido capaz de identificar a una persona o un objeto a partir de los análisis de un vestigio llevados a cabo en el laboratorio¹⁷; como tampoco existe fundamento para que realice una valoración

¹⁴ Así lo señala el *NRC Report* 2009, 7.

¹⁵ M. J. SAKS y J. J. KOEHLER, «The Coming Paradigm Shift in Forensic Identification Science», *Science*, vol. 309 (2005), 892.

¹⁶ Esta necesidad de interpretar y comunicar al tribunal los resultados de los análisis realizados en una forma apropiada ha sido fuertemente destacada por la obra de C. AITKEN y F. TARONI, *Statistics and the Evaluation of Evidence for Forensic Scientists*, 2.ª ed., Chichester, John Wiley & Sons, 2004. Y en ello ha insistido también después el *NRC Report* 2009.

¹⁷ Para un análisis detallado de las críticas recibidas por el paradigma de la individualización, remito al trabajo conjunto M. GASCÓN y J. J. LUCENA, «Razones científico-jurídicas para valorar la prueba científica: una argumentación multidisciplinar», en *Revista La Ley*, núm. 7481, 4 de octubre de 2010.

probabilística sobre la posibilidad de atribuir ese vestigio a esa persona u objeto. El resultado de un análisis de voz, por ejemplo, no dice, ni categórica ni probabilísticamente, que la voz analizada pertenezca a tal o a cual persona, sino que solo aporta *datos* que, una vez interpretados con las adecuadas herramientas estadísticas, expresan una *ratio de verosimilitud* y dicen cosas del siguiente tipo: «es *x* veces más probable que se observen los datos encontrados en el análisis si la voz pertenece al acusado que si no pertenece a él». Por eso, la extendida costumbre (particularmente arraigada en las técnicas forenses tradicionales) de declarar en términos de identificación categórica o simplemente probabilística, estableciendo así lo que debe creerse sobre la hipótesis judicial en examen, debe ser abandonada¹⁸. Por lo demás, como ha subrayado Christophe 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 erradicado¹⁹.

El llamado por Richard ROYALL *paradigma de la verosimilitud*²⁰ es —frente al de la individualización— el que se presenta como modelo a seguir en las disciplinas forenses identificativas. Dicho paradigma se erige sobre la distinción entre las tres preguntas básicas que cabe formular una vez que el perito ha realizado los análisis o exámenes pertinentes. Supongamos que, mediante una prueba de comparación de voces, se ha comparado una voz desconocida con la voz de un sospechoso y se ha obtenido un resultado positivo: se ha observado una coincidencia relevante entre las voces. Las tres preguntas mencionadas son las siguientes: ¿constituye esa observación algún *apoyo* en favor de la hipótesis de que la voz es del sospechoso?; ¿puedo *creer* a partir de ahí que la voz es del sospechoso?; ¿puedo *actuar* a partir de ahí como si la voz fuese del sospechoso? Sumariamente: (1) ¿*qué nos dicen los datos* u observaciones como prueba en favor de la hipótesis en examen? (2) ¿*qué ha de creerse* a partir de esos datos?; y (3) ¿*qué ha de hacerse*?²¹. En mi opinión, esta distinción se presenta como una importante herramienta para reconstruir la interacción entre el experto y el juez cuando se ha realizado una pericia científica. De hecho permite realizar un deslinde neto entre las tareas del perito y las del juez y recupera para este su rol como decisor. Brevemente: proporciona una razón para superar la deferencia.

¹⁸ 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 estadísticos apoyan un resultado cuando en realidad sucede justo lo contrario* (R. ROYALL, *Statistical Evidence: A Likelihood Paradigm*, Monographs on Statistics and Applied Probability, London, Chapman & Hall/CRC, 1997, Prefacio, xi). Y esto último sucede, no porque los expertos usen mal la estadística sino justamente porque domina el paradigma de la individualización.

¹⁹ Ch. CHAMPOD, *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, 8-11 September 2009.

²⁰ R. Royall, *Statistical Evidence: A Likelihood Paradigm*, *op. cit.*

²¹ Para un desarrollo más extenso del alcance de este paradigma remito a «Razones científico jurídicas para valorar la prueba científica: una argumentación multidisciplinar» (M. GASCÓN y J. J. LUCENA), *La Ley*, *op. cit.*

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 seguir con nuestro ejemplo: una prueba de comparación de voces) deberá *interpretar el resultado* de la prueba en el laboratorio (supongamos que positivo: los rasgos de las voces comparadas presentan una coincidencia relevante) dándole un alcance en relación con las hipótesis enfrentadas (supongamos que H_a : la voz pertenece al sospechoso; y H_b : la voz no pertenece al sospechoso). Muy simplemente, se trata de interpretar si los datos apoyan (y con qué fuerza apoyan) una hipótesis frente a la otra. Recordemos que la interpretación de esos resultados ha de expresarse (verbal o numéricamente) mediante una *likelihood ratio*, y es importante que se haga bien, porque es lo que se refleja en las conclusiones del informe pericial y por consiguiente lo que se va a comunicar al juez.

Lo que ha de 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 sobre la hipótesis litigiosa (es decir, cuál es su *probabilidad*) a partir de esos datos, y la respuesta dependerá no solo de la información suministrada por el experto (lo que dicen los datos una vez interpretados por él) sino también de la derivada del resto de pruebas disponibles en la causa. Esta tarea decisional expresa la *valoración de la prueba* y es evidente que corresponde al juez, por dos simples razones. En primer lugar, porque lo que debe creerse sobre la hipótesis litigiosa depende —ya se ha dicho— no solo de lo que dicen los datos obtenidos en la pericia sino también de lo que dicen las demás pruebas disponibles, y es el juez quien posee esa información. En segundo término (y decisivamente) porque son los jueces, independientes e imparciales, quienes tienen *institucionalmente* atribuida la tarea de decidir el conflicto, y en consecuencia son ellos (y no los expertos) quienes deben determinar también lo que hay que creer sobre los hechos litigiosos una vez que se ha practicado la prueba. Por eso resulta claro que el paradigma de la individualización no es compatible con este esencial principio del sistema jurídico, pues si el informe pericial usara expresiones individualizadoras del tipo «la voz desconocida analizada proviene (o es probable que provenga) del sospechoso», habría sido el perito (y no el juez) quien habría determinado lo que hay que creer sobre ese punto.

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 ambas tareas, pues propicia que el perito declare en términos de «el vestigio v_x procede de la fuente f_1 (por ejemplo, el acusado)», que son los términos en que debe expresarse el juez como resultado de haber valorado prueba, e impulsa de este modo una actitud deferencial del juez²².

²² Por lo demás, este modo de declarar los expertos no solo impulsa la deferencia sino que además da lugar a múltiples malinterpretaciones y en ocasiones a serios errores judiciales. Esto, por ejemplo, es lo que

Lo que ha de hacerse

Por último, el paradigma de la verosimilitud traza también una distinción entre lo que ha de *creerse* y lo que ha de *hacerse* a partir de la realización de una prueba científica, que resulta sumamente apropiada para reconstruir otra distinción esencial dentro del universo de la prueba: la existente entre la valoración de la prueba propiamente dicha y la decisión sobre los hechos probados. Mientras que lo que ha de *creerse* fundamentamente sobre una hipótesis se expresa como la probabilidad de esa hipótesis y se identifica con la *valoración* de la prueba, lo que ha de *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 la hipótesis para que podamos considerarla probada y actuar en consecuencia. Y así como la valoración es (o debería ser) una cuestión completamente librada a la *racionalidad epistémica*, la fijación de estándares de prueba es también una cuestión *política o valorativa*, dependiente de lo tolerante que el sistema esté dispuesto a ser con los dos tipos de error que pueden cometerse al adoptar una decisión sobre hechos: el error consistente en considerar probado lo que es falso, y el consistente en no considerar probado lo que es verdadero. Por ejemplo, el estándar de prueba en el proceso penal suele ser alto (y más alto que en el civil), porque somos menos tolerantes con el error de condenar a un inocente que con el de absolver a un culpable.

En suma, *lo que ha de creerse* sobre una cierta hipótesis depende de lo que dicen los datos resultantes de la prueba pericial una vez interpretados y del resto de las informaciones y pruebas disponibles en la causa. *Lo que ha de hacerse* depende obviamente de lo que debe creerse, 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 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. Es evidente, en todo caso, que ambas tareas corresponden al *juez*, y no al experto.

4. LA SUPERACIÓN DE LA DEFERENCIA

Desde el punto de vista del derecho la adopción del paradigma de la verosimilitud se traduce —frente al de la individualización— en un punto simple pero central: la devolución a los jueces de su papel como decisores, y por tanto la superación de la deferencia. Son los jueces quienes efectivamente (y no solo formalmente, como ocurre cuando se declara con expresiones individualizadoras) han de valorar la prueba, eva-

explica que se produzcan las conocidas *falacias del fiscal y de la defensa*. Sobre estas falacias *cf.* W. THOMPSON y E. SCHUMANN, «Interpretation of statistical evidence in criminal trials. The prosecutor's fallacy and the defense attorney fallacy», *Law Human Behav.*, vol. 11 (1987), 167 y ss. Y entre nosotros A. CARRACEDO, «Valoración de la prueba del ADN», en B. MARTÍNEZ JARRETA (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, 1999, 303. Sobre la falacia del fiscal puede consultarse más recientemente W. C. LEUNG, «The prosecutor's fallacy – a pitfall in interpreting probabilities in forensic evidence», *Medicine Science and the Law*, 42 (2002), 44 y ss.

luando su *fiabilidad*, primero, y determinando su *valor probatorio* después. Pero conseguir que su decisión en este punto sea racional implica algunas exigencias. La primera de ellas se impone sobre el derecho y se traduce en adoptar las medidas necesarias para que el control de fiabilidad de la prueba que los jueces han de realizar se haga con base en el conocimiento y no arbitrariamente. Dicho de manera breve, se trata de asegurar que el juez pueda ejercer, fundada o racionalmente, su rol como *gatekeeper*. La segunda exigencia se impone sobre los expertos y se traduce en la necesidad de que estos aporten información completa y adecuada sobre las pruebas practicadas, a fin de que los jueces puedan determinar su fiabilidad, y de comunicar los resultados de las mismas de un modo apropiado, a fin de evitar que sean malentendidos o sobrevalorados por los jueces. Dicho también de forma breve, se trata de formular adecuadamente el informe pericial y sus conclusiones. Ambas exigencias son de un interés epistémico máximo para el estado de derecho, pues con ellas se pretende garantizar que las decisiones probatorias basadas en pruebas científicas sean «racionales»: adoptadas por *jueces independientes, imparciales y correctamente informados*.

4.1. Sobre los controles de fiabilidad de las pruebas

Que son los jueces (y no los expertos) quienes deben decidir significa que son ellos, correctamente informados, quienes deben determinar el valor probatorio que cabe atribuir a la prueba y, aún antes, controlar o evaluar su fiabilidad; es decir, evaluar si es creíble lo que el experto nos dice. La fiabilidad depende de (y por consiguiente exige evaluar) varios factores²³. Depende en primer término —esto es obvio— de la *calidad técnica* de la prueba; o sea, de que se haya realizado bien. Es importante asegurar que los exámenes y análisis se han realizado correctamente: por personal cualificado, siguiendo los protocolos apropiados y rigurosos controles de calidad. Y es importante también que los vestigios analizados hayan sido obtenidos y conservados adecuadamente. Si una huella dactilar está emborronada o difuminada cuando es recogida su examen difícilmente puede ser útil. Y en relación a una prueba de ADN, si los vestigios biológicos analizados 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 vestigios ha de hacerse con sumo cuidado, y el mantenimiento de la cadena de custodia es crucial para que no pierdan su valor probatorio²⁴. Por lo demás, no hay que olvidar que las pruebas las realizan personas de carne y hueso y por tanto pueden cometer errores. No son pocos los estudios que en los últimos tiempos llaman la atención sobre los *sesgos cognitivos* que pueden existir en la realización de algunas pruebas forenses²⁵, sobre todo de aquellas tradicionales que,

²³ En su informe «Scientific Evidence in Europe. Admissibility, Appraisal and Equality of Arms», *International Commentary on Evidence*, 9 (2011), Ch. CHAMPOD y J. VUILLE realizan un exhaustivo repaso por los factores que pueden influir negativamente en el valor probatorio atribuible a los resultados de las pruebas científicas.

²⁴ 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.

²⁵ Con respecto a los sesgos cognitivos *cf.* I. DROR, «How can Francis Bacon help forensic science? The four idols of human biases», *Jurimetrics Journal*, 50 (2009), 93-110; I. E. DROR y S. A. COLE, «The vision in

como la dactiloscopia y la grafística, tienen un fuerte componente comparativo que las deja enteramente bajo la supervisión del perito²⁶.

Pero el problema más grande en relación con la fiabilidad de una prueba es si hay o no ciencia detrás de la disciplina: si la técnica usada no tiene validez científica no cabe atribuir a sus resultados mayor fuerza probatoria. Por eso, controlar que los conocimientos expertos admitidos al proceso tengan un sólido fundamento es una garantía de racionalidad, pues dicho control aspira a dejar a la pseudociencia fuera del ámbito de las decisiones judiciales. Cómo conseguir esto es lo que puede resultar controvertido.

En línea de principio pueden distinguirse dos «modelos», dependiendo de si el control está guiado por algún criterio o si, por el contrario, se deja enteramente en manos de los jueces que lo ejercen sin auxilio o guía alguna. Esta segunda opción evidencia claramente un problema: los jueces, por lo general, carecen de la formación necesaria para evaluar la calidad científica de las pruebas. Por eso dejar en sus manos la libre apreciación de este asunto, sin establecer algún criterio que los guíe, puede conducirlos a rechazar pruebas científicamente válidas que podrían haber ayudado al esclarecimiento de la verdad, o —lo que es más probable— puede conducirlos a dar crédito a pruebas que son pura charlatanería pseudocientífica y que pueden convertirse en fundamento de decisiones judiciales erróneas. No debe extrañar por ello que en algunos sistemas se hayan establecido criterios para guiar al juez en su tarea de evaluar la fiabilidad de las pruebas. Esta es la opción que se ha seguido en los Estados Unidos con lo que se conoce como *Dauber test* y también la recientemente sugerida por *The Law Commission* para el Reino Unido siguiendo los pasos del primero²⁷.

En *Dauber* la Corte Suprema ha sugerido como estándar de admisibilidad de las teorías o técnicas usadas, además de su *aceptación general* por parte de la comunidad científica concernida, la *validez científica del método*, lo que puede acreditarse evaluando los siguientes factores: *a*) si la teoría o técnica es refutable y verificable y si ha sido efectivamente puesta a prueba (verificabilidad y refutabilidad como criterios mínimos de cientificidad de cualquier teoría); *b*) si han sido sometidas a revisión de pares y publicadas en revistas especializadas, *c*) si se conocen las tasas de errores potenciales asociados al método (que la técnica empleada tenga una precisión conocida), y *d*) si existen procedimientos estandarizados y controles de calidad que rigen la práctica. No es que el *Dauber test* exija que cualquier técnica 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 relevante— escrutar de manera más estricta la «cientificidad» de muchas pruebas que hasta ahora no habían sido cuestionadas y conduce así

“blind” justice: Expert perception, judgment, and visual cognition in forensic pattern recognition», en *Psychonomic Bulletin & Review*, 17 (2010), 161-167.

²⁶ Ch. CHAMPOD y J. VUILLE, *Scientific Evidence in Europe*, op. cit., 10-11.

²⁷ El estándar fue establecido por *Daubert v. Merrell Dow Pharmaceuticals Inc.*, 113 S.Ct.2786 (1993). Los criterios sugeridos para el Reino Unido por *The Law Commission* constituyen prácticamente una réplica del criterio *Dauber* (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*, Consultation Paper No 190, 2009, 53 y ss. Disponible en http://www.lawcom.gov.uk/wp-content/uploads/2015/03/cp190_Expert_Evidence_Consultation.pdf).

a que estos recuperen su papel de *gatekeeper*²⁸. Por decirlo muy sumariamente, la sentencia *Dauber* supone un llamamiento a los jueces para que miren más críticamente las pericias científicas antes de atribuir un valor probatorio a sus resultados²⁹. De hecho, después de esta sentencia, ámbitos tradicionales de la policía científica, como las huellas dactilares, la grafística y la balística empiezan a ser puestos seriamente en cuestión.

Pero el estándar *Dauber* no es la panacea a la falta de formación de los jueces en las técnicas forenses que deben controlar; y de hecho no ha zanjado la polémica, porque su aplicación presupone que los jueces son capaces de comprender y aplicar los conceptos científicos para ello, lo que normalmente no es cierto³⁰. Eso no significa que los test como el *Dauber* no sean valiosos. Su propia existencia constituye una llamada de atención sobre la necesidad de ejercer un control crítico sobre la fiabilidad de las técnicas forenses propuestas; y desde luego ayudan a simplificar el control, a hacerlo más «manejable» para los jueces. Lo que significa es que el simple establecimiento de este tipo de test no cierra el problema, que sigue siendo la falta de una apropiada formación para decidir racionalmente o con fundamento. Por supuesto cuando la prueba ha sido ya objeto de una intensa investigación y aceptación en la comunidad científica, como sucede con la prueba del ADN, no se plantea mayor cuestión sobre su fiabilidad. Pero hay otras especialidades de la ciencia forense de mayor fragilidad y controversia que la prueba de ADN³¹ y sobre las que se fundan diariamente decisiones judiciales. 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 (o al menos en los más habituales) y de familiarizarlos con las nuevas técnicas que vayan surgiendo. Si carecen de la información y conocimientos necesarios para decidir con sentido, su decisión o será deferencial o será irracional.

En resumen, la superación de la deferencia cuando se ha realizado una prueba científica implica que son necesarias algunas medidas para garantizar que sean efectivamente los jueces (independientes e imparciales) y no los expertos quienes deciden

²⁸ 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 simplemente que el testimonio de los expertos será admisible si resulta útil para el esclarecimiento de los hechos (es decir, no exigía —al menos explícitamente— el criterio de la «aceptación general»). La sentencia *Daubert* respondió afirmativamente a esa cuestión, pero estableció también que los tribunales siguen teniendo el deber de escrutar si el testimonio de los expertos cumple esos parámetros de relevancia y de «fiabilidad», y cifró la fiabilidad en la validez científica de la «metodología» usada, expresada en los factores *Dauber* arriba mencionados. Para una clara exposición del establecimiento y desarrollo del estándar *Dauber* (así como una crítica de las confusiones y malentendidos que los «factores *Dauber*» encierran), *cf.* S. HAACK, «The Expert Witness: Lessons from the U.S. Experience», *Humana.Mente*, 28 (2015), 50 y ss.

²⁹ Con elocuentes palabras Ch. CHAMPOD y J. VUILLE 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*, *op. cit.*, 26.

³⁰ En este sentido, S. HAACK, «The Expert Witness: Lessons from the U.S. Experience», *op. cit.*, 64.

³¹ El *NRC Report 2009* señala que «en muchas disciplinas forenses no se ha hecho investigación sistemática y rigurosa para validar las premisas básicas y las técnicas» y, de hecho, «en la mayoría de las áreas forenses no existe un sistema bien definido para determinar las tasas de error, y los controles de calidad realizados muestran que algunos peritos no lo hacen bien. En algunas disciplinas, como la odontología forense, los métodos de obtención de las pruebas son relativamente no controvertidos, pero existen disputas sobre el valor y fiabilidad de las interpretaciones resultantes» (pp. 188-189).

sobre su fiabilidad; para devolverles, en definitiva, su papel de *gatekeepers*. Pero centrar la cuestión en la ausencia o presencia de test de fiabilidad seguramente desvía la atención del verdadero problema: si poseen o no los conocimientos apropiados para decidir «racionalmente» sobre estos aspectos.

4.2. Comunicación de los resultados periciales y formulación del IP

Pero la superación de la deferencia depende también (y diría que sobre todo) de una cuestión central a la que desgraciadamente no se ha prestado demasiada atención³²: cómo deben los peritos formular el informe pericial y sus conclusiones. En concreto: ¿qué deben incluir en su informe para que el juez pueda evaluar la fiabilidad de la prueba realizada? y ¿cómo deben interpretar los resultados de los test y expresarlos en sus conclusiones para evitar que su valor probatorio sean malentendido o sobrevalorado por los jueces? La cuestión —digámoslo otra vez— tiene gran importancia, pues es clave para asegurar que puedan ser los jueces (y no los expertos) quienes decidan y para que su decisión además sea racional: si el informe pericial no contiene información completa y adecuada sobre la prueba realizada o los datos obtenidos son interpretados y presentados con una exagerada o distorsionadora confianza, será difícil que el juez pueda hacer una recta evaluación de su fiabilidad y valor probatorio.

Precisamente con el fin de permitir al juez una correcta valoración de la prueba, los expertos —como indica el *NRC Report*— deberían describir como mínimo «los métodos y materiales, procedimientos, resultados y conclusiones, y deberían identificar de forma adecuada las fuentes de incertidumbre en los procedimientos y conclusiones así como la estimación de su magnitud (para indicar el grado de confianza que merecen los resultados)»³³. El informe, dicho sistemáticamente, debería: (1) explicitar adecuadamente los *métodos y procedimientos usados* (señalando las fuentes de *incertidumbre* asociadas a los mismos y, cuando sea posible, la estimación de su valor), a fin de que el juez pueda valorar la *validez y fiabilidad* de los mismos; y (2) dar cuenta rigurosa de los *datos resultantes y las conclusiones obtenidas* (señalando claramente sus *limitaciones* y, cuando sea posible, las probabilidades asociadas), a fin de que el juez pueda comprender su significado y atribuirles un *valor probatorio*. En relación con este último aspecto (las conclusiones del informe) debe subrayarse una cuestión central. De acuerdo con el paradigma de la verosimilitud, al perito no le corresponde emitir una opinión sobre el *factum probandum* sino que debe limitarse a interpretar los datos resultantes de la prueba de un modo científicamente riguroso, procurando que el juez entienda exactamente su significado y pueda valorarlos junto con el resto de pruebas disponibles. Por eso conviene insistir una vez más: las conclusiones de la prueba que el perito formula en su informe se orientan a determinar si los datos analíticos y técnicos obtenidos apoyan (y con qué fuerza apoyan) las hipótesis en consideración, y no lo que

³² De ello se lamentaba no hace mucho, entre otros, S. M. WILLIS, «Forensic science, ethics and criminal justice», en J. FRASER and R. WILLIAMS (eds.), *Handbook of forensic science*, Cullompton, UK, Willan Publishing, 2009, 523-545. El *NRC Report 2009* también presta atención a este asunto, al establecer que es absolutamente necesario sentar las bases «para la correcta interpretación y comunicación de los hallazgos forenses» (*op. cit.*, 186).

³³ *Ibid.*: 186.

hay que creer sobre esas hipótesis a la luz de los datos. Entender que las conclusiones de la prueba deben hablar en estos últimos términos (esto es, que deben decirle al juez justamente lo que él desea saber para fundar su decisión: si la hipótesis litigiosa es o no verdadera, o con qué grado de probabilidad puede afirmarse que lo es) es propio del paradigma de la individualización, que no distingue entre lo que dicen los datos (una tarea pericial) y lo que debe creerse sobre el *factum probandum* a partir de los mismos (una tarea judicial), y propicia de este modo una actitud deferencial del juez.

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 identificativas categóricas (del tipo: «la huella latente examinada *b* es del acusado») o probabilísticas (del tipo: «es muy probable que la huella latente examinada *b* sea del acusado») y debe en cambio optarse por pronunciamientos del tipo «es *x* veces más probable que se observen los datos encontrados en el análisis si la huella latente examinada *b* es del acusado que si no lo es»³⁴. Por eso las llamadas escalas verbales de probabilidad, ampliamente usadas en el cotejo de huellas digitales, examen de escritura, balística, comparación de caras, identificación de voz mediante análisis lingüístico, cotejo de huellas de calzado, neumáticos y herramientas, etc., son inapropiadas. De hecho estas escalas reproducen el paradigma de la individualización estableciendo, no lo que dicen los datos, sino lo que debe creerse sobre el *factum probandum* a partir de los mismos, y de este modo impulsan —digámoslo de nuevo— una praxis deferencial. En este sentido un importante reto de la ciencia forense 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 punto³⁵. Hay que decir que en este punto —afortunadamente— ya se están dando pasos importantes, entre los que merece destacar la reciente aprobación de un estándar europeo sobre conclusiones de informes periciales de carácter evaluativo, que resulta acorde con las comentadas exigencias avanzadas desde la comunidad científica forense³⁶.

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 estadística³⁷. Por eso, del mismo modo que

³⁴ Me permito remitir aquí, de nuevo, a nuestro trabajo M. GASCÓN, J. J. LUCENA y J. GONZÁLEZ, «Razones científico-jurídicas para valorar la prueba científica», *op. cit.*

³⁵ El *NRC Report 2009*, de nuevo, ya señalaba esta exigencia al decir que la mayoría de las disciplinas forenses necesitan también «estandarizar y clarificar la terminología usada en los informes y testimonios sobre los resultados y en la aportación de más información» (*op. cit.*, 189).

³⁶ Me refiero al documento *Guideline for Evaluative Reporting in Forensic Science. Strengthening the Evaluation of Forensic Results across Europe* (STEOFRAE), de 2015, elaborado por ENFSI (European Network of Forensic Science Institutes), que está disponible en <http://www.enfsi.eu/documents/forensic-guidelines>. En España, además, la CNUFADN del Ministerio de Justicia ha aprobado también en octubre de 2015 unas *Recomendaciones sobre el informe pericial y la expresión de resultados en materia de análisis genéticos forenses*, que recoge tanto el modo en que han de presentarse los resultados como el modo en que han de ser interpretados por los peritos.

³⁷ Por ceñirme 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 Ch. CHAMPOD y J. VUILLE, *Scientific Evidence in Europe*, *op. cit.*, 10.

—como ya se señaló— sería conveniente proporcionar a los jueces alguna formación sobre las diferentes áreas forenses para que puedan decidir con sentido si la prueba es fiable, también sería conveniente familiarizarlos con algunos conceptos estadísticos básicos, como relación de verosimilitud, tasa de error, etc., para que puedan entender correctamente el significado y alcance de los resultados de las pruebas antes de asignarles un valor probatorio. Obviamente no se trata de convertir a los jueces en matemáticos o en *amateur scientists*, sino de proporcionarles los conocimientos necesarios para conjurar el riesgo de tergiversación o de sobrevaloración de los datos estadísticos en los que muchas veces se expresa el resultado de una prueba científica.

En conclusión: la (libre) valoración que hace el juez de los datos aportados por una prueba científica y expresados en el informe pericial será racional si está fundada en el conocimiento; o sea, si el juez es capaz de entender: 1) cómo se han alcanzado esos datos (qué técnica se ha usado y cuál es su crédito científico, cómo se han obtenido los vestigios, cómo ha sido la cadena de custodia y cómo se ha realizado el análisis) y por tanto qué grado de fiabilidad tienen; y 2) cómo hay que interpretarlos, y por tanto qué valor probatorio tienen una vez ponderados con el resto de las pruebas. Ese objetivo racionalizador solo se alcanzará si se le suministra al juez la información necesaria (a través de un informe pericial completo y correcto) y se le proporciona una formación adecuada (familiarizándolo al menos con las principales áreas forenses y con los conceptos estadísticos más básicos implicados en la interpretación de los datos resultantes de las pruebas). Sin ese conocimiento la decisión del juez, o estará basada en la confianza en el experto (o sea, será deferencial), o será irracional. El hecho de que el informe pericial pueda resultar controvertido por otro informe distinto obviamente es importante, pero no cambia mucho las cosas. Puede parecer que la confrontación de testimonios y la *crossexamination* garantizan, ahora sí, una decisión judicial racional. Pero lo cierto es que el juez deberá decidir entre ellos, y si carece de la información y conocimientos apropiados su decisión *no estará fundada en elementos cognoscitivos* sino de otro tipo. Normalmente en la capacidad de una de las partes para «persuadirlo», por ejemplo mediante la presentación de una gran cantidad de expertos que la otra parte no es capaz de contrarrestar por el costo que ello supone. Es verdad que habrá decidido el juez, pero lo habrá hecho en medio de un «vacío cognoscitivo», con el consiguiente riesgo de un error. En suma, si los informes periciales no se formulan en los términos apropiados y, en general, si los jueces carecen de la formación necesaria para entender la base cognoscitiva de las pruebas que deben valorar, su decisión no estará epistémicamente fundada. No será racional.

4.3. Recapitulando: superación de la deferencia y educación de los jueces

Hasta aquí he sostenido que la deferencia del juez ante los conocimientos expertos, siendo una tendencia generalizada ante todo tipo de pericia, se acentúa especialmente cuando esos conocimientos vienen arropados por el enorme prestigio de la ciencia. He sostenido también que, en las disciplinas forenses identificativas, el extendido uso de expresiones individualizadoras para comunicar las conclusiones de las pruebas propicia o impulsa la deferencia mientras que el paradigma de la verosimilitud, por el contrario, subraya la necesidad de recuperar para el juez su papel como decisor.

Asimismo he señalado que la superación de la deferencia comporta exigencias tanto para las disciplinas forenses (que deben formular adecuadamente sus informes y sus conclusiones a fin de que los jueces puedan controlar la fiabilidad de las pruebas y entender cabalmente el valor probatorio de sus resultados) cuanto para el derecho (que debe adoptar las medidas necesarias para que el control de fiabilidad de la prueba que el juez debe realizar se haga con base en el conocimiento y no arbitrariamente). Pero —como he tratado de decir también— ni el simple establecimiento de controles de fiabilidad ni la adecuada formulación del informe pericial son suficientes. Si no se proporciona a los jueces información y adiestramiento para que puedan entender si algo ha ido mal en la prueba realizada (si la metodología usada resulta discutida, o si hay datos relevantes que han sido obviados o distorsionados en su transmisión, o si el juicio del experto puede estar afectado por sesgos cognitivos, o si el supuesto experto es tan solo un pseudo-experto, etc.) y para que puedan valorar adecuadamente su fuerza probatoria (sin infravalorarla ni sobrevalorarla) su decisión estará basada en la confianza en los expertos o será en todo caso irracional³⁸. En suma, una decisión puramente no-deferencial no es, por sí misma, racional. Lo será solo si está basada en el conocimiento, y para eso alguna educación es necesaria³⁹.

No-deferencia y educación son, pues, los dos ingredientes de una decisión probatoria de calidad, basada en el conocimiento y no en la pura confianza en el experto. La educación significa que los jueces, que en última instancia son quienes deben decidir sobre la fiabilidad de las pericias y su valor probatorio, tienen la *información* y la *formación* necesarias para hacerlo. Apunta por tanto a la necesidad de *aumentar o reforzar la base cognoscitiva* del juez. Por eso la educación puede (y debe) obtenerse por distintas vías. El propio *informe pericial*, completo, riguroso y adecuadamente formulado, es una fuente imprescindible de información. Y desde luego no cabe dudar del importante *papel auxiliar de los expertos* en la tarea de interpretación del informe, por lo que su presencia en el juicio reviste una importancia (educativa) crucial. Pero cabe pensar también en otras vías. Por ejemplo en la posibilidad de establecer alguna *instancia independiente* que pueda ayudar a los jueces a decidir en los casos complicados, como la propuesta por The Law Commission para el Reino Unido⁴⁰ o la sugerida por Christophe CHAMPOD y Joëlle VUILLE⁴¹ para Europa. Y por supuesto hay que pensar también en la propia *formación de los jueces* (a través, por ejemplo, de programas educativos orga-

³⁸ El planteamiento del problema de los conocimientos expertos en el proceso en términos de deferencia *vs.* educación ha sido introducido en la literatura por primera vez por R. J. ALLEN y J. S. MILLER, «The Common Law Theory of Experts: Deference or Education?», *Northwestern University Law Review*, vol. 87, núm. 4 (1993), 1131 y ss.

³⁹ De distinta opinión es Carmen VÁZQUEZ, que aboga por la educación solo para los testimonios de los peritos nombrados por las partes pero defiende un modelo deferencial para los testimonios del perito elegido por el juez. *Vid.* Su valioso estudio *De la prueba científica a la prueba pericial*, Madrid-Barcelona, Marcial Pons, 2015.

⁴⁰ En su informe de 2009, *The Law Commission* propone 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, citado anteriormente).

⁴¹ «Una posibilidad podría ser establecer un órgano, 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 [...] 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,

nizados por el propio poder judicial): formación en la naturaleza de la ciencia y en el alcance de sus métodos, sobre todo de los más novedosos; en las principales técnicas forenses, o por lo menos en las más habituales; en la terminología técnica empleada; o en el significado de los datos estadísticos que arrojan la mayoría de estas pruebas. En relación con esto último (la formación de los jueces) se objeta con frecuencia que no es realista instar a los jueces a adquirir un adiestramiento como el comentado. Pero tal objeción resulta, en mi opinión, un tanto apresurada. La finalidad de la formación no es (porque no puede ser) convertir a los jueces en expertos, sino tan solo ayudarles a entender los presupuestos y el alcance de las pruebas realizadas y a controlar en consecuencia si todo ha funcionado bien. Este objetivo es en cierto modo modesto, y no se ve por qué razón habría que infravalorar la capacidad de los jueces para alcanzarlo.

Naturalmente la educación, por sí sola, no es la panacea para asegurar la calidad de la decisión, pues cabe suponer que la educación siempre será imperfecta y en todo caso es difícil mantener a los jueces permanentemente educados: hay demasiados campos de experticia potencialmente relevantes y su número crece sin parar. Y naturalmente la educación no es el único modo de lidiar con los problemas que plantea la entrada de la experticia en el proceso. Susan HAACK, por ejemplo, sugiere como más efectivo concentrarse en lo que sucede «antes» con el fin de evitar que se produzcan cosas indebidas (fraudes en los laboratorios, malas prácticas en la investigación y en la promoción de una técnica, expertos incompetentes, sesgos evitables, etc.) y de que, si se producen, queden rápidamente al descubierto⁴². En definitiva, que es mejor prevenir que curar. No puedo estar más de acuerdo con ella pero aún así la educación resulta necesaria. Sin ella existirá siempre el riesgo de aceptar como conocimiento sólido lo que en rigor tiene escaso fundamento, o de terminar haciéndoles decir a los datos arrojados por las pruebas lo que no dicen ni pueden decir, con lo que la justicia de la decisión puede quedar comprometida⁴³. Sin educación la base cognoscitiva de la decisión judicial se debilita y el riesgo de error se hace más fuerte.

o cuando pareciera adecuado abandonar una forma de prueba que hubiera quedado obsoleta» (Ch. CHAMPOD y J. VUILLE, *Scientific Evidence in Europe. Admissibility, Appraisal and Equality of Arms*, *op. cit.*, 109).

⁴² S. HAACK, «The Expert Witness: Lessons from the U.S. Experience», *op. cit.*, 69.

⁴³ Cfr. por ejemplo V. GRACE, G. MIDGLEY, J. VETH y A. AHURIRI-DRISCOLL, *Forensic DNA Evidence on Trial: Science and Uncertainty in the Courtroom*, Isce Publishing, 2011, que abogan por que los jueces y los jurados estén correctamente informados de (o educados en) los inevitables márgenes de incertidumbre en los resultados de una prueba científica.